Cronfa - Swansea University Open Access Repository

_____

_____

Paper:

_____

CrossMark

Six degrees of cultural diversity and R&D output efficiency

Cultural percolation of new ideas: an illustrative analysis of Europe

Annie Tubadji[1,2] · Peter Nijkamp[3,4]

Abstract This paper seeks to highlight the efficiency of R&D output as a function of a cultural treatment (i.e. exposure) effect. The focus of our research is on the percolation of new R&D ideas from the immaterial world of creative ideas through the cultural lattice of the locality into the documented world of knowledge. Our conceptual model is illustrated with a novel numerical operationalization of the cultural percolation of ideas hypothesis, based on the six degrees of separation literature and using a dataset compiled from EUROSTAT and the European Social Survey. The estimation strategy for the current work relies on a difference-in-differences method from a network percolation approach, with a series of alternative controls and region-fixed effects. The results show a positive significant role of the stability ('no change') of the six degrees of cultural diversity (i.e., the likelihood to have six people in a row in a locality originating from a culturally different origin) as a treatment effect for R&D output efficiency (the latter being measured as the number of new ideas over a millions of euros of R&D investments). The main value added of the paper is that it offers a theoretical justification and numerical illustration on how the six degrees of separation paradigm can be used to approximate the tipping point of the percolation of new ideas

✉ Peter Nijkamp
p.nijkamp@vu.nl

Annie Tubadji
atubadji@hotmail.com

1 Department of Economics, University of Bologna, Campus Rimini,
   Via Anghera 22, 47921 Rimini, Italy

2 University of the Aegean, Chios, Greece

3 Department of Spatial Economics, VU University Amsterdam, Room 9A-76,
   De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

4 Adam Mickiewicz University, Poznan, Poland

🌱 Springer

through the local social network from the pool of ideas to efficient R&D investment decisions.

1 Introduction

Innovation through R&D is recognized in endogenous growth theory as a main driver of economic growth (see, for example, Barro and Sala-i-Martin 1995; Bresnahan and Trajtenberg 1995). While this theoretical claim has been backed by an abundance of solid empirical evidence, the same strand of research also indicates that the relationship between innovation and growth is rather heterogeneous in nature (Romer 1986, 1990a, b, 1994; Grossman and Helpman 1993; Lichtenberg 1993; Eaton and Kortum 1997). Distinct regions can be identified that were outperforming others in terms of generating a relatively high productivity compared to their level of R&D expenditure, while this efficiency of the utilization of R&D expenditure was often explained on the basis of a socially dependent diffusion rationale (see Davelaar and Nijkamp 1989; Rodriguez-Pose 1999; Bresnahan and Yin 2010). In the present study, we direct the attention to this puzzling issue one step earlier in the diffusion of knowledge process, namely, whether and how the social lattice in a locality will influence the number of R&D—produced new pieces of knowledge that will eventually appear in a locality. This process is the most essential one, as it defines the initial conditions for endogenous economic growth based on R&D innovation.

Taking this premise as a starting point of the current analysis, we refer here to the recent cultural percolation of ideas hypothesis presented by Tubadji and Nijkamp (2015).[1] In line with the cornerstones of the underlying culture-based development (CBD) concept, the authors argue that emergence of an idea from the world of ideas into the material world of documented R&D innovation resembles the physical process of percolation. Percolation is the process of penetration of a liquid through a porous lattice. Analogically, ideas pass through the social lattice in a locality in order to be selected and implemented by a local research agent, and thus to get transformed into R&D output. This filtering process is generally indifferent of the size of the investment, but it can be expected to experience a higher cultural bias with the increase of the R&D investment due to the perceived higher risk of failure (for more details on culture and

---

[1] The percolation model for the selection of new ideas for R&D investment differs from a knowledge diffusion model by the temporal relationship of the model with the choice for R&D investment. The percolation model reflects the moment when an idea is still to be selected from a pool of ideas and faces uncertainty regarding the question whether it will be selected for investment or not. The diffusion models deal with the distribution of the new idea already chosen and invested in through the R&D process. These are two related but clearly distinct processes. Our percolation model focuses on the first stage of percolation of new ideas from the field of possibilities into the real world of R&D investment. For more details see Tubadji and Nijkamp (2015).

cultural bias and risk, see Tubadji 2012, 2013; Batabyal and Nijkamp 2014). Moreover, every percolation is determined by its tipping point (a switching point between a 0 and 1 mode of existence).

The tipping point of each lattice is a function of the structure of the lattice, but even in the natural sciences such as physics and mathematics the tipping point is known only for some relatively simple lattice structures. The exact structure of the entire local social network is very difficult to be captured and is itself unobservable in its entirety. Therefore, its tipping point is practically impossible to even try to estimate. Thus, the only way to approach the cultural percolation model empirically is through an approximation of its tipping point. Some approximations for social networks structure (and not their tipping points) have been tried (see e.g. Zenou 2012) which use the labour market structure to approximate the network structure. Yet, this firstly leads to endogeneity issues if the effect of the network on economic processes related to the labour market is investigated. And second, a very strong assumption for the social network structure has to be made, in particular that it is dependent only on immediate one, two or three connections, for which the tipping point is known and possible to calculate. Instead, we propose here an alternative measure related to a purely cultural paradigm—the six degrees of separation in society—as a possible endogeneity—free approximation for the tipping point of cultural percolation.

It should be noted that existing research on the dissemination of information through a network has generated the well-known concept of the 'six degrees of separation', i.e. the importance of being able to connect a random draw of six people in a row in order to manage to transmit consistently a piece of information from person A to person B who do not know each other (Traverse and Milgram 1969; Granovetter 1973; Watts and Strogatz 1998). We reinterpret this notion as the dependence of efficient network communication on the likelihood of having six degrees of social connectedness (SDSC) between six random members in the network. More precisely, we are interested in the likelihood of having six random people in the social network sharing the same attitude levels (i.e. common shared cultural values). The higher the likelihood of having SDSC, the lower the attitudinal diversity in the locality, but also the higher the speed of percolation of certain type of ideas can be expected to be. If the notion of SDSC is empirically based on a variable related to open-mindedness of the milieu towards new ideas, these SDSC will lead to a more efficient and more productive percolation of ideas, which can be quantified as more original ideas (or more patents). Furthermore, as we know from the homogeneity literature and social networks (see for instance Lozares et al. 2014), the higher the cultural proximity, the higher the 'love' for each other. This means that the very likelihood of cultural diversity among six randomly drawn people in a locality will predict how efficiently they will be able to communicate with each other and how connected they will be in their attitudes (see Akerlof and Kranton 2000). Thus, the six degrees of separation (or social connectedness) can be approximated with the likelihood of six random people in a locality having the same cultural origin. If we use people whose cultural origin is foreign and carrying a sense of diversity, we can expect our six degrees measure to be positive related with more innovative ideas in R&D investment.

Thus, the aim of the current paper is to generalize the six degrees of separation notion from the perspective of local culture (which can be measured through a general

measure of local attitude and cultural diversity[2]) and to shed additional light on the role of the cultural factor for the R&D investment efficiency in the NUTS2 localities in the EU27. The next section offers an overview of the conceptual motivation behind this generalization of the six degrees of freedom and their involvement in the percolation mechanism explained in Tubadji and Nijkamp (2014). Section 3 presents the data for the numerical illustration about the available EU27 regions and the difference-in-differences approach applied for operationalizing our six degrees of separation and cultural percolation of R&D ideas hypothesis. Section 4 offers some concluding remarks and discussed the replicability of the numerical illustration offered here.

2 Cultural percolation of R&D ideas as a function of the six degrees of separation

Below follows an expose of the conceptual blending between a Schumpeterian quality ladder model and a local cultural relativism perspective. We discuss a Schumpeterian motivated model for percolation of new ideas through the local social network. The cultural percolation is defined as the passage of an idea from the pool of possible ideas through the social network structure, till it transforms its status into an idea selected by the decision-making processes for R&D investment. The selected idea for R&D investment becomes an output from this investment with actual innovative value added documentable by a patent (see Tubadji and Nijkamp 2015 for a more comprehensive elaboration of the cultural percolation model). Every percolation process however depends on its threshold, which operates like a tipping point and is generally very difficult to capture quantitatively, as we will elaborate on this below. Thus, our novel proposition here is that the six degrees of separation paradigm is theoretically justifiable to be used as an empirical approach to approximate the tipping point of this percolation process through the social network. Below follows a gradual elaboration of this rationale.

The essence of the cultural percolation model of new ideas lies in the filtering and selection of new ideas from the pool of possible ideas as the distinct ideas that will be invested on for research. This model is based on one of the most refined examples of modeling R&D investment decisions—the Schumpeterian quality-ladder model. As a model of innovation (Schumpeter 1934, 1942; Aghion and Howitt 1992; King and Levine 1993; Ericson and Pakes 1995; Cohen and Klepper 1996a, b; Klette and Griliches 2000), it was developed in the context of firms which choose whether to invest in a specific R&D project based on the net present value financial analysis of the investment opportunity. In this setting, the firm is interested in the expected duration of a monopoly profit confronted with the total expenditure required for R&D (without accounting for its distribution across researchers). Moreover, the Schumpeter ladders/rungs of quality express the behaviour of the firm vis-a-vis the firm-specific knowledge at a given constant stock of universal knowledge which is the basis of

---

[2] See Guiso et al. (2006) for an overview of alternative traditional measures for approximating local cultural attitudes, such as religion, trust etc. The approximation we opt to use here follows the more encompassing Culture-Based Development (CBD) notion of culture as an aggregate of local attitudes (Tubadji 2012, 2013) which summarizes the general cultural diversity notion.

its evaluation of future opportunities. Formally, the model is driven by the (i) the aggregate flow of resources expended by potential innovators in a given sector when the highest rung (highest ever addressed step of the ladder of knowledge) is addressed for investment by R&D; and (ii) the net expected return on investment. The expected return on investment however, as argued in Tubadji and Nijkamp (2015), is an output of a forecast function in which uncertainty of the prediction due to cultural bias on the interpretation of present and future possibilities is involved. Put differently, the Schumpeter quality ladders investment choice decision-making is a function of two elements: (i) the available lump sum for investment; (ii) the expectation on the quality, i.e. the expected price of the new idea if it gets implemented. While the first element is a hard matter, the second element is above all a function of the decision maker's preferences and risk aversion biases, which are both derivatives of local culture. This can be expressed formally through the limits of the choice function and the attitude to risk function, as follows:

$$\lim f(\text{Choice}) = \lim f(\text{Investment\_Z}) + \lim f(\text{Expectation for NPV}) \tag{1}$$

$$\lim f(\text{Choice}) = \text{constant1} + \lim f(\text{All\_Knowledge}_t)$$
$$\times \lim f(\text{culturally\_rel\_risk\_attitude}) \tag{2}$$

$$\lim f(\text{Choice}) = \text{constant 1} + \text{constant 2} \times \lim f(\text{culturally\_rel\_risk\_attitude}) \tag{3}$$

The above three (lines (1)–(3)) are a proof of the dependence of the limits of choice in the Schumpeterian quality ladder model on the cultural relativity of the chooser. Namely, in (1) choice is a function of some investment lump-sum (Investment_Z) and the net present value (NPV) expected from it. As the Investment_Z is a constant its limit is also a constant, the limit of the choice will vary according to the variation of the expectation for the NPV. The NPV as well is a function of the knowledge and the expectations, demonstrated in (2). However, as shown in (3), the highest level of knowledge attainable in a given moment of time is a constant for everyone who chooses. As Schumpeter's model suggests, then again its limit is a constant. Therefore, the variation of the limits of the choice is determined by the limits of the cultural relativity of risk attitudes. The latter means that as risk attitudes differ in different localities, for the same Z and same knowledge (innovation epoch), the choice will tend towards a particular number depending on the risk attitude, which has its limit that is locally varying according to local culture. Put differently, local culture determines what the limit of the function of risk attitude tends to, which ultimately determines the limit of the function of choice.

Next, we define choice as a 0 and 1 function, where 1 denotes a chosen, and 0 a not chosen idea. This means, local culture determines risk evaluation which also jumps sharply between 1 and 0. Therefore, our assumption that the choice of ideas is a cultural percolation process seems plausible. So, the question which remains is what the mechanism that shapes the cultural evaluation of risk attitudes is. To answer this question, we need to understand two things: (a) the cultural attitude function is a cultural percolation [varying between 0 and 1, which is 1 when something is successfully communicated (agreed upon) in the social network, as opposed to 0

when unsuccessful communication (non-agreement) upon the same matter occurs in the social network]; (b) the threshold the function has to tend to in order for its limit to tend to 1 or 0, in other words, what is the threshold for the cultural percolation to happen or not. These two matters are explained below.

To fully understand the mechanism of the dependence between choice and culture, we employ here the cultural percolation hypothesis. Percolation processes were introduced by Broadbent and Hammersley (1957) to model the random flow of a fluid through a medium. Percolation distinguishes from diffusion theory in the fact that diffusion is defined as a random movement in a structureless medium. The percolation process in a particular environment (a specific constant type of lattice) can be explained and generalized through the following mathematical condition for the occurrence of a percolation $X_i(e)$:

$$X_i(e) = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i, \ 0 < p_i < 1. \end{cases} \quad (4)$$

What is of special interest in this percolation process is the essential existence of the so called percolation threshold. According to Kolmogorov's zero-one law (see, for instance, van den Berg 2008), for any given p, the probability that an infinite cluster exists is either zero or one. Put differently, the percolation—or the full penetration from top to bottom of a surface—is a dummy variable; it either is a fact (denoted by 1), or is not a fact (denoted by 0). And the probability of switching between 0 and 1 is an increasing function of the connectivity in the lattice, denoted with CONN. For this function of CONN, there is a critical CONN (denoted by $CONN_c$) below which the probability of percolation is always 0, while above this critical CONN the probability of percolation is always 1. This critical CONN is the percolation threshold. Moreover, for the percolation threshold it is true that beyond the critical value of CONN, the probability of an open path from the top to the bottom to be achieved increases sharply from very close to zero to very close to one in a short span of values of CONN. Because of this last characteristic, the percolation threshold is very closely resembling the behavior of a tipping point (see Lamberson and Page 2012). The tipping point in percolation occurs when p_c is passed which allows the percolation of a liquid through the lattice. After a gradual change in CONN (from $CONN_c - 1$ to $CONN_c$), there is a change of the initial conditions allowing for a percolation state to switch from 0 to 1.

The calculation of the tipping point in percolation is a complex mathematical question. In some cases, $CONN_c$ can be explicitly calculated. Such a case is, for instance, the calculation of the percolation threshold for the square lattice Z2 in two dimensions, where $CONN_c = 1/2$ for a bond percolation (Kesten 1982). For most infinite lattice graphs, however, p_c cannot be exactly calculated. For example, p_c is not known for a bond percolation in the hypercubic lattice in two dimensions. Luckily, what is known is that there is a limit case for lattices in many dimensions (the Bethe lattice), whose threshold is at $CONN_c = 1/(z - 1)$ for a coordination number z (see Bethe 1935). More precisely, a Bethe lattice, introduced by Hans Bethe in 1935, is a connected cycle-free graph where each node is connected to z neighbours, and z is the coordination number. It can be understood as a tree-like structure starting from a central node,

with all the nodes arranged in shells around the central one. The central node is called the root or origin of the lattice. The number of nodes in the kth shell is given by:

$$N_k = z(z-1)^{k-1} \quad \text{for } k > 0. \tag{5}$$

In some cases, the above expression can be modified to specify that the root node has $z-1$ neighbours and the statistical mechanics of lattice models are often exactly solvable. As well-known from Axelrod (1997), social behavior—and culture more specifically—is a coordination game within a social network. Therefore, according to the above formula, for a critical number of coordinated entities (persons at an individual level or co-existing conflicting cultural attitudes or a given cultural distance—for a local society), the tipping point for percolation of an 'non-solid' matter through society can be calculated according to $p_c = 1/(z-1)$. Besides the logical similarity however, it is still unclear what is the economic reasoning that can motivate and serve as a plausible explanation for applying such a measurement approach with regard to ideas percolation in the social lattice. Put differently, the cultural percolation depends on its threshold and its threshold depends on z, which is dependent on the particular lattice structure. But the lattice structure, i.e. the connectivity of the social network is not easily observable. So what is z for the social lattice? How to determine the threshold of the social lattice depending on its structure of connectedness that is non-observable in its full extent?

To answer this second question is pivotal to our study.And to answer what is the threshold which drives the limits of choice by determining the percolation through the social network, we infer here the SDSC literature. A renowned measure for the connectedness in society—and an important measure for the transmission of information between people—is the well-known 'six degree' phenomenon, presented below. For the first time the six degrees phenomenon was studied by Traverse and Milgram (1969) with a sample of 296 people from Nebraska. The experimental design required a letter to be sent to a stockbroker in Boston under the rule that the letter is delivered only through first-name-basis acquaintances of the sender. Traverse and Milgram (1969) found that it takes on average six people for a letter to reach from sender to target receiver. This finding was theoretically proved by Watts and Strogatz (1998). Moreover, Watts (2004) and Dodds et al. (2003) replicated the experiment with a much larger sample of 48,000 people on the internet. The experimental design was roughly the same, with the difference that the target recipients were randomly distributed around the globe. Watts' (2004) results confirm that the average number of steps of connectedness required for an idea to reach from a sender to a receiver is six connected people. Thus, the main proposition of our study is that the 'six degrees' phenomenon represents the connectedness within a social network and characterizes the behaviour of its lattice (see also Watts et al. 2002). From another perspective, the notion of cultural diversity has been long debated as a potential source of positive or negative effects on ideas, innovation and finally on local productivity (see for example Ottaviano and Peri 2005, 2006; Bellini et al. 2008; Suedekum et al. 2010). Therefore, the change of cultural diversity can be expected to influence the world of percolation of ideas. Moreover, this effect is most likely to pass through the interaction process of local and foreign workers. Thus, we assume that the six degrees (the likelihood of six

people in a row being of a different origin than the local cultural origin) is a relevant approximation of the missing information on the unobservable structure of the social network and its threshold for cultural percolation. Put differently, we suggest that the hypothesis is made that since six connected people in a row are the requested number of people for an idea to be successfully transmitted through the social network, then it might be generalized that the successful communication on a new idea proposed for R&D investment might be also subject to dependence on a SDSC between people. Only in the case of selection of an idea, we make the additional assumption that people are not just connected acquaintances, but also that they share a cultural connectedness, a cultural minimum distance (i.e. zero cultural distance—belonging to the same culture), or a full cultural proximity (proximity on attitudes and origin equal to identity). Their connectedness based on common cultural value-based will give homophily basis for reaching an agreement on the final decisions when passing the idea through the filter of riskiness attitude evaluation.

Finally, however, we note that connectedness between people of the majority in a locality leads to its closedness and lack of receptiveness to new ideas, while more models and more diversity are associated with better innovation and efficiency of solution finding to any problem (see Hong and Page 2001). Therefore, what can be expected to have a positive effect on R&D selection of the most efficient and value adding new ideas is the likelihood of SDSC between people with non-local origin. In other words, we will transform the six degree of separation into an innovative measure of diversity, where diversity is defines as the likelihood of six random people in a row happening to be with a non-local origin.

Based on the above assumptions, we ultimately suggest that the change of the likelihood of having six people in a row being from a diverse cultural origin defines to what limit the cultural attitude in a locality will tend towards new ideas—in other words, if the limit of risk evaluation will tend to 1, the choice of the most efficient ideas for R&D is also having a limit tending to 1. Put differently, our cultural percolation of R&D ideas hypothesis embodies the following testable proposition:

H01: *The higher the likelihood of having six people in a row coming from a diverse cultural origin, the closer the limit of risk evaluation will tend to one and therefore the choice of most efficient ideas for R&D will have a limit tending to 1 for a higher number of innovative ideas.*

The next section provides an illustrative example for an empirical operationalization of this hypothesis.

3 Difference-in-differences operationalization of the cultural percolation of ideas hypothesis

3.1 Database

Our dataset to test H01 is compiled of data on economic indicators from EUROSTAT and the European Social Survey (ESS). In particular, for the purpose of our numerical example, we use the ESS data available for the two years 2002 and 2004 to quantify the

likelihood of six degrees of cultural diversity in the social network, and we complete the data pool into a panel dataset by taking the rest of the economic variables from EUROSTAT Regional Database. Our dataset is at a NUTS2 level.

From the biennially collected ESS dataset, we use the number of people at a particular NUTS2 level who have indicated that they have lived in the area concerned as immigrants who have moved into it during the last 20 years. Thus, our main explanatory variable of interest—six degrees of separation (or culturally-based social connectedness)—is defined as the probability of six random people from the pool of interviewed people in a locality being from an immigrant cultural origin. The exact formula through which we derive the likelihood of six people in a row being from an immigrant cultural origin is respectively: *likelihood_six_in_a_row_M = ((total_M − 1)/(total − 1)) × ((total_M − 2)/(total − 2)) × ((total_M − 3)/(total − 3)) × ((total_M − 4)/(total − 4)) × ((total_M − 5)/(total - 5)) × ((total_M − 6)/(total − 6))*, where *M* denotes being from a non-local cultural origin, *total_M* is the number of observation of type *M* and *total* is the overall number of people living at the NUTS2 region for which the likelihood is estimated (see Traverse and Milgram 1969; Granovetter 1973; Watts and Strogatz 1998). As our measure reflects immigrants, it indicates that the likelihood of heterogeneity and cultural diversity is higher when the likelihood of six degrees of separation is higher. That is why the expected relationship of six degrees of separation and innovation efficiency is expected to be positive.

The main indicators for the outcome variable of interest in our analysis, namely, the number of patent applications and the R&D expenditures in millions, are measured through data from EUROSTAT. Thus, our main dependent variable is the number of patent applications per R&D expenditure. It has been rightly argued in the economics of ideas literature, that ideas and innovation are by nature non-rival and non-excludable goods; and the whole concept of providing patents as a re-compensatory mechanism for innovators is highly contestable (Harabi 1995; Storper 1995). Yet, the patent practice has been widely promoted and adopted throughout—in particular—the EU as the 'rule of the game', resulting in policy-supported data collection that has been performed in this field by statistical authorities in a consistent manner. Moreover, even if hampered in the statistical representativeness of the data by confidential protectionist free-riding firm behaviours, the data on patenting of an innovative idea is still the registration of an ultimate act of cooperative behaviour, where the firm innovator tries to both take the due economic advantage and also allows for the agglomeration and spillover effects based on its innovation to take place. So, from a regional development perspective, it is indeed the patented innovations that are meaningful, as they are part of the cooperative game and are most likely to quantify properly the likelihood for the future effect of innovations on local development.

EUROSTAT is also a source for our control variables: total employment, workers with tertiary education and GDP. The number of agricultural workers approximates the economic structural differences in the localities (see Evans and Harrigan 2003). The motivation for the use of control for human capital (as workers with tertiary educational degree or higher—according to the EUROSTAT classification) stems from the endogenous growth models (see for example Uzawa 1986; Romer 1986; Lucas 1988; Rebelo 1991; Fischer et al. 2009), while GDP per capita is a gross measure of

the economic productive capacity of the place and approximates its potential to invest in R&D (see Antonelli 1990; Tödtling 1992).

3.2 Estimation strategy

Our theoretical hypothesis is operationally translated into the existence of three types of conditions over time in the NUTS2 regions: regions with an increasing, regions with stable (not changing) and regions with decreasing diversity in the social networks (namely, six people in a row being from a different cultural origin). Thus, the six degrees of diversity in the local networks is a proxy that identifies in a broader sense the culture in the NUTS2 regions. Next, each region experiences an individual path of R&D output efficiency in terms of scientific output (i.e. innovative ideas to be patented per millions of R&D investment) in the different periods. Put differently, our study focuses on a specific cultural intervention (change of diversity in terms of change of the likelihood of six random people drawn from a locality to be sharing same non-local cultural origin) as a treatment effect) that leads to differences in local R&D productivity over time. In addition, we use human capital per worker and GDP per capita as alternative controls to capture the eventual omitted variables.[3] With the use of these data, we apply a difference-in-differences approach, following Villa (2012),[4] as exemplified in model (6) below:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 T_{it} + \beta_3 X_{it} \times T_{it} + \beta_4 Z_{it} + e_{it} \qquad (6)$$

where Y is the R&D output efficiency; X_{it} indicates the belonging to the treatment group or not (i.e. X_{it} has three possible logical conditions—it either increases, decreases or remains stable between two time periods and these three possibilities define the types of the treatment and exhaust its domain; next, X_{it} is quantified as follows: when we are for instance interested in increase of diversity as a treatment, X_{it} is 1, if in the period 2002 to 2004 there is an increase of the likelihood of six people in a row being of foreign cultural origin, and 0 otherwise); T_{it} stands for identifying whether the observation falls in the past period (then the variable is 0) or in the later period (when the dummy is respectively equal to 1). The interaction between the preceding two variables helps us capture the treatment effect, reflected in the coefficient β_3. Z_{it} is a vector of the additional control variables. And e stands for the error term. The expected effects from the three types of treatment are as follows. When diversity increases, i.e. the likelihood of having an increasing probability of six random people in the population being from a foreign origin, this treatment is expected to increase the number of patented new ideas in a locality, as more innovative ideas have more

---

[3] More refined data with richer information on the sectoral specialization per sector or firm-related information—such as number of firms per NUTS2 dealing with R&D—could improve further the estimation reliability. In any case, the eventual resulting bias here stems from imperfect proxies and not from a mis-specification due to an omitted variable bias. Therefore, including the GDP measure to a big extent controls for all these concerns and our estimations are fairly reliable as a specification, as confirmed by the post estimation standard tests presented in the tables.

[4] See also Bertrand et al. (2004) on the reliability of the difference-in-differences method.

efficiently penetrated through the social lattice there. The opposite effect is expected from the decrease of the same likelihood. And no change in the treatment is supposed to be associated with no change in the efficiency of choice on ideas, i.e. in the number of percolation of really innovative ideas through the social lattice in a locality. If one of these three domains of treatment turns out to be statistically significant our hypothesis cannot be falsified. If more than one domain of treatment shows statistical significance this would be a sign for a non-linear dependence between treatment and percolation of new ideas.

Consequently, for our biennial datasets (2002–2004), we create a variable 'path-length' which expresses the six degrees of cultural diversity in the society by estimating the likelihood of having six people in a row from the interviewed people not being originally from this locality for 2002 and for 2004. We take the difference of the path-length of diversity from 2002 to 2004. The resulting change of the six degrees of diversity (i.e. our treatment) falls in three possible categories: increase, stable state and decrease of the likelihood of diversity in the local network. We therefore, specify three different types of treatment: X_1 equal to 1, when there is a positive change, and 0 otherwise; X_2 equal to 1, when there is a zero change of cultural diversity between 2002 and 2004 in the locality, and 0 otherwise; and X_3 equal to 1, when there is a negative change of the likelihood of diversity in the local social network.

Initially, a simple linear OLS regression with robust errors is estimated for each of the three possible treatment effects: increase in diversity, stability (no change of diversity) and decrease of diversity (measured as having six people in a row being a foreign origin). Operationalizing model (6), we obtain the interactions of each variable with the time indicating dummy variable T_{it}. We thus examine the coefficients of the three different interaction terms in order to find out which of the three cultural treatments generates an impact on local R&D output efficiency. Next, we will delve further into the reliability of the estimated significant treatment effect by involving real data (such as share of human capital or GDP per capita) as additional controls or dealing with the eventual omitted-variable bias artificially by estimating the model with fixed effects for each region. The reliability of the different approaches will be compared through a standard Ramsey test for omitted variables as well.

3.3 Numerical illustration results

Firstly, we compare the frequency of the three different possible types of cultural treatment occurring over the period 2002 to 2004. Table 1 presents this overview:

The descriptive Table 1 demonstrates that in our data the frequency with which the three types of treatment are observed are actually almost equally distributed. This means, that an almost identical percentage of the EU regions have experienced an increase, a stable state or a decrease of the likelihood of six people in a row being from a different cultural origin. Therefore, it is interesting if these treatments that are almost equal in likelihood to happen produce different effects on local R&D output efficiency.

Table 1 Six degrees of cultural diversity: three different treatment effects, 2002–2004

Treatment	Var.	Freq.	Percent	Cum.
>0	tincr	62	37.35	37.35
0	t0	48	28.92	66.27
<0	tdecr	56	33.73	100
Total		166	100	

The table presents group distribution of the NUTS2 regions into three different treatment effects from the change or lack of change of the likelihood of six degree of cultural diversity in the local social network. The three treatments identifying the increase, stability (zero change) or decrease of the likelihood of cultural diversity in the social network. The table presents also the percentage of regions experiencing the respective treatment

Source: Authors' calculations

Table 2 Difference-in-differences test for the effect of the three alternative treatments from six degrees of cultural diversity on patents per million investment in R&D

Spec.	>0		0		<0	
Dep. var.	pat_rndm					
	Coef.	t-Value	Coef.	t-Value	Coef.	t-Value
pt_d-i-d_eff	−0.243	−0.13	4.117	2.50	−3.213	−1.30
Treatment	0.238	0.12	−4.104	−2.49	3.213	1.30
Post	−4.569	−3.16	−5.748	−4.34	−3.502	−3.85
F(3,131) =	8.43		7.83		7.77	
Prob > F	0.0000		0.0001		0.0001	
R^2	0.12		0.16		0.15	
Root MSE	6.396		6.254		6.293	
No. of observations	135		135		135	

The table presents a naïve OLS estimation with robust standard errors for three alternative specifications of our empirical model (4) with the basic difference-in-differences application: R&D efficiency (patents per millions of investment) regressed on the dummies for treatment effect (treatment) and time period (post) (equal to 0 if year = 2002 and 1 otherwise) and their multiplication term reflecting the treatment effect. The three specifications are respectively for increase (>0), stable (zero change) and decrease (<0) change of the likelihood of having six people in arrow being from a foreign cultural background within the local social network

Source: Authors' calculations

All three different types of treatments are examined as three alternative specifications of an operationalization of model (6) in its most reduced form. The results are presented in Table 2 below.

As Table 2 shows, the cultural diversity instability—both in terms of increase or decrease of the six degrees of likelihood to encounter cultural diversity in the local network—has a negative but insignificant effect on the R&D output efficiency. It is however, the stability of cultural diversity in a locality that seems to be significantly and positively associated with the increased innovation potential of the locality. The standard statistics of the three models perform very closely to each other. This result

can be interpreted as a sign that diversity instigates innovation in a stable climate without social turbulence. Put differently, if in a given interval of time there is cultural persistence in a locality, this and only this drives the limits of the cultural percolation of new ideas in a locality. And the impact from persistent cultural diversity (i.e. from a six degrees of separation measure defined as the likelihood of having connected in a row six random culturally diverse people in the locality) on innovation efficiency (measured in number of patents) is found positive. Thus, this test result fails to reject our working hypothesis and indicates a linear dependence between the limit of the function of choice of ideas from the pool of potential ideas and the limit of the function of local cultural attitude.

Yet, the results from an OLS regression with robust standard errors may still involve an omitted variable bias. In an attempt to deal with this problem in estimating our treatment effect of interest, we perform an estimation of four alternative specifications, as presented in Table 3 below.

Table 3 presents the specifications: involving a share of agricultural workers and a share of human capital in the locality as controls; then a parsimonious version of it involving only the share of human capital as a control; next a control variable GDP per capita is introduced in the operationalization of model (6); and lastly, model (6) is run in its fully parsimonious form but in combination with regional fixed effects to take away all the eventual omitted locality-specific characteristics.

The additional economic prosperity capturing variables such as share of human capital and GDP per capita report significant but negative coefficients and t-values. This is in line with the expectation that when the investment volume can be bigger, the risk evaluation gets more risk-avoiding and thus the less risky and respectively less innovative ideas are opted for, which ultimately results in a decrease in (or a negative departure from the optimal) innovative potential in richer localities. Naturally, the standard statistics of the empirical models improve with the inclusion of the control variables. The results of our four specifications are consistent, while an improvement in the omitted variable Ramsey RESET test is observed from specification 1 to 4. This second set of results confirms our first impressions from the numerical example considered here. Indeed, our hypothesis that the variation of the likelihood of having SDSC of culturally more diverse people in a locality is a statistically significant predictor of the limit of the function of choice for innovative R&D ideas tending to one more often seems in agreement with our empirical findings.

These results, though potentially intriguing, would have to be considered within certain data limitations imposed by the ESS data used. The ESS data needs to be double weighted to come close to the real representation of the likelihood of an individual to be chosen and to adjust to country-specific population sampling. Therefore, we avoid weighting the data and we refer to the direct number of observations, as the weighting procedures themselves do not add much to the reliability of the results. This numerical illustration presented here serves as a model for empirical application of the model of interest and aims only to aid the replication of the model with a better representative sample. Therefore, we suggest the analytical part of the paper to be understood as an example for potential operationalization of a mixed approach between difference-in-differences and six degrees of separation in a diversity context. It is possible that

Table 3 Stable six degrees of cultural diversity effect and omitted variables tests

Spec	1		2		3		4	
Dep. var.	pat_rndm							
	Coef.	t-Value	Coef.	t-Value	Coef.	t-Value	Coef.	t-Value
pt_d-i-d_eff	3.852	2.39	3.856	2.40	3.818	2.47	4.372	2.09
t0	−3.956	−2.48	−4.050	−2.54	−2.730	−1.93	−5.436	−2.58
Post	−5.501	−4.32	−5.474	−4.31	−5.740	−4.50	−6.581	−4.17
agri_s	4.607	1.03	–	–	–	–	–	–
hc_s	−12.060	−2.24	−13.489	−2.70	–	–	–	–
gdppc	–	–	–	–	0.000	−3.13	–	–
_cons	7.653	3.61	8.371	4.29	8.444	4.50	6.599	4.18
Fixed effect	–		–		–		Yes	
F	(5,129) = 4.82		(4,130) = 5.52		(4,130) = 5.85		(62,58) = –	
Prob > F	0.0005		0.0004		0.0002		–	
R^2	0.19		0.18		0.23		0.60	
Root MSE	6.202		6.190		5.989		6.462	
No. of observations	135		135		135		135	
Ramsey RESET test								
F	F(3,126) = 2.29		F(3,127) = 2.19		F(3,127) = 8.18		F(3,55) = 628.07	
	Prob > F = 0.0814		Prob > F = 0.0925		Prob > F = 0.0001		Prob > F = 0.0000	

The table presents a test for the effect of the stability (zero change) of the six degrees of cultural diversity (likelihood of selecting six people in a row always from a foreign origin) in the share of patents produced per million euros of R&D investment in the locality. Difference-in-difference estimation for the 2002–2004, for four alternative specifications is shown, with varying treatment of the omitted variable bias: (i) involvement of control variables share of human capital and GDP per head; (ii) control variable share of human capital per head; (iii) use of control variable DGP per head; (iv) introduced fixed effects on regional level to artificially remove the local economic input differences

Source: Authors' calculations

this original methodological solution can be interesting to be replicated with bigger representative datasets on a regional level.

4 Conclusion

The key message from our difference-in-differences test is that the creation of new knowledge in a locality seems to be positively determined by a stable likelihood of having connected in a row six people who are carriers of cultural diversity in the local network. This result is a confirmation of the main cultural premise of three lines of argument: (i) culture (measured with an innovative mixed measure combining diversity and six degrees of separation between people in a social network context) is a critical factor for the limits of the function of choice of R&D ideas determining the innovative process in a locality; (ii) if we combine the messages from our results, it seems that they agree with our proposition that the six degrees of separation paradigm approximates the tipping point for the unobservable structure of the whole social network in a locality; i.e. a strong connection between the tipping point of the functioning of the social network communication and the local value system of this locality seem indeed to exist; this conclusion stems from the fact that the tipping point for the percolation of new ideas seems to be determined positively and statistically significantly, when the cultural diversity in the locality is used to approximate the six degrees of separation driven tipping point for the percolation of new R&D-investment-ideas; (iii) our results are filling in a caveat to the endogenous growth discourse by showing that, depending on the local cultural values, an objectively important output (such as R&D ideas) may be different, even if the economic investment and other input factor differences (such as the share of human capital) are controlled for.

In summary, given all relevant statistical constraints in our approach, and taking into account the available data limitations, we observe that the results from our numerical illustration report a confirmation of the main cultural impact on innovation hypothesis of our inquiry. Yet, it seems to be stability rather than an increase or decrease of diversity in the local networks that stimulates R&D output in the present case.

Above all, the data limitations do not allow for generalized conclusions. However, the current study presents an original conceptual blending, and offers an empirical operationalization for the purpose of exploring the intersection of three streams of literature: (innovation, diversity and six degrees of separation) as a determinant for the selection of new ideas in decision making processes related to investment in R&D.

Appendix

See Table 4.

Table 4 List of the EU27 regions included in the sample of the numerical illustration

NUTS 2	Name of locality	NUTS2	Name of locality
AT11	Burgenland (AT)	HU10	Közép-Magyarország
AT12	Niederösterreich	HU21	Közép-Dunántúl
AT13	Wien	HU22	Nyugat-Dunántúl
AT21	Kärnten	HU23	Dél-Dunántúl
AT22	Steiermark	HU31	Észak-Magyarország
AT31	Oberösterreich	HU32	Észak-Alföld
AT32	Salzburg	HU33	Dél-Alföld
AT33	Tirol	LU00	Luxembourg
AT34	Vorarlberg	NL11	Groningen
CZ01	Praha	NL12	Friesland (NL)
CZ02	Strední Cechy	NL13	Drenthe
CZ03	Jihozápad	NL21	Overijssel
CZ04	Severozápad	NL22	Gelderland
CZ05	Severovýchod	NL23	Flevoland
CZ06	Jihovýchod	NL31	Utrecht
CZ07	Strední Morava	NL32	Noord-Holland
CZ08	Moravskoslezsko	NL33	Zuid-Holland
EL12	Kentriki Makedonia (NUTS 2010)	NL34	Zeeland
EL14	Thessalia (NUTS 2010)	NL41	Noord-Brabant
EL21	Ipeiros (NUTS 2010)	NL42	Limburg (NL)
EL23	Dytiki Ellada (NUTS 2010)	PL11	Lódzkie
EL24	Sterea Ellada (NUTS 2010)	PL12	Mazowieckie
EL30	Attiki	PL21	Malopolskie
EL41	Voreio Aigaio	PL22	Slaskie
EL43	Kriti	PL31	Lubelskie
ES11	Galicia	PL32	Podkarpackie
ES12	Principado de Asturias	PL33	Swietokrzyskie
ES13	Cantabria	PL34	Podlaskie
ES21	País Vasco	PL41	Wielkopolskie
ES22	Comunidad Foral de Navarra	PL42	Zachodniopomorskie
ES23	La Rioja	PL43	Lubuskie
ES24	Aragón	PL51	Dolnoslaskie
ES30	Comunidad de Madrid	PL52	Opolskie
ES41	Castilla y Leon	PL63	Pomorskie
ES42	Castilla-la Mancha	PT11	Norte
ES43	Extremadura	PT15	Algarve
ES51	Cataluña	PT16	Centro (PT)
ES52	Comunidad Valenciana	PT17	Área Metropolitana de Lisboa
ES53	Illes Balears	PT18	Alentejo

Table 4 continued

NUTS 2	Name of locality	NUTS2	Name of locality
ES61	Andalucía	SI01	Vzhodna Slovenija (NUTS 2010)
ES62	Región de Murcia	SI02	Zahodna Slovenija (NUTS 2010)
ES70	Canarias (ES)		

The table shows the 83 NUTS2 regions from EU27 countries, which are available from the ESS dataset. After matching with the EUROSTAT Regional dataset the regions from Greece were dropped due to lack of data on patents for that country

Source: Authors'

References

Aghion, P., Howitt, P.: A model of growth through creative destruction. Econometrica **60**, 323–351 (1992)

Akerlof, G., Kranton, R.: Economics and identity. Q. J. Econ. **CXV**(3), 715–733 (2000)

Antonelli, C.: Induced adoption and externalities in the regional diffusion of information technology. Reg. Stud. **24**, 31–40 (1990)

Axelrod, R.: The dissemination of culture: a model with local convergence and global polarization. J. Confl. Resolut. **41**(2), 203–226 (1997)

Barro, R., Sala-i-Martin, X.: Economic Growth. McGraw-Hill, New York (1995)

Batabyal, A., Nijkamp, P.: Biodiversity prospecting over time and under uncertainty: a theory of sorts. ERSA 2014 Paper (2014)

Bellini, E., Ottaviano, G., Pinelli, D., Prarolo, G.: Cultural diversity and economic performance: evidence from European regions. HWWI Research Paper 3-14 (2008)

Bertrand, M., Duflo, E., Mullainathan, S.: How much should we trust differences-in-differences estimates? Q. J. Econ. **119**(1), 249–275 (2004)

Bethe, H.: Statistical theory of superlattices. Proc. R. Soc. Lond. Ser. A **150**, 552–575 (1935)

Bresnahan, T., Trajtenberg, M.: General purpose technologies 'Engines of growth'? J. Econom. **65**(1), 83–108 (1995)

Bresnahan, T., Yin, P.: Reallocating innovative resources around growth bottlenecks. Discussion Papers 09-022, Stanford Institute for Economic Policy Research (2010)

Broadbent, S., Hammersley, J.: Percolation processes I. Crystals and mazes. Proc. Camb. Philos. Soc. **53**, 629–641 (1957)

Cohen, W., Klepper, S.: A reprise of size and R&D. Econ. J. **106**, 925–951 (1996a)

Cohen, W., Klepper, S.: Firm size and the nature of innovation within industries: the case of process and product R&D. Rev. Econ. Stat. **78**(2), 232–243 (1996b)

Davelaar, E., Nijkamp, P.: Industrial innovation and spatial systems: the impact of producer services. Serie Research Memoranda 0045. VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics (1989)

Dodds, P., Muhamad, R., Watts, D.: An experimental study of search in global social networks. Science **301**(5634), 827–829 (2003)

Eaton, J., Kortum, S.: Trade in ideas: patenting and productivity in the OECD. NBER Working Papers 5049, National Bureau of Economic Research, Inc. (1997)

Ericson, R., Pakes, A.: Markov-perfect industry dynamics: a framework for empirical work. Rev. Econ. Stud. **62**(1), 53–82 (1995)

Evans, C., Harrigan, J.: Distance, time, and specialization. NBER Working Papers 9729, National Bureau of Economic Research, Inc. (2003)

Fischer, M., Bartkowska, M., Riedl, A., Sardadvar, S., Kunnert, A.: The impact of human capital on regional labor productivity in Europe. Lett. Spat. Resour. Sci. **2**(2–3), 97–108 (2009)

Granovetter, M.: The strength of weak ties. Am. J. Sociol. **78**(6), 1360–1380 (1973)

Grossman, G., Helpman, E.: Innovation and Growth in the Global Economy, 1st edn., vol. 1, no. 0262570971. MIT Press Books, The MIT Press, Cambridge (1993)

Guiso, L., Sapienza, P., Zingales, L.: Does culture affect economic outcomes? J. Econ. Perspect. **20**(2), 23–48 (2006)

Harabi, N.: Appropriability of technical innovations: an empirical analysis. Res. Policy **24**(2), 981–992 (1995)

Hong, L., Page, S.E.: Problem solving by heterogeneous agents. J. Econ. Theory **97**(1), 123–163 (2001)

Kesten, H.: Percolation Theory for Mathematicians. Birkhäuser, Boston (1982)

King, R.G., Levine, R.: Finance and growth: Schumpeter might be right. Q. J. Econ. **108**(3), 717–737 (1993)

Klette, T.J., Griliches, Z.: Empirical patterns of firm growth and R&D investment: a quality ladder model interpretation. Econ. J. **110**(463), 363–387 (2000)

Lamberson P., Page, S.: Tipping points. SFI Working Paper 2012-02-002 (2012)

Lichtenberg, F.: R&D investment and international productivity differences. NBER Working Paper No. 4161 (also Reprint No. r1813) (1993)

Lozares, C., Verd, J., Cruz, I., Barranco, O.: Homophily and heterophily in personal networks. From mutual acquaintance to relationship intensity. Qual. Quant. **48**(5), 2657–2670 (2014)

Lucas, R.: On the mechanics of economic development. J. Monet. Econ. **22**(1), 3–42 (1988)

Ottaviano, G., Peri, G.: Cities and cultures. J. Urban Econ. **58**, 304–337 (2005)

Ottaviano, G., Peri, G.: The economic value of culture diversity: evidence from US cities. J. Econ. Geogr. **6**, 9–44 (2006)

Rebelo, S.: Long-run policy analysis and long-run growth. J. Polit. Econ. **99**(3), 500 (1991)

Rodriguez-Pose, A.: Innovation prone and innovation averse societies: economic performance in Europe. Growth Change **30**, 75–105 (1999)

Romer, P.: Increasing returns and long-run growth. J. Polit. Econ. **94**(5), 1002–1037 (1986)

Romer, P.: Endogenous technological change. J. Polit. Econ. **98**(5), 71–102 (1990a)

Romer, P.: Capital, labour and productivity, pp. 337–367. Bookings Papers on Economic Activity, Microeconomics (1990b)

Romer, P.: The origins of endogenous growth. J. Econ. Perspect. **8**(1), 3–22 (1994)

Schumpeter, J.: The Theory of Economic Development. Harvard University Press, Cambridge (1934)

Schumpeter: Capitalism, Socialism and Democracy. Harper, New York (1942 [1975])

Storper, M.: Regional technology coalitions. An essential dimension of national technology policy. Res. Policy **24**, 895–911 (1995)

Suedekum, J., Wolf, K., Blien, U.: Cultural diversity and local labour markets. IZA Discussion Paper Series, DP No. 4619 (2010)

Tödtling, F.: Technological change at the regional level: the role of location, firm structure, and strategy. Environ. Plan. **24**, 1565–1584 (1992)

Traverse, J., Milgram, S.: An experimental study of the small world problem. Sociometry **32**(4), 425–443 (1969)

Tubadji, A.: Culture-based development: empirical evidence for Germany. Int. J. Soc. Econ. **39**(9), 690–703 (2012)

Tubadji, A.: Culture based development—culture and institutions: economic development in the regions of Europe. Int. J. Soc. Soc. Sci. **5**(4), 355–391 (2013)

Tubadji, A., Nijkamp, P.: The cultural percolation of new knowledge: a regional analysis of the cultural impact on knowledge creation in EU27. In: Bakens, J., Poot, J., Nijkamp, P. (eds.) The Economics of Cultural Diversity. Edward Elgar, Cheltenham (2015, forthcoming)

Uzawa, H.: Optimum technical change in an aggregative model of economic growth. Int. Econ. Rev. **6**(1), 18–31 (1986)

van den Berg, J.: Approximate zero-one laws and sharpness of the percolation transition in a class of models including two-dimensional Ising percolation. Ann. Probab. **36**(5), 1880–1903 (2008)

Villa, J.: Simplifying the estimation of difference in differences treatment effects with Stata. MPRA Paper No. 43943 (2012)

Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks. Nature **393**(6684), 440–442 (1998)

Watts, D., Dodds, P., Newman, M.: Identity and search in social networks. Science **296**(5571), 1302–1305 (2002)

Watts, D.: The "New" science of networks. Annu. Rev. Sociol. **30**, 243–270 (2004)

Zenou, Y.: Networks in economics. CEPR Discussion Papers 9021 (2012)