A Spatial Analysis of Innovation in Europe

Tonnerre, Antoine

10 July 2017

Spatial Analysis of Innovation Europe: an Empirical Study

Antoine Tonnerre[1]

Abstract

The purpose of this paper is two-fold. Firstly, it surveys different works related to the analysis of Innovation and the way it spreads locally. We will see that theoretical predictions and empirical facts are rather contradictory. This will lead us, secondly, to shedding light towards the concentration of innovative activities in the European Union, at a more precise level of analysis: the NUTS2 nomenclature. We contradict the usually accepted empirical fact that innovative activities tend to be less and less concentrated, thus supporting the theory of knowledge spillovers, which is probably due to the important changes the European Union went through in terms of trade structure. This claim relies on a rather simple but well-founded analysis. Patent applications to the European Patent Office will be used as a proxy for the level of Innovation. These patent applications have a geographical component that will be exploited to determine whether spatial autocorrelation of innovative activities is present. After showing the importance of considering spatial autocorrelation in an analysis of Innovation, we will propose an empirical methodology to assess how it spreads. Regarding this, no results are presented, as it goes beyond the scope of the paper.

Key words: *innovation, knowledge spillovers, spatial concentration, spatial autocorrelation, empirics, patents, European Union, NUTS2*.

JEL codes: R12, O31.

I. Introduction

Innovation is a subject of great matter, as it influences various aspects of economic and social activities. First, it impacts on the consumers as it brings new goods and services that make their everyday-life easier and better. Secondly, Innovation is often considered as one of the main growth engines: the Endogenous Growth theory, with founding articles from authors such as Romer (1986), Lucas (1988) and many others, internalizes the effects of innovation, and more broadly of human capital and knowledge since the end of the 1980s. Such theory implies that policy-makers should support Innovation to recover a strong growth from the 2008 crisis. Thirdly, Innovation can be considered as one of the most efficient investment. Indeed, in *Rates of return to investment in science and innovation* from Frontier Economics (2014), the authors explain that average private returns are 30% and average social returns are 90% for such investments. This last figure is explained by the numerous spillovers that innovation generates across firms, industries, regions, etc. Finally, it is trough Innovation that the most challenging social issues for the future will be solved: aging population, climate change, resource exhaustion, and others, will require high levels of innovation, both in the idea that will solve it (new approaches, new methods, etc.) and in the application of the idea (new technologies, new business processes, etc.). Supporting innovation is a very effective way to reduce unemployment as well.

We henceforth bring some justifications and comments on why focusing our analysis on the area of the European Union. An evident reason is its geographical characteristics. As explained by Acemoglu et al. (2011), Europe is a rather homogenous area. It is wide enough to observe spatial phenomenon, but small enough for the analysis to stay coherent. Moreover, the countries and regions that compose the European Union

[1] Affiliations: Paris School of Economics, Paris, France / Novosibirsk State University, Novosibirsk, Russia. Contacts : antoine.tonnerre@gmail.com / 22 rue des Boulangers, 75005 Paris, France.

are relatively close in their level of GDP, of development, etc. They thus form a coherent area in the sense that the differences observed within the European Union will not be outliers. Additionally, while Eurostat provides unified statistic measurements for the whole area, different policies are carried out in the different regions. This characteristic should make the analysis both convenient and reliable at the same time, since different levels of Innovation and of its determinants should be observed, while no measurement distortions should be present. Furthermore, we are midway through the Europe 2020 strategy [2], which introduced the Innovation Union. It is interesting to observe if it had any impact during its first half of existence, to try to predict its ending in 2020: by looking at whether or not the prediction realized today for the level of Innovation at the end of the Europe 2020 strategy will turn out to be coherent, we will be able to state that the spatial relationships observed do or do not have an impact and that the model will be correct or not. Lastly, we remind that the European Union is constantly getting closer to the technology frontier, as explained by Zilibotti (2011), and this has important implications in terms of innovation policy: "*The closer an economy is to the world technology frontier, the higher the relative importance of innovation relative to imitation as a source of productivity growth.* […] *Closer to the technology frontier, there is less room for copying and adoption of well-established technologies; consequently, there is an equilibrium switch to an innovation-based strategy* […]." (Acemoglu et al., 2006). Indeed, the closer an economy is to the technology frontier, the more necessary it is to push the frontier further in order to create growth.

How a spatial analysis is relevant for studying Innovation in Europe should be explained. It is quite obvious that Innovation is produced at every geographical level: from European industrial groups such as Airbus Group to students creating new software, Innovation is everywhere and at every scale. Localities are parts of bigger entities and can thus communicate and exchange a lot, whether for commercial purpose or not: it is those flows that carry what is necessary for Innovation, such as workforce, resources, technology, knowledge, etc. An easier way to explain the relevance of a spatial study is the clear pattern observed in *Figure 1*. In 2012, a clear concentration of applications to the European Patents Office (EPO) is observed in the area between the east of France, the south of Scandinavia and the United Kingdom, the west of Poland, Czech Republic, Slovakia, Hungary and Croatia, and the north of Italy. The number of applications then spread peripherally around this area. This is a good sign of spatial concentration of Innovation in the European Union. We will later explain how the number of patent applications is a good proxy for the level of Innovation of a given area.

[2] According to the European Commission, the Europe 2020 strategy is "the European Union's ten-year growth strategy", its main priority is "to move beyond the crisis", and one of its pillar is a "smart growth", supported by investments in education, research and innovation. Innovation is one of the five goals expressed by the Commission.

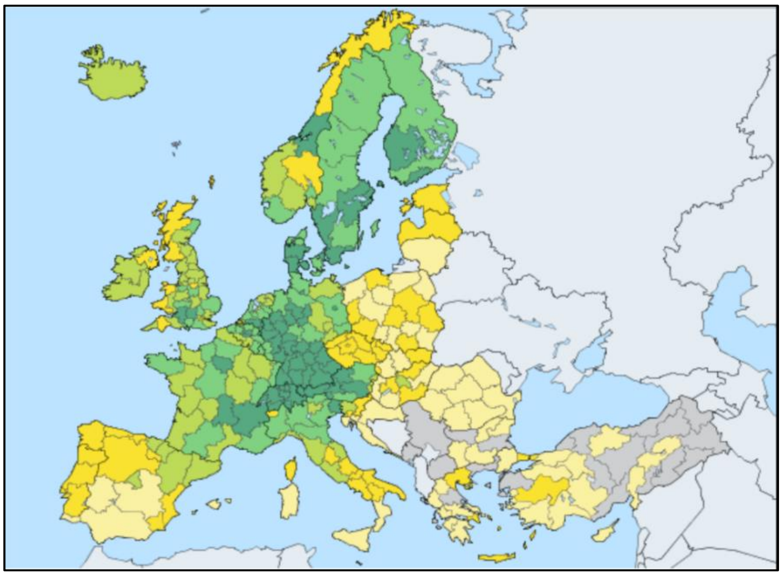


Figure 1. Patents Applications to the European Patents Office in 2012. Dark green is the most applications, light yellow the least. Source: Eurostat.

We finish the introduction of the paper by stating its ultimate goal: to provide to the European policy-makers a description of the spatial Innovation relationships that cause such concentration in the European Union and the keys to try to develop Innovation homogenously across the European Union, because one of themain objectives of the European Union in its economic policy is to implement a coordination toward common goals [3].

We will now first provide a review of the literature. This will give us theoretical explanations of how Innovation can be so concentrated, and a basis for the tools to use. From this will follow some critics and a description of the data available, leading us to a suggestion of what should be done in order to better understand the spread of Innovation in the European Union and of how Innovation could be developed across the European Union at the regional level.

II. Review of the existing literature

1. Location of Innovation

We begin our review with an explanation of why it is relevant to observe Innovation from a geographical point of view. The most important reason is because some of the determinants of Innovation are local. Scott (2006) introduced the main arguments for such point of view. The most obvious concerns the research budgets: research budgets are not same in every region. These research budgets will influence on

[3] See article 119.1 of the Treaty establishing the European Economic Community (EEC)

many characteristics of the creation of Innovation, such as the quality of the local universities, laboratories and infrastructure. Moreover, quality laboratories will only be efficient when occupied by highly-skilled scientists, which also is a local characteristic: the skills and knowledge of the scientists vary across regions. Moreover, they potentially have a certain degree of specialization that will lead them to a specific path of research, thus creating a local path-dependency.

This idea of path-dependency illustrates the way the creative field works: it is not "frozen in time and space", which means that when an innovation is produced at the local level, it will influence back the local field that produced it more than any other field. It can be seen as a web: the creation of innovation takes place inside webs that are located at specific geographical places, so that when an innovation occurs at some place, the web located at this particular place will be modified and evolve in a particular way. This gives a clear understanding of why the level of innovation and knowledge differ across the globe. Finally, the author stresses the crucial role played by the entrepreneur in the process of the creation of Innovation. He is the one that will organize a production, and this production frequently involves the production of ideas and knowledge. The incentives for the entrepreneur to do so are essentially local (business opportunities, observed demand, etc.), such that he will accommodate his behaviour with his location.

In order to go further, Autant-Bernard et al. (2007), the authors controlled for the local characteristics that are independent from the firms' strategies. It can be, for instance, public research budgets. Their discovery was that, even when controlling, the single proximity with another firm have repercussion in the decision to innovate: firms evolving nearby influence each other and, when it comes to innovation, they adopt akin policies. The implication of such results is that purely local peculiarities influence the production of innovation by themselves, in virtue of local knowledge externalities, such as universities, laboratories, or transportation infrastructure: their variety and extent will entail unique rational decisions.

This relates well to the concept of National Innovative Capacity, developed by Porter and Stern (2001), which they define as "*a country's potential, as both a political and economic entity, to produce a stream of commercially relevant innovations*". It is composed of three important notions. Firstly, the notion of Common Innovation Infrastructure, which corresponds to human and financial resources allocated to science and technology at a given geographical scale, public policies and the level of technology in the economy. There are a wide range of public policies: subsidies, intellectual property protection, antitrust enforcement (implemented in order to motivate competition using innovation), etc. Secondly, the notion of Cluster-Specific Environment for Innovation, which assimilates high-quality and specialized inputs, both common investments and competition, sophisticated local demand, and the presence of related industries. Lastly, the quality of the linkages between the two, i.e. the existence of an efficient university system to prevent the knowledge from diffusion abroad (or in other regions). Undoubtedly, the more efficient each of the three points, the greater the National Innovative Capacity. What is interesting for our paper is that this concept could easily be adapted into a Regional Innovative Capacity, because regions do have the opportunity to stimulate each of the three points.

Therefore, as much at the regional than at the national scale, location matters in the sense that each region can build its own Innovative Capacity to try to attract the best scientists, to keep its own, as well as to increase financial inflows for innovation, from public or private investments.

The 2014 Regional Innovation Scoreboard reports a similar observation of the phenomenon of spatial concentration of Innovation in the European Union, which is aware that Innovation is not homogenous across all regions, nor is it random. This means that there are relations at stake, and a need for understanding those relations. Moreover, the phenomenon of spatial concentration is actually increasing, making the need for a policy more and more urgent in order to avoid innovative deserts which would impoverish the concerned regions. This signals a paradox of information technology [4]. Additionally, the paradox of regional innovation[5] keeps on going, as most of the regions that receive important financial support to innovate stay the less innovative. Similarly, regions with the same Innovative Capacity do not necessarily have the same actual level of Innovation. As decentralization is increasing in the European Union, this inconsistency between the Innovative Capacity and the actual level of Innovation will be more and more important, because policy-makers at the regional level will have to be more aware of the relationships of innovation production at the regional level. Once again, the scoreboard insists on the impact of the positive externalities of the location in the production of innovation, with four main points: heterogeneity of the inputs, efficiency of the innovation system, tacit knowledge (which is spatially sticky, this is about the way Innovation spread and it is a topic we will address in the next part) and closeness of firms (which brings reductions of transaction costs, of transport costs, etc.). Close firms also can be incited to form co-colocations to reduce their supply costs by shortening their supply chains and to increase their access potential to talents.

2. *Movement of Innovation*

We showed that location is a very important characteristic in the production of Innovation. We will now focus on how Innovation diffuses itself. Following from Autant-Bernard et al. (2007), there are other explanations to such correlation between close firms. According to the authors, there also exist underlying and impalpable elements: this is what they qualify as a "local state of mind", that the local community experiences. A parallel can be drawn with the work of Vazquez-Barquero (2003), who stressed that "*a strong local identity along with social recognition of the business activity can account for the birth of productive activity and its*

[4] The *paradox of information technology*, or *productivity paradox*, was defined as the "discrepancy between measures of investment in information technology and measures of output at the national level." (Wetherbe, Turban, Leidner and McLean (2007)). It comes from Brynjolfsson (1993).

[5] "The regional innovation paradox refers to the apparent contradiction between the comparatively greater need to spend on innovation in lagging regions and their relatively lower capacity to absorb public funds earmarked for the promotion of innovation and to invest in innovation related activities compared to more advanced regions." (Oughton et al. (2002)).

continuation even in times of need and high risk. Trust among entrepreneurs favours cooperation […]. *Finally, the work ethic contributes to upgrade human resource skills,* […]". He also maintains that urban development and cities encourages all of the determinants of Innovation, such as its diffusion, entrepreneurship, skilled human resources (because of the networks), the concentration of factors, the infrastructures, and the historical and cultural heritage. The "local state of mind" gives explanation on why the predicted death of distance [6] did not occurred, and probably will not. According to Jacks (2009), it can indeed be seen that the role of distance as an obstacle to trade kept stable or increased during the recent years. However, from the work of Li and Minondo (2014), the opposite can be observed for knowledge spillovers among scholars. Ultimately, the total effect on Innovation must balance. As Autant-Bernard et al. (2007) propose, this is most probably due to the fact that telecommunications do not have the ability to transport the underlying elements that constitutes the local state of mind. Following this logic, the concentration of innovation would be caused by the absence unconstrained networks of knowledge and relation. This constriction is the result of imperfect factor mobility, both capital and labour. Finally, they make the distinction between knowledge spillovers, as the environment that surrounds the decision maker, and spatial mimicry, as the fact of being influenced by the decision marker nearby.

Innovation is a phenomenon that spreads rather than simply appear randomly in a given area, at the microeconomic level. Acemoglu et al. (2011) explained that when the signal quality is different between different firms, delays in the spread of Innovation differs. This is due to the way information circulates differently among the firms. Indeed, information about successful experimentations, which themselves motivate further experimentations, do not spread similarly to each firms. Formally, it can be represented as a function that defines the time at which a given firm innovates according to the quality of the signal, such that firms receiving a better signal will create more Innovation.

If we relate to the graph theory, we expect information to move faster between close nodes than between distant nodes. The shape of the graph does not have to be given, it only has to be stated that, whatever the way of transmitting the information (there exists as many graphs as there are tools for communication, i.e. internet, publications in scientific journals, seminars, student exchanges, meeting at different spatial scales, etc.), we can easily imagine a final, weighted abstract graph that will sum up all of the different interactions between actors in the Innovation domain according to every possible way of communication. However, our focus will be on one important determinant of the quality of the signal: distance. According to graph theory, we define distance as the shortest path to connect two nodes: some paths almost are independent from distance (internet, even though some information only is accessible according to where one tries to access the

[6] "*The death of distance* is a phrase coined by Frances Cairncross, which means that any activity which relies on electronic devices for information or communication handling, can be carried out anywhere in the world." (University College London, Urban Transformation in the Information Age).

information), others are greatly dependent (it is much easier to plan a scientific seminar with scientists from the same cluster than from the whole world).

Factors' mobility also is a prominent determinant. It can take the form of technological transfers for the capital, or mobility of the highly-qualified ones for the labour. For any given level of mobility, we can assume that the neighbouring firms and regions will have the most exchanges.

One last point to make is about the different ways of qualifying distance. Distance can be though as non-geographical: we can imagine concepts such as economic distance (differences in GDP between regions, in profits between firms, etc.), knowledge distance (differences in education between regions, in highly-qualified labour force between firms, etc.), or market potential in a perimeter of a given dimension (reachable GDP in less than a 100 kilometres, etc.). We will thus try to explain the level of Innovation given a certain notion of distance, with a focus on the geographical distance. We will, however, propose later an alternative way of measuring distance. For now, the way of measuring Innovation should be explained.

3. *Patents as a proxy for the level of Innovation*

		Others	
		Innovate	Do not innovate
Individual	Innovate	5 ; 5	-5 ; 10
	Do not innovate	10 ; -5	0 ; 0

Figure 2. Representation of the problem of producing innovation in the absence of patents.

In the absence of patents, there is a market failure in the process of producing innovation. In terms of Game Theory, it can be represented in the form of a prisoner's dilemma as in *Figure 2*. The process of innovating is indeed very costly, and since there are spillovers in the production of innovation, each individual has much interest in free-riding, because if someone else innovates, one will still benefit from the innovation without having participated. As in the classical prisoner's dilemma game, the solution of the game is that no one innovates. In *Figure 2*, one would replace the pay-offs according to the potential benefits from innovating.

Moreover, Laginier and Moschini (2002) explained that the nature itself of a patent must bring innovation: patents must be novel, useful, non-obvious and will be disclosed. Consequently, trying to tend to monopoly by seeking for patents will in fact stimulate social benefits. This is the main reason why it is the object of our study: we can reasonably assume that the number of patents delivered is a great representation of the level of innovation of the region studied, because of the requirement they must fulfil in order to be delivered. However, as the authors suggest, we must keep in mind a certain number of inherent limits to using patents as tools for estimating the level of innovation. First of all, in almost every cases, patents concern innovations with market potential. Secondly, a firm can decide to not patent its innovation in order to keep it secret, when the cost of the disclosure outweighs the benefit of acquiring a position of monopoly for a given time. Finally, some of the biggest discoveries where not patentable. Indeed, ideas, mathematical methods or scientific theories belongs to the innovations that cannot be patented.

Nevertheless, we focus in this paper in the creation of economic wealth, which is the main drive for economic agents to research and create innovative processes. Another reason why we focus on the production of patents is because they are the only way to manage the production of knowledge using the market, as developed by Arrow (1962). In this way, we can try to understand how the market can create knowledge by understanding the production of patents.

4. A similar approach to the paper

We end our review with Moreno et al. (2003), who proposed an empirical study of Innovation in Europe. Their finding is that there exists a spatial dependence in Europe regarding the production of Innovation. This dependence is axed around a central-periphery distribution, with its centre in North and Centre Europe. However, they find that this concentration tends to decrease over time, which contradicts the Regional Innovation Scoreboard 2014. Then, they propose a spatial model which allows them to conclude that, in Europe, Innovation mostly spreads between regions of the same country, i.e. the national effect of spatial dependence dominates the European effect. Clusters of innovation are spotted, the most important being located in West Germany.

III. Remarks and critics

A number of remarks can be formulated towards the article from Moreno et al. (2003).

As it is an empirical study, the results are only certain for the data used. This is even more true in Economics as policies are designed to influence the data used as explained variable. In this article, the most recent data is almost twenty years old, as the period studied covers the years from 1978 up to 1997. Despite the meticulousness of the methodology, because of the data we can reasonably question the validity, at the present time, of the results proposed, even more when the recent Regional Innovation Scoreboard 2014 claimed an increase in the spatial concentration, contrarily to Moreno et al. (2003).

To support this hesitation, one can look at the history of the European Union, which in the recent time grandly evolved. Chronologically, it can be remarked that approximately 80% of the period covered occurred before the signature of the Maastricht Treaty [7] in 1992. This treaty includes free circulation and residence of the citizens inside the European Union. One could rationally expect a greater mobility of the labour force within the European Union from such treaty. It also includes a destruction of the barriers to the mobility of commodities, services and capitals. Such measure most probably allowed technological transfers within the European Union. Three other major evolutions of the European Union took place after the covered period. First of all, the Amsterdam Treaty [8], which consists of a revision of the Maastricht Treaty, included the possibility to joint national programs in scientific research and innovation. It is quite obvious that such measure can definitely impact on the relationships at stake in the production of Innovation and the spatial dependence of regions. We can already wonder why it did not reduce the concentration of Innovation in the European Union. Secondly, in 2002 was introduced the unique currency, the euro, which, according to Baldwin et al. (2008), increased the exchanges inside the European Union. Once again, this theoretically should have increased the volume of technological transfers. Finally, the Lisbon Treaty [9] was ratified in 2009 and amended the Maastricht Treaty. It gave to the European Union the power to lead science and innovation programs as an independent entity. One expected policy from the European Union would then be to try to harmonize the level of Innovation in the European Union by leading programs soliciting regions with low levels of Innovation. This stimulation could, again, greatly redefine the way Innovation emerges and spreads in the European Union.

Another argument that favours a critical approach is the countries used to realize the study. Before going further, it is important to keep in mind the following dates, from the beginning of the covered period until today: in 1981 entered into force the Athens Treaty 1979, for the adhesion of Greece; in 1986 entered into force the Madrid/Lisbon Treaty, for the adhesion of Spain and Portugal; in 1995 entered into force the Treaty of Accession 1994, for the adhesion of Austria, Finland and Sweden; in 2004 entered into force the Athens Treaty 2003, for the adhesion of Malta, Poland, Lithuania, Czech Republic, Latvia, Hungary, Estonia and Slovenia; in 2007 entered into force the Luxembourg Treaty, for the adhesion of Bulgaria and Romania; finally, in 2013 entered into force the Brussels Treaty, for the adhesion of Croatia. One last remark can be formulated about the enlargement of the European Union after the end of the period covered in the article. Between the beginning of the covered period and the end of the covered period, 6 countries joined the European Union. Between the end of the covered period and today, 13 countries joined the European Union. In total 19 countries were added to the Union.

---

[7] One can access the Maastricht Treaty, or *Treaty on European Union*, in its completeness using the following url: http://europa.eu/eu-law/decision-making/treaties/pdf/treaty_on_european_union/treaty_on_european_union_en.pdf

[8] One can access the Amsterdam Treaty, or Treaty of Amsterdam amending the Treaty of the European Union, the Treaties establishing the European Communities and certain related acts, in its completeness using the following url: http://www.europarl.europa.eu/topics/treaty/pdf/amst-en.pdf

[9] One can access the Lisbon Treaty, or *Reform Treaty*, in its completeness using the following url: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:C:2007:306:FULL:EN:PDF

We now propose some geographical critics. The first one is that the area covered did not took into consideration the frontier of the European Union, whose existence could have largely influenced the relations between the regions within the European Union. The article does only consider the European area (which is different from the European Union area), but one will find it difficult to properly decide which countries to include or not in that case. Including all countries of Scandinavia but not Czech Republic or Poland for instance is open to debate. Indeed, there was no common policy between all the countries studied in the article. Using the frame of the European Union allows to a have a clear definition and to give evident reasons to include or not a country in the study, motivated by the common market and policy. The second one is about the use made of the NUTS. This critic also can be made to the Regional Innovation Scoreboard. Both use a mixture of different levels of NUTS. Moreno et al. (2003) use NUTS 0, 1 and 2 to create harmonious regional areas, because they considered at the time (and certainly with reason, because the NUTS was mostly considered as a gentlemen's agreement until 2003) that the nomenclature was artificial. Since then, however, Eurostat has worked to build a nomenclature that, we believe, gives a correct and coherent division regarding the economic and administrative controls allowed to each region. To support that, we remind that the NUTS obtained juridical status in 2003 [10] and was multiple times amended to tend to a better geographical division. Additionally, in an attempt to provide comparable studies, using an official nomenclature is more appropriate, because future studies will not have to manage a complete redefinition of the geographical division in order to compare the results. A last remark will concern the way distance is measured. Since no precision is given, we assume that distance is calculated "as the crow flies". However, such notion of distance could be reductive in the analysis of innovation. There might be a more realistic representation of distance regarding the interactions between the economic agents. The area of the European Union is shaped by seas, mountains, etc., that could alter the interactions proposed by Moreno et al. Moreover, distance "as the crow flies" may be prone to spherical distortions on areas as wide as the European Union. Lastly, this distance does take not into consideration the quality and density of the transportation infrastructure. One regions could be close to another as the crow flies, but distant in reality if its access is complicated. Inversely, one region could be far as the crow flies, but close in reality if they can be joined quickly.

To address the remarks made previously and efficiently use the literature reviewed, we will now propose the methodology, and present what we can learn from the research we realized so far.

IV. Methodology and Research

We will adopt a similar approach to the one chosen by Moreno et al. (2003). We want to study spatial dependence of the regions of the European Union regarding the production of Innovation. Therefore, we will first use Ordinary Least Squares (OLS) method, and test its errors. If they reveal spatial autocorrelation, then

[10] For more information on the evolution of the NUTS, please visit: http://ec.europa.eu/eurostat/web/nuts/history

we will not use OLS since it is not efficient. Consequently, we will have to use spatial econometrics techniques. We do not know yet which of the Spatial Lag Model (in the case of lagged value, OLS is not efficient) or Spatial Error Model would be more appropriate. As a reminder, we here give the form of the models:

$Y = \alpha + \rho WY + X\beta + \epsilon$, for the Spatial Lag Model.

$Y = \alpha + X\beta + \lambda Wu + \epsilon$, for the Spatial Error Model.

With W a weight matrix, whose way of calculating will be discussed later. By doing so, we will be able to observe spillover effects and to obtain unbiased estimates.

Regarding the data, we will use the most recent data provided by Eurostat for the most recent regional nomenclature (NUTS 2 level), which gives us 268 individuals, and the period 2008-2012, which gives us 5 time periods. Such period has multiple advantages. First of all, the structure of the European Union has been stable during those years. Secondly, almost all regions provided data for this time period. Thirdly, compared to the number of regions (268>>5), the number of time period is much lower: this allows us to obtain the advantages of working with a panel data while preventing us from serial correlation. Indeed, short panels are by definition not prone to serial correlation. For each region, we have the following data: Patents Applications to the European Patent Office (per million inhabitants), Total intramural R&D expenditure in euro (per inhabitant), Employment in technology and knowledge-intensive sectors, Demography and GDP.

We will be able to transform the data, so that for every category it will be expressed in the same scale.

Henceforth, we provide the first results to determine whether or not we are in presence of spatial autocorrelation. We use Moran's Index for that purpose. We remind that Moran's Index is computed the following way:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} . \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(Y_i - \frac{1}{n}\sum_{i=1}^{n} Y_i)(Y_j - \frac{1}{n}\sum_{i=1}^{n} Y_i)}{\sum_{i=1}^{n}(Y_i - \frac{1}{n}\sum_{i=1}^{n} Y_i)^2}$$

Using softwares *QGis* and *GeoDa*, we extract from the shapefile for the whole NUTS provided by Eurostat the NUTS 2 level. We join our data set to that NUTS 2 shapefile. Using the data for 2012, we computed the Moran's Index with different kind of weight matrices and different parameters.

From Figure 3, Figure 4 and Figure 5, we can observe that in any method, there is a quite strong spatial autocorrelation in the applications for patents to EPO in 2012. The weight matrix used in Figure 3 takes into account the nearest neighbours, with k the number of nearest neighbours. For instance, using *k=3*, the level of Innovation of each region will be associated with the level of Innovation of its 3 closest neighbours when performing the calculations. In Figure 4 are represented the Moran's Indexes computed using the *k* contiguous neighbours of each region. As an illustration, for *k=2*, the level of Innovation of the direct neighbours of each region's direct neighbours will be taken into account. Finally, in Figure 5 the Moran's Indexes are computed using the same principle as the previous one but including the lower degrees: again, for *k=2*, the level of Innovation of the direct neighbours of a given region's direct neighbours, and of the direct neighbours of the given region themselves, will be taken into account. Examples of the matrices are in the Appendix.

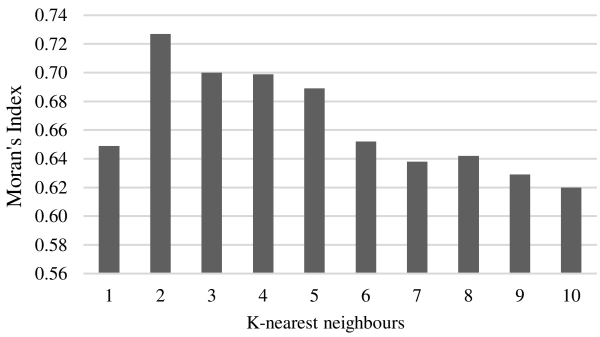


Figure 4. Moran's Indexes with k-nearest's method

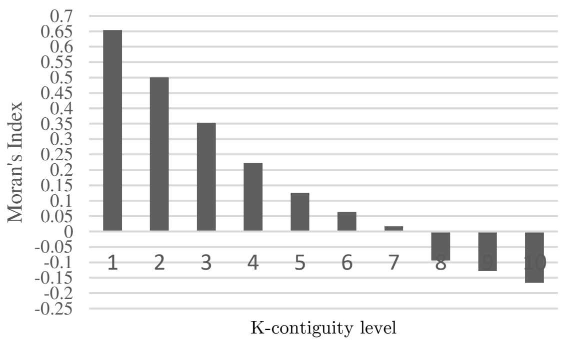


Figure 4. Moran's Indexes with k-contiguity's method

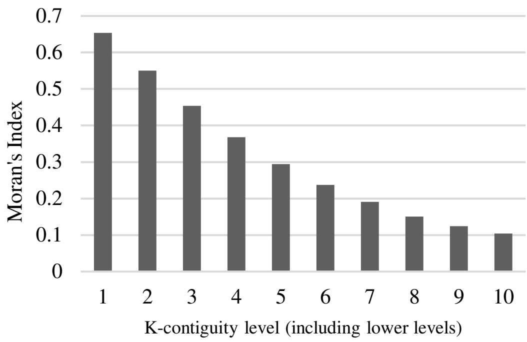


Figure 5. Moran's Indexes with k-contiguity including lower levels'

We notice that the highest spatial autocorrelation appears when using a weight matrix for the two nearest neighbours. In this configuration, the Moran's Index reaches approximately 0.72. Unfortunately, it is impossible to use this weight matrix for further analysis because it is not symmetrical. However, the ones computed with contiguity levels (Figure 4 and Figure 5), including the lower levels, can be used. The highest level of spatial autocorrelation is achieved at first degree of contiguity, which is obviously the same for both methods and ranges at around 0.65, which is still quite high. We will thus use this weight matrix in the first place to provide a deeper analysis of the spatial autocorrelation.

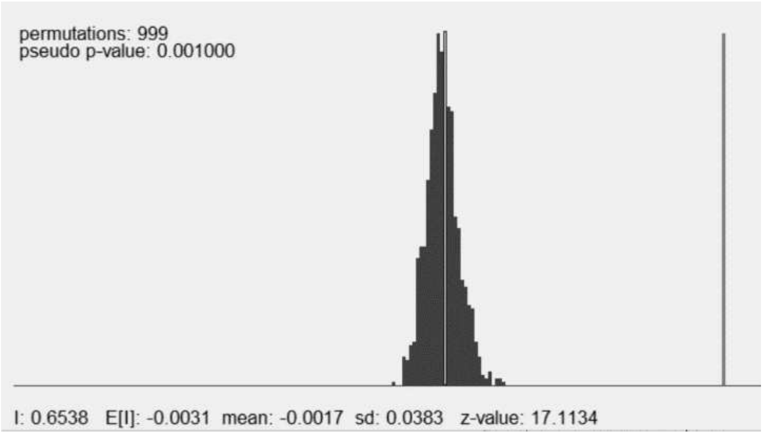


Figure 6. Test for significance of the Moran's Index

Figure 6 was obtained with *GeoDa*. It tests how strong the Moran's index (I) is, using random permutations of the data across all regions covered: the levels of Innovation of each region are randomly assigned to other regions. Here, 999 random permutations are realized. The standard error of the Moran's Index is 0.0383. The mean obtained is the average Moran's Index for the 999 simulations. It is equal to -0.0017. The mean is an estimator of the expected Moran's Index, $E(I) = -\frac{1}{n-1}$, which is equal to -0.0031. The difference between the two gives the sample error of 0.0014. Comparing the Moran's Index of 17.1134 and the expected Moran's Index of -0.0031, we can note that this is clue for spatial autocorrelation. However, we can only be assured that this is not the result of sampling errors, and actually is a feature of the observed variable, by the significance test.). The pseudo p-value is obtained as $p' = \frac{M+1}{S+1}$, with *M* the number of random permutations which obtained a bigger Moran's Index than the one obtained with our actual data, and *S* the number of random permutations. We can conclude that *M=0* here, since *S=999* and *p'=0.001*, which is the lowest possible p-value for 999 permutations: in any of the random permutations of our data did the Moran's Index get as high as the one obtained with the actual data. The Moran's Index obtained thus is high (*I=0.653785*) and very significant (*p'=0.001*).

The z-value is obtained using $z = \frac{I-E(I)}{\sqrt{V(I)}}$, according to Anselin (1995), with $V(I) = E(I^2) - E(I)^2$. The z-value is equal to 17.1134, which is higher than the ones obtained by Moreno et al. (2003): at contiguity *k=1*, they obtain z-values of 8.083 for 1981-1983, 9.734 for 1998-1990 and 10.022 for 1995-1997. We provide the Moran's scatter plot for the data for 2012 with the weight matrix for contiguity of degree 1 in Figure 7.

It represents on the horizontal axis the standardized values for the number of applications to the EPO in 2012, and on the vertical axis the spatial lag for.

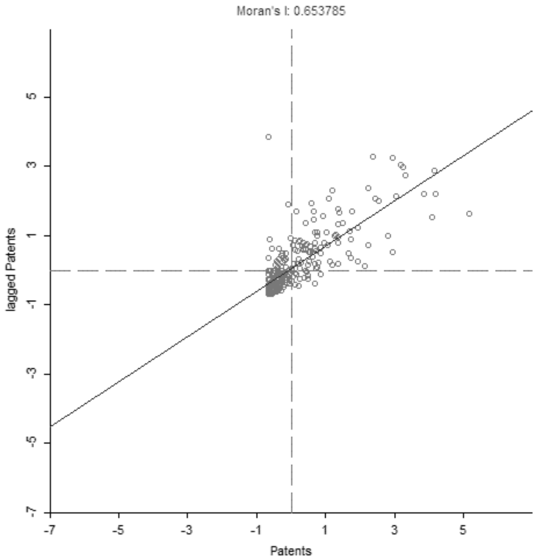


Figure 7. Moran's Scatterplot for the levels of Innovations

We can observe the relation between the number of application to the EPO of a given region and of its direct neighbours: in the north-east quarter are represented regions that positively correlate to their direct neighbours with high levels of Innovation, in the south-west quarter are the regions that positively correlate to their direct neighbours with low levels of Innovation, and in the north-west and south-east quarters are the regions that negatively correlate to their direct neighbours. The slope displayed is the Moran's Index, which represents the global autocorrelation. It is here positive, which means that globally there is positive spatial autocorrelation: regions with high levels of patents applications are often close to regions with high levels of patents applications, and regions with low levels are often close to regions with low levels. This is good sign of clustering.

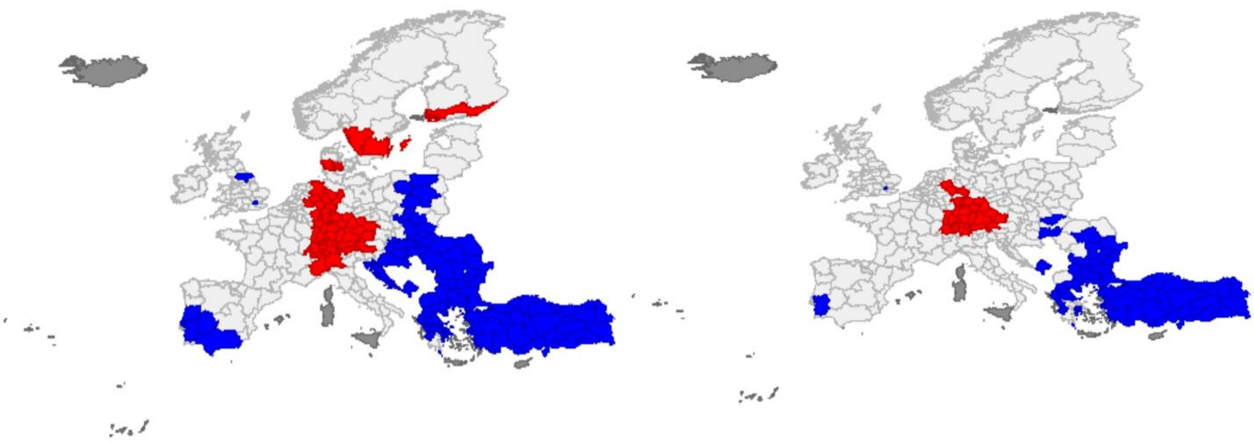


Figure 8. Clustering maps of Innovation in the European Union, at the 5% degree of significance on the left, 1% on the right.

In Figure 8 are represented clustering maps realized with *GeoDa*. The red colour indicates positive clusters and the blue colour negative clusters. At the 5% degree of significance, the cluster of innovation that was previously located in west Germany has extended. Moreover, new clusters seem to have appeared in Northern Europe. This could be explained by successful innovation policies in the countries of Scandinavia (the most of the population of these countries is located in the south). At the 1% degree of significance, only the cluster located around Germany persists. It seems that the spatial concentration has increased, as it appears that clustering is greater than the one found by Moreno et al. (2003). It also seems that the cluster's got concentrated in its South compared to the one obtained by Moreno et al. (2003)[11], which located in the West of Germany.

To confirm the presence of spatial autocorrelation in the European Union regarding Innovation, some "pre" regressions have been run. The absolute values that follows should not be considered, however, one can observe an improvement in the well-known statistics. The most efficient model was the Spatial Lag Model. We can compare it with an OLS regression.

SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION

Dependent Variable	:	Patents	Number of Observations	:	320
Mean dependent var	:	38.1105	Number of Variables	:	4
S.D. dependent var	:	55.6276	Degrees of Freedom	:	316
Lag coeff. (Rho)	:	0.694237			
R-squared	:	0.653878	Log likelihood	:	-1592.37
Sq. Correlation	:	-	Akaike info criterion	:	3192.73
Sigma-square	:	1071.05	Schwarz criterion	:	3207.81
S.E of regression	:	32.7269			

Variable	Coefficient	Std.Error	z-value	Probability
W_Patents	0.6942365	0.0391988	17.71066	0.00000
CONSTANT	2.412011	2.870913	0.8401546	0.40082
Employment	0.09834249	0.01580624	6.221752	0.00000
Demography	-3.656976e-005	6.570812e-006	-5.565485	0.00000

SUMMARY OF OUTPUT: ORDINARY LEAST SQUARES ESTIMATION

Dependent Variable	:	Patents	Number of Observations	:	320
Mean dependent var	:	38.1105	Number of Variables	:	3
S.D. dependent var	:	55.6276	Degrees of Freedom	:	317
R-squared	:	0.240714	F-statistic	:	50.2487
Adjusted R-squared	:	0.235923	Prob(F-statistic)	:	1.10728e-019
Sum squared residual	:	751859	Log likelihood	:	-1695.98
Sigma-square	:	2371.8	Akaike info criterion	:	3397.96
S.E. of regression	:	48.7011	Schwarz criterion	:	3409.26
Sigma-square ML	:	2349.56			
S.E of regression ML	:	48.4723			

Variable	Coefficient	Std.Error	t-Statistic	Probability
CONSTANT	22.09111	3.920095	5.63535	0.00000
Employment	0.2108826	0.02245637	9.39077	0.00000
Demography	-8.059174e-005	9.400643e-006	-8.573003	0.00000

In the Spatial Lag Model, the coefficient of determination improved greatly, the standard error reduced significantly, and both Akaike and Schwarz criterion, even if they remain high, diminished.

From all those results, we can affirm that spatial autocorrelation is present in the European Union regarding the production of Innovation at the regional level. We must now model it in order to better understand the relations at stake. As a response to the previous critics we made, we would like to build an ideal weight matrix composed of indexes of distance based on all kind of available transportations weighted by their cost, degree of availability, etc. Unfortunately, such data is not available for now. However, according to the Service for Observation and Statistics of the French Ministry for Environment, Energy and Sea, road transportation represents 90% of total transportation in Europe. Moreover, its share is constantly increasing. Consequently,

[11] See Appendix

we could consider driving distance as a great proxy for this ideal index. Using driving distance, we even make sure to get rid of the problems of spherical distortions, as distance is calculated on the road. To do so, we use the NUTS 2 shapefile we previously extracted from the official NUTS shapefile by means of *R*. In *R*, we use the package *maptools* to first obtain the GPS coordinates of all 268 centroids of the regions. One regions' centroid is computed as the mean position of each point forming the shape of the region. We can then use those GPS coordinates with *Open Source Routing Machine* (OSRM), which is an open-source project that gives the shortest road distance (including ferry transportation on short and widely used sea paths, such as between Italy and Albania). It is possible to automatize OSRM calculations for every shortest route between each centroid, using the package *osrm* in *R*. By doing so, and after some manipulations of the data obtained, we have in our possession a 268 by 268 matrix, with 0 diagonal, containing the 71824 shortest driving distance. Finally, we inverse the matrix and set the diagonal as 0 in order to give more weight in the regression to the closer centroids.

We will now propose what should be done next. It would be interesting to obtain even more recent data on the explained and explanatory variables. We hope that by some little time, Eurostat will have published some. We should use that newly built weight matrix in a regression model. It would be interesting to compute different weight matrices, such as for market potential. Moreover, the data, without modification, has been tested for normality and did not passed it. A power transform using natural logarithm was, however, quite effective. After that, we will make the final regressions in different configurations, and compare the results obtained with different weight matrices. We will provide visual support and analyze the results in order to understand the spatial relations between regions when producing Innovation. Empirical and theoretical explanations will have to be given. A comparison with the work of Moreno et al. (2003) will determine whether or not the recent policies and evolutions of the European Union have had an impact. This will, ideally, provide a framework for the future research on this topic.

V. Conclusion

In this paper, we tried to shed some light on the spatial correlation between regions of the European Union in the recent years when producing Innovation.

The stakes of this topic were presented in three points: Innovation, the European Union and the spatial analysis. We started our study from a simple observation of the apparent concentration of patent applications to the EPO in 2012. From that, we reviewed the existing literature on Innovation, on its emergence, on its spread and on the tools that we can use to explore it. The article from Moreno et al. (2003) was introduced as the basis on which our work will rely on. We then formulated some critics and remarks, mostly oriented toward the recency of the already existing work, which appears critical for an empirical study. We also observed some contradiction between the older and more recent research. To address these critics and remarks, we proposed a methodology and some primary results, which confirmed the high and significant positive spatial correlation of the production of Innovation in the European Union. Before being represented in clustering maps in order to better visualize the patterns, these results were explained and interpreted. The z-score obtained was higher than the ones obtained by Moreno et al. (2003), suggesting a more important spatial autocorrelation than before. The maps, with different significance levels, revealed that concentration at least did not decreased and at most increased. We thus note both a higher concentration of the Innovation activity in the European Union, and a greater spatial correlation. We then verify in a very basic model the positive impact of the spatial model to explain the levels of Innovation on the main indicators. A first step into the following of this research was the construction of a particular weight matrix based on driving distance. Finally, some propositions were made regarding the future of this research work.

VI. Appendix

Weight matrix used in Figure 4

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & 1 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ 1 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 & 1 \\ \vdots & \ddots & \ddots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 & 0 \end{bmatrix}$$

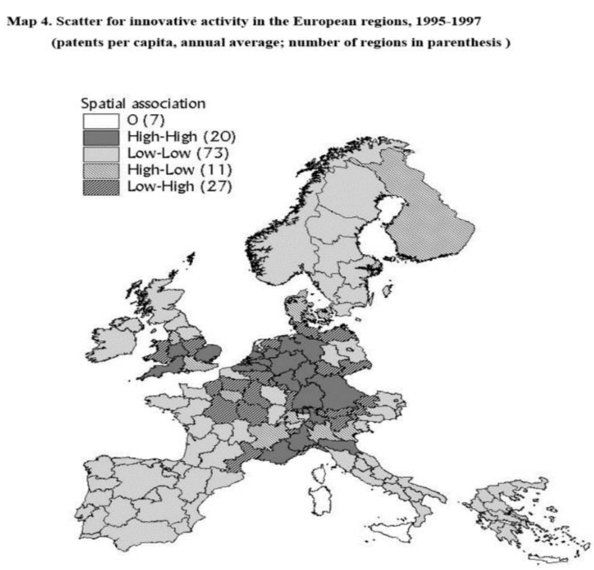
k=1 k=2

Weight matrix used in Figure 5

$$\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 & 0 \\ \vdots & \ddots & \ddots & 1 & 0 & 1 \\ 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \ddots & \ddots & \vdots \\ 1 & 1 & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 1 & 1 \\ \vdots & \ddots & \ddots & 1 & 0 & 1 \\ 0 & \cdots & 0 & 1 & 1 & 0 \end{bmatrix}$$

k=1 k=2

Clustering map from Moreno et al. (2003)



VII. Figures

VIII. References

Acemoglu, D., Aghion, P., & Zilibotti, F. (2006, March). Distance to Frontier, Selection, and Economic Growth. *Journal of the European Economic Association*, pp. 37-74.

Acemoglu, D., Bimpikis, K., & Ozdaglar, A. (2011, February). Experimentation, Patents, and Innovation. *American Economic Journal: Microeconomics, Vol. 3*, pp. 37-77.

Anselin, L. (1995, April). Local Indicators of Spatial Association. *Geographical Analysis, Volume 27, Issue 2*, pp. 93-115.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Inventinos. *The Rate and Direction of Inventive Activity: Economic and Social Factors*, pp. 609-626.

Autant-Bernard, C., Lesage, J. P., & Parent, O. (2007, December). Firm Innovation Strategies: A Spatial Cohort Multinomial Probit Approach. *Annales d'Economie et de Statistique, N° 87/88, Spatial Econometrics, Innovative Networks and Growth*, pp. 63-80.

Baldwin, R., DiNino, V., Fontagné, L., Santis, R. A., & Taglioni, D. (May 2008). *Study on the Impact of the Euro on Trade and Foreign Direct Investment.* Economic and Financial Affairs Directorate-General of the European Commission.

Brynjolfsson, E. (1993). The Productivity Paradox of Information Technology. *Communications of the ACM*, pp. 66-77.

Frontier Economics Ltd. (2014, July). *Rates of return to investment in science and innovation.* Prepared for the Department for Business, Innovation and Skills (bis), London.

Hollanders, H., Es-Sadki, N., Buligescu, B., Leon, L. R., Griniece, E., & Roman, L. (2014). *Regional Innovation Scoreboard.* European Union.

Jacks, D. (2009, September 12). *Revisiting the Death of Distance*. Retrieved from VOX: http://voxeu.org/article/revisiting-death-distance

Langinier, C., & Moschini, G. (2002, February). The Economics of Patents: An Overview. *Working Paper 02-WP 293, Center for Agricultural and Rural Development*.

Li, Y. A., & Minondo, A. (2014, February 25). Has Distance Died with Networks? - An Empirical Study of Scholarly Citations and Genealogy in Knowledge Spillovers. *HKUST; University of Deusto*.

Lucas, R. E. (1988, February). Mechanics of Economic Development. *Journal of Monetary Economics*.

Moreno, R., Paci, R., & Usai, S. (2003, December). Spatial Spillovers and Innovation Activity in European Regions. *WP CRENoS*.

Oughton, C., Landabaso, M., & Morgan, K. (2002). The Regional Innovation Paradox: Innovation Policy and Industrial Policy. *Journal of Technology Transfer*.

Porter, M. E., & Stern, S. (2001, Summer). Innovation: Location Matters. *MIT Sloan Management Review*.

Romer, P. M. (1986, October). Increasing Returns and Long-Run Growth. *The Journal of Political Economy, Vol. 94*, pp. 1002-1037.

Scott, A. J. (2006, February). Entrepreneurship, Innovation and Industrial Development: Geography and the Creative Field Revisited. *Small Business Economics*, pp. 1-24.

Vazquez-Barquero, A. (2003). Chapter 2: An approach for action. *Endogenous Development: Networking, Innovation, Institutions and Cities*, 21-35.

Wetherbe, J. C., Turban, E., Leidner, D. E., & McLean, E. R. (2007). Information Technology for Management: Transforming Organizations in the Digital Economy (6th ed.). New York: Wiley.

Zilibotti, F. (n.d.). Distance to technology frontier and European economic growth.