Doctoral Thesis of

Matteo Tubiana

"STUDIES ON THE ORGANIZATION OF KNOWLEDGE PRODUCTION"

Doctoral Student in Economics of Innovation, Knowledge and Complexity

XXX Cycle of Vilfredo Pareto Doctorate in Economics

Dipartimento di Economia e Statistica Cognetti De Martiis, Università degli Studi di Torino, Torino

Supervisor: professor Cristiano Antonelli

Year 2018

Acknowledgments

The most of my gratefulness goes to my supervisor, Cristiano Antonelli, who led me wisely in the Ph.D. journey.

I want to thank all my co-authors, Alessandra Colombelli, Ernest Miguélez and Rosina Moreno, who gave me the opportunity to learn from their precious human and academic experience.

A special, sincere thought goes to Enrica Maria Martino, Gianluca Orsatti and Massimiliano Coda Zabetta, friends, colleagues and companions in the last four years and for the years to come.

Index

A theory of interactions between the economics of information and knowledge

Matteo Tubiana[a,b]

[a] *Department of Economics and Statistics Cognetti de Martiis, University of Torino*

[b] *BRICK, Collegio Carlo Alberto*

Abstract

In this essay, I will bring evidence of the specific role interactions have in shaping the socio-economic system dynamics, from the standing points of the economics of knowledge and the economics of information. After recalling some significant steps in the literature of reference, the exercise will lead to the formulation of a basic theory of interactions in economics, building on a two-partite definition of what an interaction is according to i) what an interaction conveys and ii) which is the *locus* where it happens. The embeddedness of individuals in the social space will be appreciated and formalised. These arguments will lead to exploit the concepts of emergence and scaffoldings structures in order to contribute to a major issue in economic theory: coordination.

1. Introduction

The bulk of economic theory considered, and still considers today, the dynamics of socio-economic systems as driven by impersonal transactions between contractors. According to this view, market mechanisms regulate the exchanges of economic value, and each individual's behaviour respond to a set of exogenous preferences. In this narrative, there is little space for considering the role of social interactions between economic agents (Antonelli and Scellato, 2013). The rise of the economics of information firstly disputed such a perspective, indicating that the role of non-market based information exchange and their behavioural consequences are of foremost relevance to understanding economies' functioning (Akerlof, 1970; Stiglitz, 2000). Similarly, the economics of knowledge, appreciating knowledge creation as a recombinant process, eventually evidenced that the engine of economic growth, i.e. knowledge production and the consequent introduction of innovations, is a collective process (Antonelli, 2000).

In the following pages, I will recollect some important contributions from the heterodox economics of the last thirty years, in order to sketch a theory of interactions. It will oppose the depiction of economic reality as dominated by impersonal market transactions. Rather, it will stress the centrality of interactions, primarily from the point of view of the economics of knowledge, hence it will concentrate on the specific role interactions have within the dynamic, complex process that is the law of motion for economic development: innovation (Hanusch and Pyka, 2006). The adopted methodology consists of defining interactions along two coordinates. As a former step, I will try to define what an interaction is according to what it conveys. Secondly, I will recall the traditional notion of impersonal market transactions, compare it with the defined notion of interaction, and introduce each's archetypal *locus* where it takes place. Indeed, interactions are not detached by their context, rather they are strongly embedded, structurally and functionally, in the environment.

Figure 1. *An economic theory of interaction in figure.*

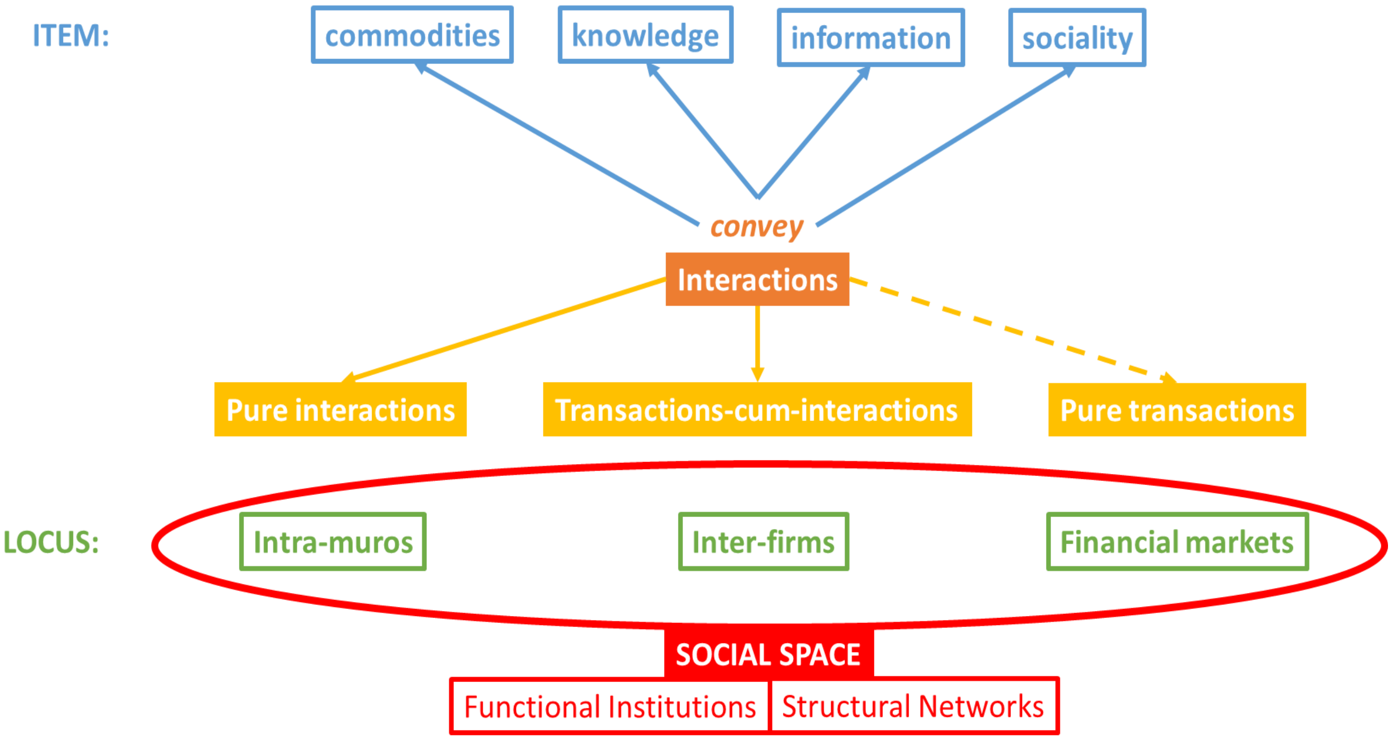


Figure 1 summarises the theory of interactions I will thoroughly treat in Section 3. The primary objective of this exercise is to affirm the socio-economic substance of interactions – complex socio-economic dynamics rest upon the functioning and developing of interactions, as opposed to impersonal transactions, among heterogeneous agents (Lane et al., 2009; Landini, Gallegati and Stiglitz, 2015). This is in opposition to the dominant Coasian perspective, dictating that under perfect conditions, i.e. no transactions costs and well-identified property rights, perfect efficiency is possible irrespective of resource allocation (Coase, 1960). Coase Theorem, as insightful as it has been for economic theory, de facto undervalues if not excludes the role of interactions.

Addressing the definition of interactions with a description of their functioning leads to consider their broader role in the economy at large (Glaeser and Scheinkman, 2000; Arata, 2017). In so doing, I will contribute to the topic of economic coordination, positing that the diffusion and efficacy of platforms for interactions can explain differences in countries' economic performance. However, economic coordination is a thorny topic in need of a thorough introduction.

At the very beginning of the economic discipline, Adam Smith introduced the problem of market coordination. It could reduce to the study of how *market* prices correlate with *natural* prices, whereby coordination happens when the two coincide. When market prices exceed natural prices, producers are enforcing consumers, via the price mechanism, to pay for their eagerness; when market prices, instead, are lower than natural prices, producers are likely to fail and production do not take place. In both cases, market forces alone do not coordinate the two entities of a market economy (supply and demand) inducing a substantial loss for society as a whole, e.g. inequalities and unemployment. Smithian masterpieces – *The theory of Moral Sentiment* (Smith, 1759) and *Wealth of Nations* (Smith, 1776) – are part of a research project aiming to understand how and when markets of autonomous individuals can coordinate without external interventions. Matthew Watson (2005) summarises Smith's enquiry, showing the cardinal role of moral philosophy in shaping Smith's answer. Indeed, justice is the pillar upon markets should reside for them to survive; individuals, intrinsically embedded in the social space, must fight to behave accordingly to a deontological rationality driven by propriety and equity, in contrast with an instrumental rationality striving for accumulate wealth, Smith says. The issue of market coordination cannot be taken, therefore, within pure economics but in-between moral philosophy and political economy, according to Smith.

Coordination soon became one of the central questions of economics. Daniel Klein and Aaron Orsborn (2009) went through the history of the concept of coordination, pointing out a clear temporal and semantical divide, originating from Thomas Schelling's work on game theory. The first acceptation is what they call *concatenate coordination*, building on the transitivity of the verb "to coordinate". Born into the analysis of the firm, it refers to the tuning of each individual onto a *plan*. The autonomous, seemingly unconscious – on parts of individual actors – emerging coordination produces a pleasing result, not exclusively at the economic level. Initially, this kind of coordination was thought as centrally planned, "top-down", since Friedrich Hayek brought the concept to a wider, systemic reach, introducing the idea of a spontaneous order arranged by prices (von Hayek, 1933). It does not mean that the firm perspective faded away, indeed the two (system and firm layers) coexisted, but some incongruence emerged as Robert Coase did note:

> "Marshall introduces organization as a fourth factor of production; J.B. Clark gives the co-ordinating function to the entrepreneur; Knight introduces managers who co-ordinate. As D.H. Robertson points out, we find "islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk." But in view of the fact that it is usually argued that co-ordination will be done by the price mechanism, why is such

organization necessary? Why are there these "islands of conscious power"? Outside the firm, price movements direct production, which is co-ordinated through a series of exchange transactions on the market. Within a firm these market transactions are eliminated, and in place of the complicated market structure with exchange transactions is substituted the entrepreneur-co-ordinator, who directs production. It is clear that these are alternative methods of co-ordinating production. Yet, having regard to the fact that, if production is regulated by price movements, production could be carried on without any organization at all, well might we ask, Why is there any organization?" (Coase, 1937)

The second acceptation is that of *mutual coordination*, referring to the intransitivity of the verb "to coordinate". In this case, the subject is the individual agent (rather than the system or the firm as a whole) whose strategies seek the mutual benefit from the interaction with other individuals. This mutuality operates mainly via conventions and traditions – that is, around a focal point and in virtue of common knowledge (Lewis, 1969). Mutual coordination is a game theoretical operationalization of the wider concept of coordination, whereas concatenate coordination reflects more what we now address as social welfare or system efficiency.

However, the fruitful discussions around the coordination issue did not characterise the modern paradigmatic approach to economic analysis. What instead became the traditional attitude towards market coordination is the analytical apparatus of General Equilibrium Theory (GET) firstly theorised by Leon Walras in *Éléments d'économie politique pure* (Walras, 1874). As Watson (2005) argues, if in the first three editions Walras theorized a within-time groping process of adjusted prices towards equilibrium by part of market agents, in the last edition the author opened to the idea of an *ex-ante*, a-historical moment where demand and supply intentions meet, equilibrium prices are set, then production begins. In this process, distributional neutrality is assumed, therefore justice and equality follow by assumption. Moreover, coordination is no more a procedural achievement; instead it is a postulate. Even though Walras declared that his exercise was purely artificial, prone to demonstrate the logical possibility of equilibrium in a market society, this approach became *de facto* paradigmatic in pure economics, settling aside any spur of moral philosophy. The achievements of more than one hundred years of economic science testimony the beneficial consequences of employing a mathematical and rigorous approach to economic matters, especially in the field of policy evaluation. However, all of this inheritance resides on the theoretical assumption that coordination happens *ex-ante*: it postulates that markets work. Instead, Watson rightfully states, the inability to demonstrate mathematically the existence of a procedural equilibrium by part of Walras should lead to the conclusion that there is no reason to believe that markets coordinate autonomously. The tradition of neoclassical, Walrasian GET totally forgot the existence of concatenate coordination, focusing instead on the theoretical consequences of different kinds of bargaining under various circumstances, in an exclusively mutual coordination attitude of thought.

At the end of Section 3, I will suggest that the implementation of a theory of socio-economic interactions provides the ground to explain the variance in systems' innovative capabilities, hence in their economic performances, defined as a coordination process of the relevant and necessary (technological) knowledge and information. Far from being a brand new statement, such conclusion is, instead, a logical consequence of the past evidence I will deploy in Section 2 and 3. Indeed, whereas Section 3 will introduce a proposed theory of interactions, Section 2 will recall some important contributes of the economics of knowledge literature necessary to appreciate the centrality of interactions in the knowledge production, diffusion and appropriation processes.

2. The Economics of Knowledge

The most important contribute economics of innovation gave to economics core theory has been pushing hard on the idea that *change* cannot be dismissed merely as exogenous. Indeed, the most significant evolutions in human societies happened when new technologies – new ways of doing things and perceiving them – stepped in, as Joel Mokyr (2002) outlines. Technological creation and diffusion are not a matter of *manna* from the heaven, rather they are food for social scientists. At the very ground of innovation studies stand the vast literature on creation, diffusion, management and characteristics of knowledge. The terms *knowledge* and *information* refer to different domains. The word *knowledge* refers to technological and/or scientific knowledge, rather than to Samuelson's acceptation of structured information about others' beliefs (Samuelson, 2004). Cristiano Antonelli describes the Arrowian acceptation of knowledge as "a broad array of overlapping activities including research and learning, competence, experience, know-how, as well as information about technological, organisational and scientific procedures that can be embodied both in tangible and intangible products ranging from capital goods to services. As such knowledge exhibits varying levels of tacit and codified contents embedded not only in protocols and routines but also, and primarily, in skills" (Antonelli, 2016).

Two of economics of knowledge godfathers are Richard Nelson and Kenneth Arrow, who firstly formalized the idea of knowledge as a non-ordinary good qualified by idiosyncratic characteristics, such as limited appropriability, poor excludability and, consequently, low reproduction costs (Arrow, 1962), together with the implications following from these peculiarities on competitive markets (Nelson, 1959). The "Arrovian hypothesis" predicts knowledge underproduction induced by the inability of producers to fully internalise revenues. Knowledge is then appreciated as a public good. However, Zvi Griliches' seminal contribution shed light on the positive side of non-appropriability, namely knowledge externalities affecting the agent's capacity of producing new knowledge positively (Griliches, 1979). The peculiarity of knowledge entails its non-rivalry, meaning that it can be used all over again once created at much lower costs than its production cost. Griliches suggested the formalisation of a knowledge production function (KPF) stating that new knowledge creation rests upon R&D activities as well as external knowledge acquisition. Therefore, if new knowledge partly spills out of the producer's hands, it will enlarge the potentiality of third parties in creating novelties. Hence, from the Arrovian perspective knowledge externalities are a disincentive for production, whereas in the KPF they entail a production cost reduction: a potential trade-off is in place (Antonelli, 2017).

Innovation and knowledge are not separable entities in a systemic analysis of growth. Once Robert Solow (1957) identified the role of TFP in explaining growth patterns of U.S.A. in the first half of the century, the production of technologies or the implementation of production processes, able to enhance factors productivity, became the central issue for the new economics of innovation's agenda. New technologies, new processes – in a word, *novelty*, depends on knowledge creation. The literature originating in Grilliches' contribution widely empirically demonstrated the correlation between productivity and knowledge exploitation thus making of knowledge externalities a key ingredient for growth (Antonelli, 2016). Probably Crepon, Duguet and Mairesse (1998) building on earlier research by Pakes and Griliches (1998), have done the most notable work in this respect, showing, at the firm level, the circular feedback mechanism that links R&D activities, innovation creation and productivity growth. The so-called CDM sets forth the systemic feature of knowledge creation.

At this point, knowledge has been appreciated as a partially public, partially private good, with the former causing both positive and negative consequences on production. From a systemic point of view, knowledge has proved decisive in explaining growth trends via productivity enhancements. A great thinker as Karl Polanyi actively contributed to refining further the definition of knowledge in economics, introducing the concept of *tacitness*. However, Polanyi never refers strictly to techno-scientific knowledge. Instead he is concerned with the epistemology of discovery, therefore referring to *knowing*. Indeed, Polanyi (1958) is a book on language, communication, the motivation for research efforts, written in opposition to the positivistic approach in science. From the realm of language, the definition of tacitness drifted to epistemic considerations on the nature of knowledge, embodying the critique to the algorithmic approach to knowledge creation advanced by the "enculturation" model of scientific activities. The former asserted the existence of a universal, accessible and declarative source of knowledge, whereas the latter suggested the primary importance of contextualization to frame knowledge creation and transmission, i.e. in order to define knowledge (Cowan, David and Foray, 2000). In modern economics, Polanyi's teachings generated a strict dichotomy, namely tacit versus codified knowledge, reducing the potential richness of the categories to a simplified antagonism involving a component of knowledge as a proprietary good, which can be easily articulated and transmitted, and a tacit component embedded in owner's expertise, beliefs and insights. The latter component cannot be sold, but, sometimes, can be absorbed, depending on the absolute ability of the individuals involved and on the relative cognitive distance between sender and receiver (Cohen and Levinthal, 1990).

As Cowan et al. (2000) outline, the New Growth Theory pioneered by Romer (1990) and the new innovation studies following up attribute a pivotal role to the so-called stock of knowledge (and to the knowledge base) of a territory in order to account for territorial growth rates. The hidden assumption entails universal access to such a stock, therefore falling under the category of codified knowledge in an algorithmic sense. However, if knowledge is, instead, contextual, whether it is codified or tacit depends, indeed, on the context. First, a rationale to set the boundary between the two categories is a classical costs evaluation: certain skills or know-how might be articulable, but it might cost too much to operationalise and, subsequently, codify them from the owner's standpoint. Certainly, there is some capacity that cannot be articulable. Second, possessing a code does not mean being able to use it. Such a statement opens up to the idea of complementarity between different *knowing*, e.g. know-what and know-how. It follows that hardly we can define knowledge as a universal stock; therefore, a problem of access arises.

This is just a part of the articulated topology of knowledge developed by different publications of Cowan, David and Foray, but serves to account for a literature plugging in heterogeneity concerning the use of knowledge in organizations. Moreover, it introduces the issue of the access to that stock of knowledge so valuable for new knowledge creation, as pointed out in the KPF approach. Indeed, the context-dependent quality of knowledge, its complementarity, together with its heterogeneity require the analysis of the mechanism governing its transmission and use. A theory of interactions and networks at the societal level can accommodate this necessity.

In order to complete the picture and push further the definition of the concept, we must go through a theory of technological knowledge creation and evolution, i.e. dealing with the heterogeneity in production. The two pillars, here, are Martin Weitzman and Brian Arthur, the former being the path-breaker but leaving to the latter the task to focus and synthesise a theory. Weitzman suggests, referring to the application of the first achievements of the economics of knowledge within the New Growth Theory inaugurated by Paul Romer,

> “[..]the unopened black box contains the "production function for new knowledge"-conventionally postulated to be of one or another reduced form but never really given proper microfoundations. "New ideas" are simply taken to be some exogenously determined function of "research effort" in the spirit of a humdrum conventional relationship between inputs and outputs.” (Weitzman, 1998)

Weitzman proposes, instead, a theory of *recombinant* or combinatoric creation of new knowledge, in analogy with new cultivated varieties creation in agricultural stations. Novelty is the result of a selection of fertile combinations over the unbounded set of possible combinations of old ideas, most of them being useless. Creating new technological knowledge appears, therefore, as a strongly path-dependent process resting on the stock of past knowledge produced. In Weitzman’s growth theory, “the ultimate limits to growth may lie not so much in our ability to *generate* new ideas, so much as in our ability to *process* an abundance of potentially new seed ideas into usable form” (Weitzman, 1998).

Arthur (2009) has the virtue of providing a comprehensive theory of technological creation and evolution. The baseline is that every technology is meant to solve some problem, and, in order to do so, it exploits some scientifically discovered natural phenomena. Every invention is a response to a need and the result of a combinatoric process involving precedent technologies. Moreover, each invention becoming a technology contributes to creating new problems to be solved or new needs to be perceived, so that a new combinatoric research effort is put in place. Eventually, some line of technological development may die out when all problems at present are solved, leaving place for other, new ventures to be tackled. In short, technological creation is both a recombinant and a *recursive* process involving existing technological knowledge.

How do this recombinatory and recursive process happen? In other words, how do ideas meet and fuse in an unprecedented way, contemporarily occurring with a need to be satisfied by them? There is not a unique and comprehensive way to treat such a topic. From a meso or macro perspective, the answer the literature gives is the externalities mechanism. Co-location in space provides inhabitants with the opportunity to experience positive pecuniary externalities, making the access to new knowledge easier and cheaper (Gehringer, 2011). Altogether, co-location might provoke encounters with a heterogeneous pool of other agents and resources, i.e. it creates the opportunity for diversities to match (Jacobs, 1969). It is clear, however, that answering from this perspective does not really satisfy the questioner because the word "externalities" is a black box. This is why the economics of knowledge had gone deeper into microanalyses of knowledge exchange. Stefano Breschi and Francesco Lissoni (2009) showed that mobility of inventors is the main driver for knowledge externalities measured as patent citations, significantly reducing the role of spatial proximity promoted by the pioneering research of Adam Jaffe, Manuel Trajtenberg and Rebecca Henderson (1993). Breschi and Lissoni (2009) paved the way for a branch of the literature paying particular attention to the role of networks in shaping knowledge diffusion.

Summing up, knowledge peculiarity regards the quasi-public and quasi-private nature, setting the dynamic for a potential trade-off on the verge of the level of appropriability, with lots of implications for optimal production at the system level. Further, knowledge is recognised as a heterogeneous good both in use, because of the tacit/codified context-dependent dialectic, and in the production, given its combinatoric, path-dependent and recursive nature. At both levels, interactions among knowledge agents are pivotal. On the first level, interactions candidate as the mean to access the stock of knowledge embedded in a territory, bridging tacit and codified knowledge, thus enhancing a more extensive exploitation of the resources at disposal. On the second level, interactions provide agents with the necessary bundles of unpossessed knowledge to be recombined recursively. The two levels clearly intertwine. Without interactions, agents' reach for new potential knowledge components would severely dampen, and even if acquired, lots of knowledge would be useless (or at least less productive) because of the lack of tacit complementary components. The interactivity permeating the whole process of knowledge production leads us finally to recognise knowledge as a collective good (Antonelli, 2000).

3. A *trait d'union*: a theory of interaction between the economics of information and knowledge

In any discourse on the role of interactions among agents in economics, a due credit must be rewarded to the ground-breaking contributions of Joseph Stiglitz, Michael Spence and George Akerlof, who, with their work, paved the way for a new paradigm in economics to rise: the economics of information (Stigler, 1961; Akerlof, 1970; Stiglitz, 2000; Spence, 2002). This paradigm developed the consequences of a full account of the presence/absence of information for neoclassical equilibrium theory formally. One of its most important takeaways is that possessing – or lacking – information influences a person's perception of reality, i.e. the set of prospect choices, and actual behaviour (Stiglitz, 2002; Samuelson, 2004). Other than the amount and kind of information owned by each individual, the channels for information to pass from hand to hand emerged as a crucial feature of every socio-economic system. The possibility and conditions of access to information proved essential characteristic to understand the evolution of structural properties of economies such as unemployment or income inequalities (Ioannides and Loury, 2004; Calvó-Armengol and Jackson, 2007).

A literature has grown to tackle the role of social structures in disseminating information, and even further, towards models of endogenous network formation. Out of this stream, and probably one of the first researchers to take the consequences of imperfections in information, behaviours and systems seriously, is Alan Kirman. Notwithstanding the fact that the information paradigm has been the most severe and disapproving approach to neoclassical thinking in economics from the mid '90s, Stiglitz and colleagues' research mostly tried to maintain the neoclassical analytical toolset, framing the agenda in order to look for equilibria (which has not been an easy task). Alan Kirman, instead, took a much more aggressive attitude towards core theory, pushing a little forward the achievements of the information paradigm, plugging it onto the concepts of networks, heterogeneity and learning, adaptive and complex systems. Probably, Kirman would not place his name among the ranks of information economists, if any such a label conveys some informative content. Indeed, Kirman's reference is the complexity theory applied to socio-economic systems.

One of the core critics moved by Kirman towards the general equilibrium theory is about the inconsistency of the representative agent in macroeconomic models, stemming from the intrinsic heterogeneity of individual agents (even if he is not the only one to stand this position, as Blundell and Stoker, 2005 survey). As he states,

> "there is no plausible formal justification for the assumption that the aggregate of individuals, even maximizers, acts itself like an individual maximizer. Individual maximization does *not* engender collective rationality, nor does the fact that the collectivity exhibits a certain rationality necessarily imply that individuals act rationally. There is simply no direct relation between individual and collective behavior." (Kirman, 1992)

The mechanism linking individuals' behaviour and systemic trends is networking. Individuals interact for a variety of reasons, consequently forming networks on which the economic activity builds on.

> "A modern economy is characterised by interaction, both direct and indirect, between individuals. Three aspects of this interaction are important. Firstly, there are the different ways in which individuals interact. Secondly, there is the fact that agents learn over time, from their previous experience, about the consequences of particular interactions with other individuals. Thirdly interactions take place through networks. These networks may involve trade, observation or active communication." (Kirman, 1997)

Most importantly, networks evolve over time because their nodes – the agents – evolve, i.e. they learn. Kirman dismisses the rationality axioms typical of Arrow-Debreu General Equilibrium Theory, modelling individuals that exchange information through networks, learn from experience and from nearest neighbours, finally gauge their choices according to procedural and bounded rationality relying on the possessed information. A strong feedback mechanism between micro and macro dimension is in place (Kirman, 1997):

> "My basic argument is that interaction changes the nature of the relation between individual and aggregate behaviour and the nature of aggregate behaviour itself." (Kirman, 1997)

Kirman's contribution to the information paradigm is in that he introduces *creativity* in agents' behaviour as a consequence of *learning*, which is the mechanism for information internalisation. Information flows do not become behaviour following an incremental process of accumulation and digestion. Instead, un-deterministically, such flows induce individual creative reactions to the alteration of the settings, i.e. the amount of information in possession of an agent to choose. In turn, unexpected reactions intensify and perturb information flows activating a cascade effect among economic agents, so that the whole system is unsettled and in motion. Learning processes of evolving but rationally bounded individuals generate a coherent set of information to interpret complex environment and take decisions. Decisions produce new information and, in parallel, affect the structure of the economy, i.e. the network, affecting the way information is collected and thus how learning takes place. Recently, Matthew Jackson provided sizeable empirical evidence of the role of networks in modern societies, from criminal activity to inequalities, cooperation and public good production (Jackson, 2014; Jackson, Rogers and Zenou, 2017). Jackson remarks that networks affect economic behaviour and are endogenous respect to it.

Both the economics of information and the economics of knowledge, developing the understanding of their topic of interest (information and knowledge) and the intricate and complex relationships between these and the other factors characterising socio-economic processes, acknowledged the centrality of interactions to understand information and knowledge dynamics. If the economics of information and knowledge started as subsets of the discipline, with the declared aim of turning some problematic specific aspect of mainstream out, the apparatus (that is both analytical and substantial) they generated stretched back enlightening economics as a whole. Economic crises, globalisation, but also more intense fertilisation across disciplines evidenced the inadequacy of neoclassical theories of growth, as well as of its micro foundations. Studies about information, behaviour, knowledge, communication, declined at various levels of analysis (from international trade to local governance) provided – and are providing – new tools for empirical analysis and policy induction.

The purpose of this section is to outline a simple theory of interactions unifying the achievements of the two paradigms. In order to do so, I start with a simple question of a practical nature: what do interactions convey? I will try to articulate on four contents I judge to be the more encompassing set. Each element of this set is connected with the others in a mutual reference; however, each of them has its autonomy.

1) The first content is the more classical and trivial: *commodities*. Not even a thousandth of social interactions in the world happen because of exchange of commodities, but these few have particular importance for economics. Neoclassical economics is used to depict any form of market exchange as pure and frictionless; namely, a transaction. However, in between the two extremes of

physical interactions and pure transactions, the category of transaction-cum-interaction particularly suits for the economics of innovation: these are occasions of market value exchange entailing some form of personalized interaction between agents (Antonelli and Scellato, 2013).

2) Second, *knowledge* is exchanged via interactions, defining interactions as inventors teams, enterprise-university collaborations, or embedded in hierarchies (Gehringer, 2011). Breschi and Lissoni (2009) explicitly stated that inventors' mobility is the *modus* of knowledge diffusion and insemination across regions. Wutchy, Jones and Uzzi (2007) brought evidence of the dominance of team in academic and patent production. Economics is progressively paying attention to the internal dynamics of teams (the list of reference is vast; see Harrison and Klein, 2007; Boschma and Frenken, 2010; Häussler and Sauermann, 2014 as some examples) recognising that the typology of interaction influences knowledge exchanges, whether it is codified or tacit.

3) From the Nobel lecture of Stiglitz (2000): "It is not only that the market for information is markedly different from the market for apples or oranges or chairs, but also that information issues are intertwined with the production and sale of traditional commodities. In traditional economics, prices convey all the relevant information (between consumers and producers, say about the scarcity value of resources). We now realise that there are a variety of other ways in which economically relevant information is conveyed and that prices convey information other than that about scarcity": *information* is our third item. The difference between knowledge and information is that knowledge is a procedural content while information is behavioural. Knowledge has to do with new content creation, while information pertains to the domain of successful exchange (coordination). Once again, knowledge is a matter of production; information is a matter of organization. Neither of them spills freely in the air, rather they are embedded in tokens or individuals and they mobilize via interactions.

4) Last but not least, interactions contribute to the creation of sociality. The main core of economics favour an epistemological approach centred on the *ego* – the very same novel approach named *network analysis*, clearly sponsoring a more systemic view on social phenomena, is practically more focused on the centrality of the *ego*. Such a perspective derives from the methodological individualism inheritance of social science, which has never been proved wrong (but it has been hardly criticised, e.g. by Durkheim). However, such a philosophical foundation of social sciences enquiries should not impede us to appreciate the fact that individuals do not move in a free space, and interactions do not just convey economic value – in many forms as information or knowledge

or commodities – but they contribute to the creation of a societal texture or a *milieu* for other exchanges to happen. Sociological studies first drew the very fertile, multifaceted (and criticised) concept of *social capital*, which may help in understanding this fourth quality of interactions. Both that we embrace the instrumental definition of social capital as resource to a mean (*à la* Bourdieu, but also Coleman and Burt might reside in this alley) or one more oriented toward civicness (*à la* Putnam and Fukuyama) the rationale is that interactions among members of a community provide the necessary endowment for *other things* to happen (Trigilia, 2001; Durlauf and Fafchampes, 2004; Tosini, 2005).

As anticipated, the four pillars interlace. The *ego*-network can be enforced to access information, which in turn may be useful to plan new knowledge production or to engage in strategic bargaining for market transactions. In turn, every transaction entails some measure of information collection and exchange; territories where markets do work intensively might favour the proliferation of non-market interactions fostering the creation of social capital. Embracing altogether the four items possibly conveyed by any interaction, it appears why an accurate study of their happening, intensity, modality, is valuable for the economics of innovation. Out of the four of them, only the first two have been subject to attentive studies, via the full exploration of the concept of *externalities*. On the one side, the very idea of increasing returns stemming from innovative efforts in industrial districts builds on the voluntary and involuntary interactions taking place between entrepreneurs residing in the district, driven by the intensity of transactions. Thus, we speak of transactions-cum-interactions when transactions pave the way to knowledge spillover. On the other side, and in parallel, a vast literature spent efforts to define the difference between Jacobsian and Marshallian knowledge externalities, discussing which is better for the innovative outcome, i.e. if it is better specialization and agglomeration or heterogeneity and diffusion.

Up to now, in the economics of knowledge, interactions among individuals have been represented solely as a channel to exchange a (whatsoever special) product. Interacting individuals are those who have access to more resources to be exploited in the production process. Interactions are typically instrumental and *ego* centred. However, such a vision is narrow-minded if exclusive, because it neglects the other two items: information and sociality. Both of them have significant consequences for the environment where transactions-cum-interactions take place. In order to fully understand the issue, we ought to recall the concept of knowledge *complexity*. The recent development of the economics of knowledge and innovation provided a complexity theory of knowledge creation (Arthur, 2009) claiming that the mechanisms underpinning the spurring of new knowledge qualify as complex dynamics. In a nutshell, generating new knowledge is not just a matter of juxtaposing knowledge pieces following a rule, instead it is a process involving the observation of natural phenomena, the (fortuitous or strategic) recombination of existing knowledge, the presence of appropriate incentives and gaps to be filled. The whole process happening in a recursive fashion, whose chances of success are unpredictable. Evidently, such a process cannot be coerced to fit into a mere exchange of resources among players.

Lane and Maxfield (2005) suggested that

> "[..]ontological uncertainty is endemic in innovation processes, because innovative agents participate in constructing large-scale transformations of agent-artifact space, and the transformations that result from their interactions do not, indeed cannot, correspond to intentions at the level of individual actors"

Agents involved in a creative process face a complex foresight horizon activity (Lane and Maxfield, 1997) which is an activity where the attributes assigned by each individual to their interpretative structure of the world (their narratives) lose coherence or opens up. This may come for good, as they state:

> "With this opening, it may happen that participants can jointly construct *new* meanings. Generating new meanings, particularly new attributions of functionality for artifacts and new attributions of identity for agents, is an important part of innovation. *Becoming* uncertain in this sense may be consistent with actors' directedness, in interaction streams in which participants seek to explore their attributional heterogeneity, without precise foreknowledge of just where the heterogeneity lies and where its exploration may lead."

Activities characterized by complex foresight horizons can be better handled in the frame of a relationship, they say, which eventually may become *generative*. The generativity of the relationship stems from the empowered capacity of the interacting individuals to build new interpretative structure of reality, i.e. to see old things with new eyes. Generativity has five prerequisites, according to Lane and Maxfield, specifying that there must be some commonality of intentions, benevolence in approaching the other's differences, confidence to share doubts and inspirations, and some heterogeneity – whatever kind it is, thus not necessarily in terms of competence or skills. Indeed, the fascinating aspect of Lane and Maxfield theory of innovation is that it does not only look at invention and innovation as a resource matching, as most of the economics of knowledge does, rather they picture it as a behavioural and cognitive *experience*.

The concept of information may help to integrate the "narratives approach" formulated by Lane and Maxfield into a theory of interactions for innovation dynamics. Indeed, the information paradigm in economics opened up to the idea that individual choices depend on the set of information the individual gather, pertaining prices (and quality) but also other's choices (and preferences) or the characteristics of the environment surrounding (and the trust they put in it). The individual's set of preferences defining utilities does not coincide with the individual's interpretative structure of reality, however, they overlap to some extent. At the cross point of these two theories, information flows in a social system take up two roles: 1) they serve as a meaning for coordination and 2) they affect behavioural attitude. In both cases, interactions are the vessel. In both cases, they are pivotal for innovation dynamics. On the first account, Antonelli (2006) provides a dictionary for translating the information paradigm into innovation studies' terms, and specifies the role of information in governing knowledge creation and diffusion; indeed

> "Neither pure markets nor pure hierarchies can provide the necessary levels of coordination and division of labor. An array of knowledge governance mechanisms, ranging from coordinated transactions and constructed interactions to quasi-hierarchies, however, has progressively emerged according to the characteristics of knowledge and the costs of using markets and organizations".

As for the behavioural effect of information flows, Lane and Maxfield's theory of narratives provides the background to plug in the behavioural considerations formulated by information economics into a theory of interactions for innovation studies, as sketched above.

Furthermore, interactions do not only convey knowledge and information. As introduced in the description of the fourth item, they might build sociality, which can take two interrelated forms from the social sciences' analytical point of view. On the one side, a vast literature on applied network analysis found its mouth into the ocean of economics in the recent articles of Jackson (2014) where the author states that networks are endogenous to economic activity because humans are inherently social beings, i.e. interactive entities. On the other side, the literature on social capital leads us to appreciate the societal effect of interactive milieus, supporting trust and compliance, reducing transaction costs thus permitting information externalities to move smoother.

In conclusion, we should dedicate more effort to theoretically frame the role of interactions in economic activities, especially dealing with reactive and creative behaviour. Even if Schmookler (Schmookler, 1954, 1957) claimed that the lone genius is not disappearing, evidence in patent and academic activity demonstrate the absolute centrality of interactions in knowledge creation (Wutchy et al., 2007). The significant contribution in these Schmookler's works has been highlighting that innovation does not happen freely in space because it is embedded in the social space. There is no doubt that strict individualism has proved fundamental for the development of economics, however, recent advances in research suggest that it might be time to shift from a representative *agent* to a representative *relationship*. We should move on from considering interactions only as instrumental, to study them as endogenous, endemic, intrinsic of socio-economic systems (Grazzini, 2013; Landini et al., 2015; Arata, 2017). It means to abandon the ideal reference of smooth markets, perfect information and rationality when theorizing about real world functioning. Coase (1937, 1960) best represent such ideal referencing that excludes a priori the intrinsic socio-economic value of interactions. Indeed, the internal organization of production, which may rephrase as teamwork in the inventive setting, is understood solely as a consequence of market imperfections that cannot be straightened, hence inefficient by definition (similarly to Jones, 2009). On the contrary, the recent researches outlined above induce to consider interactions – organization, teamwork – as an intrinsic, emergent, necessary and *generative* characteristic of economies.

3.1 Interaction loci

Up to now, I sketched an analytical definition of what an interaction is. Another relevant question is: where do interactions happen? The relevance of this question comes from the fact that the *locus* of an interaction gives important hints on the reasons why it happens.

A tripartite categorization might come up useful to localize interactions in the marketplace: on the two far extremes, we find pure transactions and pure interactions, separated half-way by transactions-cum-interactions. Pure transactions are those market operations where no individual is directly involved in bargaining or value exchange. Financial markets are their prototypical *locus*, even if it is easy to cast doubts on the pertinence of such an assignment, given the acknowledged centrality of stock exchange centres in leading financial traffic. Millions of operations happen in a pure transaction form, but the market as a whole is not a totally disembodied *locus*. Another example is online retailing, which escalated in the last decade. On the opposite, pure interactions are those punctual or iterative relationships that do not entail necessarily market value exchange, i.e. interactions *not necessarily with* transactions. As for the *locus*, firm's intra-muros relationships fit the case. Workers interact in the workplace because of a variety of purposes: in order to accomplish a demanded task or in an unsupervised, purely socialization-oriented way, or because of learning purposes. Even if these relationships do not entail market value exchange, they happen within the market and influence its functioning. Lastly, transactions-cum-interactions are those situations where two agents exchange value in a personalized manner. The transaction is not necessarily oriented to interact, but it necessarily ends up in it in order to reach an accomplishment. Inter-firms relationship is the *locus* of such occurrences, which is a label covering a wide range of events: from industrial districts commercial activity to industry-academia collaboration and R&D joint-ventures.

Intra-muros and inter-firms relationships are the two *loci* of interactions, accepting a strict definition of pure transactions as disembodied occurrences. Interactions do not happen in free space; instead, they are institutionalized instances within the market. This is important because institutions are not just contexts: institutions influence the direction and motivation of an action. Therefore, we cannot look at *loci* as simple backgrounds to individuals' extrinsic motivation to interact; rather they have a *functional* role in orienting interactions. Moreover, as outlined in the previous section, individuals typically move onto the ties they created in previous interactions – or that they inherited from companions. Once again, there is not such a thing as a *free* space, rather interactions happen along the structure of the *social* space, which we stylize with two pillars: *functional* institutions and *structural* networks. We must mind that the social space is not fixed – it is inherently dynamic in time.

In order to appreciate such stylization, we should reintroduce the topic of most interest in this note, i.e. knowledge creation and dissemination as the principal propulsion of innovation. Because of the peculiar characteristics of knowledge, which is sticky, partly or totally tacit, non-exclusive and non-rival, pure transactions are not the suitable medium through which a positive knowledge dynamics may activate. As suggested above, some form of interaction is needed for coordinated efforts to take place and for recombination dynamics to emerge: this makes knowledge a collective, rather than public, good.

According to this perspective, firms make use of networks because of three reasons: 1) reduce transaction costs to enter and use the market for knowledge; 2) collect information *about* knowledge; 3) acquire (which is a matter of availability) and absorb (which is a matter of capacity) new knowledge. Transactions-cum-interactions accommodate these three necessities, happening within a set of formal institutions such as industrial districts, knocking down absorption costs. Within a firm, instead, individuals exploit their network 1) in order to collect the relevant information to coordinate the research effort of the firm; 2) steer the screening of the firms on the market; 3) carry out the recombination of knowledge, thus fostering new knowledge creation. The networks in the two *loci* are different, i.e. the structure of the social space is different: the relevant network for intra-muros knowledge creation is the ego-network of the individual, while the crucial relational structure for inter-firms transactions-cum-interactions is the ego-network of the firm. The two might be causally linked if it is a personal link of an inventor to steer a collaboration between firms, but the differences in the nature of the interaction – motives, context, objectives – remain because the institutional setting differs. Such a differentiation merits attention because, as we introduced above, not every relationship gets generative, and generativity is a crucial boost for dealing with the complexity of knowledge production. Moreover, the costs associated with knowledge spillovers, opportunistic behaviour related to non-appropriability of knowledge, and coordination in the production of new knowledge are heterogeneous in these settings, and so it is the value of information.

Summing up, we suggested that almost every invention, as a proxy of knowledge creation, occurs in an institutional setting; that interactions (pure or within transactions) are crucial in the process of knowledge coordination and recombination, and, furthermore, that the nature of every interaction depends on the social space it happens in. In these processes, the idea of externalities gains centrality. Even if much of market relationships are normed by contracts, the bulk of value added comes from knowledge externalities overcoming barriers of compositeness, indivisibility and tacitness of knowledge, as well as from informational externalities lowering transaction costs and signalling future developments of the relationship’s outcome. Interestingly, the complex and interactive process of new knowledge creation may lead to unforeseen outcomes that cannot find a satisfactory level of appropriability within the hosting institution. A more appropriate institutional setting might emerge, in the form of a new enterprise, or a new contract. According to this perspective, new firms (in knowledge-intensive sectors) sprout in order to generate a more accommodating social space for the appropriation of present and future returns of the inventive dynamic. The functionality of the institutional setting emerges more clearly now. The connection between interactions and institutional environment, as intra-muros activities, happens to be bi-directional.

It is clear, then, that the social space is a dynamic entity – interactions trespass time constraints as well as institutions do. Hence, webs of interactions and the institutions looking after them might be considered as the vehicle for synchronic and diachronic externalities too. Indeed, if the continuity of the knowledge recombination process can be described in terms of diachronic change of the social space, the gatekeeper of the flow of knowledge during time is the web of iterative interactions, which is, *de facto*, a measure of intensity. Decoupling the concept of externalities from that of interaction voids the former of procedural content. Ideas do not spill free into the air, rather they travel via some unintended medium at a lower cost than market price (if a dealing with a price can be guessed). This medium is the social space made of institutions and webs of interactions, which are complements rather than substitutes: every individual agent is embedded, simultaneously, in a variety of levels where knowledge and information flow, that may be ascribed to the institutional or interactional facet of the social space.

3.2 The issue of coordination

As anticipated in the Introduction, once a theory of interactions has been spelt out, moving from micro-level considerations to the organizational and system levels, the delicate theme of economic coordination may be handled. The economics of information vastly debated this topic. Indeed, both Kirman and Stiglitz repeatedly asserted that unemployment is a typical coordination failure consequence (Kirman, 1997; Stiglitz, 2000). The whole information economics agenda headed towards the claim that, even within that mathematical apparatus of GET, information asymmetries and flows make many equilibria or none possible, breaking the assumption of ex-ante unicity. On the other bank, economics of innovation and knowledge suggested the existence of disequilibrium dynamics, originating from creative reactions of individual agents, which might modify the environment following an evolutionary, territorial, path-dependent trajectory (Hanusch and Pyka, 2006). Recognizing that innovations modify the environment where production takes place – they affect factor prices but also the ways of exploiting them – and that individuals react to the surrounding conditions in possibly creative manners, the theoretical tools for studying disequilibrium as an ontological state of economies are set. From a coordination perspective, disequilibrium means that the forces driving economies do not match and clear out, leading to unemployment, distributional inequality, growth differentials, poverty niches, both at the regional and international level (see Bogliacino, Piva and Vivarelli, 2012 and Vivarelli, 2014 for some examples of the interest of the literature).

The achievements of the information paradigm in terms of individual's choice theory and rationality are the basis to understand the development of a theory of knowledge creation and diffusion as a reactive, creative process nurturing innovation outbursts. Models of knowledge governance for innovation systems are one-armed without accounting for the role of information flows in shaping behaviours. Inversely, if the law of motion of an economic system is innovation via knowledge creation, the information paradigm is enriched of a dynamic element, propeller of signals and asymmetries, to be taken into account in order to provide comprehensive behavioural models. The insights both paradigms generated converge in the refusal of ex-ante coordination and market clearing equilibrium postulates (Stiglitz, 2002).

However, does the existence of miss-coordination mean that no coordination is possible? The observation of reality points out differences in levels of coordination across countries. The proposition of this essay is that part of this variance might be explained by a theory of interaction brought at the system level. The Lane and Maxfield (2005)'s scaffolding theory makes the purpose. In their treatise of the central role of ontological uncertainty for innovation, the authors theorize that market systems emerge through the construction of scaffolding structures. These structures serve for controlling the continuous entering in the market of new agents and related artefacts, and, primarily, as a place for agents' ontological uncertainties to meet and stabilize. In other words, what generative relationships do at the micro level of individuals interactions, scaffolding structures do at the system level: the otherwise paralysing uncertainty of knowledge creation and implementation becomes bearable within environments characterized by partially closed, oriented interactions. The topology of scaffoldings is enormous: it comprehends unions and voluntary organizations, consumer associations and industrial groups, political parties and universities consortium, incubators and industrial districts. Lane and Maxfield (2005) suggest that scaffoldings structures also act as ambiences where conventions mutate and emerge; in this sense, journals and communication media are other examples of scaffoldings. This description of scaffoldings fits with the idea that interactions convey sociality as a fourth item. The sociological literature on Social Capital is a reference point to start diving into the role of sparse or formal/constrained interactions in distributing socio-economically relevant resources.

Before wrapping the scaffolding structures with the topic of coordination, another concept might find a relevant place in the discussion: that of *emergence*. John Foster and Stan Metcalfe (2012) trace the unfortunate path of this analytical concept in the history of economics, as it was embedded in the Classics – Adam Smith in particular – but basically neglected afterwards until evolutionary economics' resurgence in the late '80s. This conceptual journey resembles that of concatenate coordination, indeed the two do correlate strictly. What Foster and Metcalfe (2012) argue is that economic activity is the history of the continuous emergence of market structures in an open, uncertain environment. This process displays "unpredictability in principle", "genuine novelty" and "irreducibility":

> "The formation of radically new bundles of rules is "genuine novelty" and can take the form of: capital goods (technological rules), productive networks (organizational rules), contracting systems (institutional rules) and human skills (procedural rules). Enacting new bundles of rules involves a process of 'self-organization' which is "unpredictable" with regard to the patterns of structure that ultimately form. Over time, such unpredictability is diminished by a process of 'competitive selection' whereby particular combinations of technologies, organizational structures, institutions and procedures come to dominate. Such dominant structures are, necessarily, "irreducible," because of the irreversible character of dissipative economic systems operating in historical time."

In this complex dynamic, uncertainty is the rule and the law of motion. Individual agent's beliefs and intrinsic motivations nurture the power of his imagination, eventually leading him to creative behaviour. Entrepreneurship becomes the use of imagination to "conjecture that the present economic order can be arranged differently":

> "The provision of new knowledge is widely accepted as a driver of growth but, just as important, is what we do not know. In complex economic systems characterised by uncertainty, ignorance is pervasive. But it is this lack of knowledge that allows imagination to do its work. What matters is that meso-rules exist that permit a knowledge gradient to exist and be exploited. New knowledge formation is a virtual construction project based upon beliefs, hunches and intuitions rather than objective information. And it is the lack of a connection between aspiration formation and objective knowledge that allows emotions to play such an important role".

Therefore, the complexity of the process does not end up in a chaotic mass, instead it fits a temporary order if some "meso-rules" are in place as a conduit, and the "unpredictability" of the system evolution is diminished when "selection" stands on some "patterns of structure".

Here we have all the necessary, coherent ingredients to make a contribution to the topic of coordination. At the one extreme, dismissed the hypothesis of the existence of an ex-ante equilibrium and a representative agent to study it, we now see economic development as a history of emergent rules. Disequilibrium dynamics is the law of motion, originating from the continuous, disruptive agency of innovations. On the other extreme, abandoned the perfect rationality and information assumption, we recognize bounded rational individuals as embedded in a fickle environment characterized by uncertainty, within which they can learn, react creatively and unexpectedly (exploiting available resources but also the intrinsic power of imagination) finally modifying the environment with their behaviour. Individuals' action is not autonomous or isolated; rather it is plunged in a social space defined by structural networks and functional institutions. Information and knowledge flow through the social space, specifically via individuals interactions, and, at the same time, they are incentives for its maintenance and improvement. Without information flows resources are not allocated efficiently; without knowledge flows innovation is not possible. The feedback mechanism linking the environmental condition and the individuals' reaction is at the core of the continuity of the emergence dynamics. The bridging tile of the two extremes is the scaffolding theory. Scaffoldings are the manifestation of the social space, i.e. they are the conduits where information gathers and is redirected. They are the substratum where relevant information about knowledge is processed so that knowledge agent can activate the recombinant process. Scaffoldings are structures enhancing trust and social capital for transaction-cum-interactions to happen smoothly. Simultaneously, scaffoldings are the organizational systemic entity containing the flux of novelties originating from the system dynamics. They serve as a metabolic domain for digesting the new rules emerging from interacting agents' behaviour, i.e. from innovation.

Put another way, scaffoldings structures are the level where coordination takes place. It is at this level that information about resources reaches potential users, whether resources are knowledge, commodities or information. Without a proper scaffoldings structure, the territorial system could not circulate knowledge in an efficient way, e.g. variety would die out and growth with it (Jacobs, 1969). Alternatively, transaction costs stemming from radical informational asymmetries would make knowledge creation too expensive. The multitude of novel inputs, signals, knowledge components, artefacts generated by innovations (from within or outside the system) would disperse in a chaotic environment. A temporarily stable organizational structure is indispensable for emergent properties to set.

Every territorial system qualifies with a specific scaffoldings structure, consequently showing up a different level of internal coordination of resources, therefore displaying different capabilities to innovating and managing the consequences of innovations. One more relevant question hangs on: are scaffoldings spontaneous market structures, or are they purposefully built by planners? Answering this question would deserve an independent work apart of this essay. However, some considerations can be drawn. First, scaffoldings are not autonomous entities that can be established on a system; rather they emerge from the entrance of new agents and artefacts, with the consequent re-shaping of networks and re-definition of institutions (existence, purpose, positioning in the social space). Interactions maintain vivid the scaffoldings. The complexity of the dynamics cannot be managed and dominated by any system or international institution. Nonetheless, development trajectories dragged by innovation are not smooth or linear, but bumpy and disruptive. The potentiality of a system to absorb shocks and turbulence depends on the quality of its scaffoldings structure. Policy makers may want to favour its formation, consolidation and fluidity. Nevertheless, there is no equation or universal recipe for this achievement, because scaffoldings are highly idiosyncratic.

Concluding, the coordination concept is in need of a restyling in light of the recent advances in information, knowledge, innovation and complexity economics. Certainly, full coordination cannot be assumed, rather its plausibility should be demonstrated. Indeed, if in an equilibrium framework full coordination is logically possible, in a disequilibrium dynamics, instead, it does not necessarily. Instead, like many other manifestations of harmony in nature, some level of coordination could be an emergent property of the system, utterly subject to idiosyncratic factors. In turn, these factors are not only the pure, macroeconomic forces of supply and demand and prices but also the quality of the social space and the moral intentions as a guidance for individuals' behaviour.

4. Overview of the next chapters

Three empirical papers will follow this first chapter, whose role is that of a theoretical cover introducing to more detailed, circumscribed analytical studies. Indeed, each of the three empirical works mirrors one of the theoretical aspects introduced in this first chapter, but they do not pretend to back it with an empirical demonstration. Instead, the empirical papers will bring some evidence of how relevant is to focus on interactions in order to understand the functioning of knowledge creation and diffusion and its economic consequences. In particular, they will develop some insights presented above. Chapter 2 will expand the idea that the team-oriented organization of technological knowledge production is not only a necessary but emergent and effective condition of knowledge creation. It will treat teamwork in inventive activity as a reaction to the intrinsic characteristics of knowledge creation, hence linking the organization of knowledge production with territorial knowledge productivity. Chapter 3, instead, starts from the considerations about the importance of the *loci* where knowledge interactions happen and the embeddedness of every individual in a multiplicity of social layers. It will investigate the role of the different levels where knowledge externalities take place on inventor's knowledge production. Finally, Chapter 4 is set in motion by the latest arguments about coordination and the social structure. It extends the most advanced theory of knowledge externalities appropriation on the firm account – the Knowledge Spillover Theory of Entrepreneurship (Acs et al., 2009) – with the theory of Social Capital. In so doing, it provides a new research avenue to study knowledge resources coordination with a tool that specifically measures the multidimensional notion of interactions.

References

Acs, Z.J., Braunerhjelm, P., Audretsch, D.B. and Carlsson, B. (2009) The knowledge spillover theory of entrepreneurship. *Small Business Economics*, **32**, 15–30.

Akerlof, G. (1970) The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, **84**, 488–500.

Antonelli, C. (2000) Collective Knowledge Communication and Innovation: The Evidence of Technological Districts. *Regional Studies*, **34**, 535–547.

Antonelli, C. (2006) The Business Governance of Localized Knowledge: An Information Economics Approach for the Economics of Knowledge. *Industry & Innovation*, **13**, 227–261.

Antonelli, C. (2016) Digital knowledge generation and the appropriability trade-off. *Telecommunications Policy*.

Antonelli, C. (2017) The derived demand for knowledge. *Economics of Innovation and New Technology*, **26**, 183–194.

Antonelli, C. and Scellato, G. (2013) Complexity and technological change: knowledge interactions and firm level total factor productivity. *Journal of Evolutionary Economics*, **23**, 77–96.

Arata, Y. (2017) Endogenous business cycles caused by nonconvex costs and interactions. *Journal of Economic Interaction and Coordination*, **12**, 367–391.

Arrow, K. (1962) Economic welfare and the allocation of resources for invention. *The rate and direction of inventive activity: Economic and social factors* pp. 609–626. Princeton University Press.

Arthur, W.B. (2009) *The Nature of Technology: What It Is and How It Evolves*, 1. Free Press pbk. ed. Free Press, New York.

Blundell, R. and Stoker, T.M. (2005) Heterogeneity and aggregation. *Journal of economic literature*, **43**, 347–391.

Bogliacino, F., Piva, M. and Vivarelli, M. (2012) R&D and employment: An application of the LSDVC estimator using European microdata. *Economics Letters*, **116**, 56–59.

Boschma, R. and Frenken, K. (2010) The Spatial Evolution of Innovation Networks: A Proximity Perspective. *The handbook of evolutionary economic geography* (eds R. Boschma), & R. Martin), pp. 120–135. Cheltenham: Edward Elgar.

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

Calvó-Armengol, A. and Jackson, M.O. (2007) Networks in labor markets: Wage and employment dynamics and inequality. *Journal of Economic Theory*, **132**, 27–46.

Coase, R.H. (1937) The Nature of the Firm. *Economica*, **4**, 386–405.

Coase, R.H. (1960) The problem of social cost. *The journal of Law and Economics*, **56**, 837–877.

Cohen, W.M. and Levinthal, D.A. (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, **35**, 128.

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

Crépon, B., Duguet, E. and Mairesse, J. (1998) Research, Innovation And Productivi: An Econometric Analysis At The Firm Level. *Economics of Innovation and New Technology*, **7**, 115–158.

Durlauf, S.N. and Fafchampes, M. (2004) Social Capital. *CSAE WPS/2004-14*.

Foster, J. and Metcalfe, J.S. (2012) Economic emergence: An evolutionary economic perspective. *Journal of Economic Behavior & Organization*, **82**, 420–432.

Gehringer, A. (2011) Pecuniary knowledge externalities and innovation: intersectoral linkages and their effects beyond technological spillovers. *Economics of Innovation and New Technology*, **20**, 495–515.

Glaeser, E.L. and Scheinkman, J. (2000) Non-market interactions. *NBER Working Paper Series 8053*.

Grazzini, J. (2013) Information dissemination in an experimentally based agent-based stock market. *Journal of Economic Interaction and Coordination*, **8**, 179–209.

Griliches, Z. (1979) Issues in Assessing the Contribution of Research and Development to Productivity Growth. *The Bell Journal of Economics*, **10**, 92.

Hanusch, H. and Pyka, A. (2006) Principles of Neo-Schumpeterian Economics. *Cambridge Journal of Economics*, **31**, 275–289.

Harrison, D.A. and Klein, K.J. (2007) What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, **32**, 1199–1228.

Häussler, C. and Sauermann, H. (2014) The Anatomy of Teams: Division of Labor and Allocation of Credit in Collaborative Knowledge Production. *SSRN Electronic Journal*.

von Hayek, F.A. (1933) The Trend of Economic Thinking. *Economica*, 121.

Ioannides, Y.M. and Loury, L.D. (2004) Job information networks, neighborhood effects, and inequality. *Journal of economic literature*, **42**, 1056–1093.

Jackson, M.O. (2014) Networks in the Understanding of Economic Behaviors. *Journal of Economic Perspectives*, **28**, 3–22.

Jackson, M.O., Rogers, B.W. and Zenou, Y. (2017) The Economic Consequences of Social-Network Structure. *Journal of Economic Literature*, **55**, 49–95.

Jacobs, J. (1969) *The Economy of Cities*. Vintage Books, New York.

Jaffe, A.B., Trajtenberg, M. and Henderson, R. (1993) Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, **108**, 577–598.

Jones, B.F. (2009) The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, **76**, 283–317.

Kirman, A.P. (1992) Whom or what does the representative individual represent? *The Journal of Economic Perspectives*, **6**, 117–136.

Kirman, A. (1997) The economy as an evolving network. *Journal of evolutionary economics*, **7**, 339–353.

Klein, D. and Orsborn, A. (2009) Concatenate coordination and mutual coordination. *Journal of Economic Behavior & Organization*, **72**, 176–187.

Landini, S., Gallegati, M. and Stiglitz, J.E. (2015) Economies with heterogeneous interacting learning agents. *Journal of Economic Interaction and Coordination*, **10**, 91–118.

Lane, D.A. and Maxfield, R. (1997) Foresight, complexity and strategy. *The Economy as a Complex Evolving System 2* p. Addison-Wesley.

Lane, D.A. and Maxfield, R.R. (2005) Ontological uncertainty and innovation. *Journal of Evolutionary Economics*, **15**, 3–50.

Lane, D.A., Pumain, D., van der Leeuw, S.E. and West, G. (eds). (2009) *Complexity Perspectives in Innovation and Social Change*. Springer, Dordrecht.

Lewis, D.K. (1969) *Convention: A Philosophical Study*, Nachdr. Blackwell, Oxford.

Mokyr, J. (2002) *The Gifts of Athena: Historical Origins of the Knowledge Economy.* Princeton University Press, Princeton.

Nelson, R.R. (1959) The Simple Economics of Basic Scientific Research. *Journal of Political Economy*, **67**, 297–306.

Pakes, A. and Griliches, Z. (1998) Patents and R&D at the Firm Level: A First Look. *R & D and productivity: the econometric evidence* pp. 55–72. University of Chicago Press, Chicago.

Polanyi, M. (1958) *Personal Knowledge: Towards a Post-Critical Philosophy*, Nachdr. Univ. of Chicago Press, Chicago.

Romer, P. (1990) Endogenous technological change. *Journal of Political Economy*, S71-102.

Samuelson, L. (2004) Modeling knowledge in economic analysis. *Journal of Economic Literature*, **42**, 367–403.

Schmookler, J. (1954) The Level of Inventive Activity. *The Review of Economics and Statistics*, **36**, 183.

Schmookler, J. (1957) Inventors Past and Present. *The Review of Economics and Statistics*, **39**, 321.

Smith, A. (1759) *The Theory of Moral Sentiments*. Gutenberg Publishers, Kapaau (T.H.).

Smith, A. (1776) *The Wealth of Nations*, Bantam classic ed (ed E Cannan). Bantam Classic, New York, N.Y.

Solow, R.M. (1957) Technical Change and the Aggregate Production Function. *The Review of Economics and Statistics*, **39**, 312.

Spence, M. (2002) Signaling in retrospect and the informational structure of markets. *The American Economic Review*, **92**, 434–459.

Stigler, G.J. (1961) The Economics of Information. *Journal of Political Economy*, **69**, 213–225.

Stiglitz, J.E. (2000) The contributions of the economics of information to twentieth century economics. *The Quarterly Journal of Economics*, **115**, 1441–1478.

Stiglitz, J.E. (2002) Information and the Change in the Paradigm in Economics. *The American Economic Review*, **92**, 460–501.

Tosini, D. (2005) Capitale sociale: problemi di costruzione di una teoria. *Quaderno 32 del Diparimento di Sociologia e Ricerca Sociale, Università degli Studi di Trento*.

Trigilia, C. (2001) *Il Capitale Sociale*. Il Mulino.

Vivarelli, M. (2014) Innovation, Employment and Skills in Advanced and Developing Countries: A Survey of Economic Literature. *Journal of Economic Issues*, **48**, 123–154.

Walras, L. (1874) *Éléments d'économie Politique Pure*. Corbaz.

Watson, M. (2005) What makes a market economy? Schumpeter, Smith and Walras on the coordination problem. *New Political Economy*, **10**, 143–161.

Weitzman, M.L. (1998) Recombinant growth. *The Quarterly Journal of Economics*, **113**, 331–360.

Wutchy, S., Jones, B.F. and Uzzi, B. (2007) The Increasing Dominance of Teams in Production of Knowledge. *Science*, **316**, 1036–1039.

Teamwork and Recombinant Knowledge. An investigation of the hidden link between the nature of knowledge creation, teamwork and knowledge productivity.

Matteo Tubiana[a,b]

[a] *Department of Economics and Statistics Cognetti de Martiis, University of Torino*

[b] *BRICK, Collegio Carlo Alberto*

Abstract

The present paper posits that teamwork in inventive activity has to be considered as a necessary, emergent and effective reaction to the intrinsic recombinant nature of knowledge creation. The notion of *generative relationship* underlines the necessity of interactions among knowledge workers to cope with knowledge synchronic and diachronic complementarities. Empirical evidence from patent data at the NUTS2 territorial level shows that teamwork is also an effective behaviour for knowledge production purpose. In particular, teamwork is understood as a positive mediator between the recombinant nature of knowledge, measured through modular complexity, and territorial knowledge productivity. In so doing, the present analysis offers an alternative, richer perspective on the organization of knowledge production than that known as the "knowledge burden hypothesis".

1. Introduction

Teams are an increasing presence in academic as well as not academic science (Wutchy, Jones and Uzzi, 2007). However, it is not perfectly clear why teamwork is increasingly chosen as the most suitable organizational mode of knowledge creation. One of the very few attempts to explain this trend is the “burden of knowledge hypothesis” (Jones, 2009). This hypothesis couples the idea of an expanding knowledge space harder and harder to manage with the consequent falling knowledge capacity of individuals, eventually in need to team up in order to cope with their lack of specific competencies. It proposes a strong and causal relationship between the nature of the knowledge creative effort, its organization and its productivity – topics foremost pioneered by Schmookler (see e.g. Schmookler, 1954, 1957). After the pioneering works of Griliches (1979, 1990) and Pakes and Griliches (1998) many studies addressed the consequences of the knowledge characteristics respect to knowledge production (e.g. see Antonelli et al., 2016) or to economic productivity (e.g. see Quatraro, 2010). Similarly, others explored the relationship between the shape of knowledge workers network and that of the knowledge it produced (Graf, 2012). However, a specific knowledge-based account of the organization of knowledge production seems to be underestimated.

In the last two decades, the economics of knowledge reached a deeper understanding of the mechanisms governing knowledge creation, defined as a collective, recombinant process (Weitzman, 1998; Antonelli, 2000; Fleming and Sorenson, 2001). Indeed, the complexity and complementarity qualifying this process evidenced the importance of reconsidering the role of interactions between knowledge agents (Patrucco, 2008). The collective effort seems to emerge as a virtuous behaviour in dealing with the uncertainty and difficulty that the complex, recombinant process intrinsically and increasingly carries on (Lane and Maxfield, 1996). The following pages will provide an account of the knowledge effort – organization – productivity chain that builds strongly on this economics of knowledge’s inheritance. An account of the role of teamwork in the knowledge dynamics that is parallel but richer than the “burden of knowledge hypothesis”.

The aim of this study is to show that an increasing intensity of interactions between knowledge workers, framed as an organizational behaviour of the system, might be regarded not only as a necessary consequence of the intrinsic characteristics of knowledge creation processes but also as an effective and sustainable one. The operational rationale for this aim consists of shaping teamwork as a mediator between recombinant knowledge and knowledge production efficacy. Therefore, Causal Mediation Analysis is performed (Baron and Kenny, 1986; Zhang, Zyphur and Preacher, 2009; Imai, Keele and Tingley, 2010). Results confirm the presence of a partial, or total depending on the specifications, mediation effect. Interestingly, robustness checks suggest the specificity of the teamwork and the recombinant nature of knowledge effects on knowledge production adjusted by quality, and effect that disappears in knowledge production at large. Such results may indicate future avenues of research, where not only the quantity of knowledge production is evaluated but also, and mainly, its value or quality.

The paper is structured as follows. In Section 2, the link between the characteristics of technological knowledge and the organization of inventive activities will be made explicit. The empirical implementation of our hypothesis and the description of the relevant variables will be presented in Section 3. Finally, in Section 4 the results of regression analysis will be presented and discussed. Section 5 will conclude.

2. Theoretical background

The economics of knowledge made clear of much central knowledge is for economic growth (Arrow, 1962; Griliches, 1979, 1990; Romer, 1990). Most importantly for this research, it evidenced that knowledge is both the input and the output of a Knowledge Production Function (KPF; see Griliches, 1979; Jaffe, 1986; David, 1992). This led to the proliferation of an extensive literature investigating the existence, characteristics and use of knowledge externalities (Antonelli, 2008; Patrucco, 2008; Antonelli, 2011; Gehringer, 2011).

This paper focuses on the internal dynamics of the knowledge generation process. Knowledge itself has been defined as a heterogeneous good, whose creation is the successful outcome of a recursive process of recombination of different knowledge pieces (Weitzman, 1998; Fleming and Sorenson, 2001; Arthur, 2009). The recombinant nature of knowledge evidences the salient complementarity qualifying each knowledge bundle. A distinction arises between the diachronic complementarity (knowledge pieces piling up, determining the viable future paths, i.e. path dependence) and the synchronic one (among knowledge pieces in the same bundle). Both kinds are finally appreciated. Transactions and interactions, or a mixture of the two, become the spine of the recombinant process where no one can handle the whole disposable knowledge (Antonelli and Scellato, 2013). The KPF, which already built on the diachronic complementarity of knowledge, can be augmented with the synchronic complementarity when the relationship between internal and external knowledge is understood as multiplicative rather than additive.

Knowledge creation comes at the costs of assimilation of external knowledge and integration with the already owned one. This is true because of the modular, recombinant nature of knowledge (Fleming and Sorenson, 2001). The system capacity to keep producing technological knowledge is dependent on the rate of transactions-cum-interactions taking place within it, that is, the rate at which knowledge is exchanged and recombined (Antonelli, 2011; Antonelli and Scellato, 2013). An interesting consequence of this reasoning is that the two forms of knowledge complementarity call for a definition of knowledge as a collective, rather than (quasi) public good (Antonelli, 2000; Patrucco, 2008).

Interactions between individuals, i.e. teamwork, emerge as the most effective, if not the only way to deal with the complexity of knowledge production. Another argument in favour of this evidence has been brought by Lane and Maxfield (1997, 2005). Their idea is that a powerful instrument at disposal of individuals to cope with complex foresight horizons, i.e. activities qualified by ontological uncertainty, is the relationships with others. The treatise they offer is very articulated, but it may be synthesized and adapted to the present argument as follows: heterogeneous knowledge recombination is an uncertain process because its outcome is impossible to foresee and because the information about the knowledge space is limited (Fleming and Sorenson, 2001; Fleming, 2001; Fleming and Sorenson, 2004). Under this uncertainty, individuals do not know which strategy to tackle, i.e. how to behave. When individuals team up, however, their relationship may shield a generative potential – if some broad conditions are satisfied, such as heterogeneity and aligned and mutual directedness – such that the "relationship [..] can induce changes in the way the participants see their world and act in it, and even give rise to new entities" (Lane and Maxfield, 1996, see also Lane et al., 1996).

Therefore, we have two parallel and encompassing arguments sustaining that teamwork can be framed as a strategic reaction to deal with knowledge production defined as a recombinant, ontologically uncertain process. One emerges from the economics of knowledge and its long history of theoretical deepening of the knowledge production nature. It complements and enriches the "burden of knowledge" hypothesis (Jones, 2009), which is linear in nature and do not absorb the range of theoretical achievements exposed above. The other, instead, bridges the economics of knowledge and that of information (Stiglitz, 2000) offering a behavioural account of the existence of teams.

2.1 Measuring recombinant knowledge

Antonelli et al. (2016) and Balland and Rigby (2017) provide an excellent review of the various efforts made up to now in the exploration of knowledge composition. For a long time, the empirical investigations of the KPF exploited the raw count of patent applications as a proxy for the knowledge capacity of a region. When the theoretical advancements revealed the necessity to deepen further our understanding of the knowledge base, multiple measures stepped in, in order to measure a variety of compositional aspects. All these measures exploit the technological classes referenced in each patent document. At the very base of most of these measures lays a “co-occurrence” matrix. The idea is that two technological classes frequently co-occurring show some common characteristic. Once Weitzman (1998) had introduced the idea of recombinant knowledge, Fleming and Sorenson (2001) could implement such an idea exploiting the co-occurrence of technological classes in the same patent document, introducing the measure of “modular complexity”. This index appears as the most suitable to test the effect of recombinant knowledge on teamwork.

The construction of the modular complexity index relies on the conceptualization of invention as a search in a technological knowledge landscape. Landscapes are made of components, which in turn are measured by technological classes. The position in the landscape represents a combination of components, with an associated fitness value. Creativity takes the form of a movement on the landscape until a position with a higher fitness appears. The outcome of the search process depends on one factor: the interdependence among technological components. The concept of interdependence roughly coincides with that of modularity or coupling, that is, when two entities are interdependent, a small change in one component calls for changes in the other component for the combination to work correctly. In the words of the authors, “technological opportunities increase [..] as interdependence increases. Inventors can take advantage of increasing sensitivities and opportunistic couplings between components. Unfortunately, the difficulty and uncertainty of the recombinant search process increase along with the opportunity” (Fleming and Sorenson, 2001). The modular complexity index, therefore, grasps both the recombinant routine of knowledge creation and its ontological uncertainty.

2.2 The institutional factor

Even though the locus of the invention is the individual inventor, it is unreasonable to assume that inventors operate and interact, create and exchange knowledge, in the vacuum. They are, instead, radically embedded in a social space, other than questing on the technological knowledge landscape. On this regard, Schmookler (1954, 1957) provided one of the first accurate treatises of the professionalization of inventive activity. Schmookler's aim was actually to show that individual, autonomous inventors in the U.S.A. of the Fifties were not really disappearing and that the rising centrality of R&D laboratories was due to ascending dominance of technologists among inventors (for a very similar but more recent argument, see Melero and Palomeras, 2015). Technologists are well-educated, experienced and hired workers. From survey data, Schmookler brings evidence that team working is much higher among these professionals than among independent, non-graduated inventors. The author does never mention the characteristics of technological knowledge; rather, his full attention rests on the organization of production and on the typology of the knowledge workforce.

Schmookler's empirical results date back almost seventy years, but what is still prominent of his thesis is the stress on the institutional side of inventive activity, somewhat transcending the inventor level. The institutional factor must be considered one of the leading shaping forces of the organization of knowledge production.

2.3 Literature on teams

Earlier, no mention of any literature on the determinants of team formation has been provided because, indeed, the issue did not gain much attention. However, a growing amount of research work is dedicated to various aspects of teams, especially on their performance, their composition, their role in knowledge diffusion. In the following paragraph, some of these works are reported.

To start with, many researchers devoted their attention to the interplay between networks and knowledge, without touching the topic of teamwork as an organization (e.g. Powell and Grodal, 2005; Sorenson, 2005; Sorenson, Rivkin and Fleming, 2006; Cassi and Zirulia, 2012). As for those piece of work directly tackling teams, some aim at evidencing the impact of individual characteristics on team performance (e.g. Schettino, Sterlacchini and Venturini, 2013 on Italian patent data). On this venue, Crescenzi, Nathan and Rodríguez-Pose (2016) is a prototypical enquiry on how "physical, organisational, institutional, cognitive, social, and ethnic proximities between inventors shape their collaboration decisions". Others investigate the role of teams in the system's knowledge creation

(Bercovitz and Feldman, 2011). Cassi and Plunket (2014) bridge the two strands addressing the proximity paradox, i.e. different kinds of proximities have diverse and even antithetical effects on the creation of networks and their subsequent performance (see also Boschma and Frenken, 2010). From the concept of cognitive proximity (Nooteboom et al., 2007), to that of diversity among interacting individuals (Jacobs, 1969), seeping through the concept of absorptive capacity (Cohen and Levinthal, 1990) the jump is straightforward. Indeed, the teamwork literature concentrates a lot on the diatribe of specialization vs diversification, homogeneity vs heterogeneity, e.g. building up a further dichotomy of lone inventor vs teams. Singh and Fleming (2010), with an innovative analysis of the tails of the distribution, and Lee, Walsh and Wang (2015) go in depth into the creativity issue (Taylor and Greve, 2006 and Harrison and Klein, 2007 are two references to a much larger debate on variety and creativity). In this vein, Melero and Palomeras (2015) represent an interesting counter-part of Jones (2009)'s hypothesis. There are many other disciplines dealing with the topic of teamwork, generating many interdisciplinary cross-fertilizations. In particular, influences from psychology and management sciences are evident in works on team's internal processes of decision, division of labour, information exchange, but also on creativity (Reagans and Zuckerman, 2001; Stewart, 2006; Häussler and Sauermann, 2014).

3. Hypotheses and empirical analysis: variables and methodology

As discussed in the previous Section, the present paper aims at investigating two different but intertwined processes. The first relates the recombinant complexity nature of technological knowledge with teamwork as a strategic reaction. The second, instead, investigate the effect an increasing team-oriented organization of knowledge production has on the efficacy of production. Therefore, the first hypothesis takes the following form:

H1. The more the system knowledge is complex in a recombinatory sense, the more knowledge production is characterized by teamwork.

In the recent literature, there is growing concern about the capability of investments in knowledge creation to be efficient, i.e. not only to produce enough output, of enough quality, to sustain economic growth via innovation, but to do it without exponentially increasing costs (Jones, 2009; Bloom et al., 2017). In the arguments deployed above the topic of knowledge productivity is (not even so much) implicit, hence the second hypothesis follows:

H2. Since teamwork is the most effective way to cope with increasing knowledge complexity in a recombinatory sense, the more teamwork in knowledge production, the higher the knowledge productivity of the system.

The selected level of analysis for the empirical model is the NUTS2 region. One may wonder why the preferred level of analysis is the region rather than the individual. The reason is two-fold: i) data availability about individuals throughout time is very poor. With the data sources we are aware of, it would be difficult to control for unobserved individual heterogeneity, and to infer a directionality for causality. ii) As the theoretical section suggested, knowledge is a collective good, whose production is collective as well, since knowledge carriers need to interact systematically. Knowledge production, strictly depending on externalities, is strongly localized, meaning not only that it depends on space, but also on the territorial governance (there are very many studies to reference to, e.g. Audretsch and Feldman, 1996; Asheim and Coenen, 2005; Antonelli, Patrucco and Quatraro, 2008). It follows that the system level of analysis is not only suitable but also relevant.

Patent data stands out as the most suitable tool to test the hypotheses. Notwithstanding their limitations, patents have been shown to be a reliable proxy for innovative activities at the regional level and offer unique opportunities to study the compositional dynamics of the knowledge base (Acs, Anselin and Varga, 2002; Quatraro, 2010). Every patent application document shows a set of valuable information: 1) the names and relevant address of inventors; 2) the name and relevant address of the applicant (in EPO system, this is the owner of the future patent grant); 3) a set of dates, the pivotal one being the priority date, that is the one that most likely correspond to the period of invention; 4) the technological classes the invention pertains to; 5) forward and backward citations (patents citing the focal patent and citations that the focal patent did of previous inventions, respectively).

In order to empirically investigate hypotheses H1 and H2, a system of two equations is proposed as follow:

$$TeamShare_{i,t} = \beta_1\ Complex_{i,t} + \beta_2\ FirmShare_{i,t} + \beta_3\ UniShare_{i,t} + \beta_4\ No.\ Inv_{i,t} + \varepsilon_{i,t} \tag{1}$$

$$Prod_{i,t} = \beta_1\ Complex_{i,t} + \beta_2\ TeamShare_{i,t} + \beta_3\ R\&Dpc_{i,t} + \beta_4\ HighTech\%_{i,t} + \beta_5\ Inv\ HC_{i,t} + u_{i,t} \tag{2}$$

Where subscripts *i* stands for the NUTS2 region and *t* stands for time, βs are regression coefficients and ε and u are error components. Equation (1) describes the pervasiveness of teams in the knowledge production as a function of the complexity of the regional technological knowledge (*Complex*), the institutional composition of the applicants (*FirmShare* and *UniShare*) and the number of inventors (*No. Inv*). Equation (2), instead, models regional knowledge productivity as a function of the complexity of the regional technological knowledge (*Complex*), the pervasiveness of teams (*TeamShare*), R&D expenditure per capita (*R&Dpc*), the share of high tech employment over total employment (*HighTech%*) and inventors HC (*invHC*) measured as the mean quality of inventors' past applications. Each variable will be described in the next sections.

3.1 Data

In the following paragraphs, the data construction rationales will be described. The databases used to implement the empirical investigation are the CRIOS DB of the Bocconi University, based on disambiguated and augmented EPO files, and EUROSTAT (Coffano and Tarasconi, 2014).

Equation (1) describes the share of applications made by teams over the total number of application, as a function of knowledge complexity. Following Fleming and Sorenson (2001), a knowledge "interdependence" value is computed for each patent application, which is indeed the result of technological knowledge combination. The procedure counts two steps: first, the "ease of recombination" is computed for each technological class-year of the dataset, being technological classes specified as 4-digit IPCs. The Ease of Recombination is the ratio between the count of classes previously combined with the focal class, and the number of applications referencing to the focal class. In the second step, the interdependence index is calculated for each patent application document: the count of technological classes of the focal patent is divided by the sum of eases of recombination of these classes – an inverse weighted average. Once every patent application has its interdependence index, the average patent modular complexity is taken for each NUTS2-year.

__Figure 1__. Modular Complexity Index. Each box is a nation, dots are NUTS2 regions.

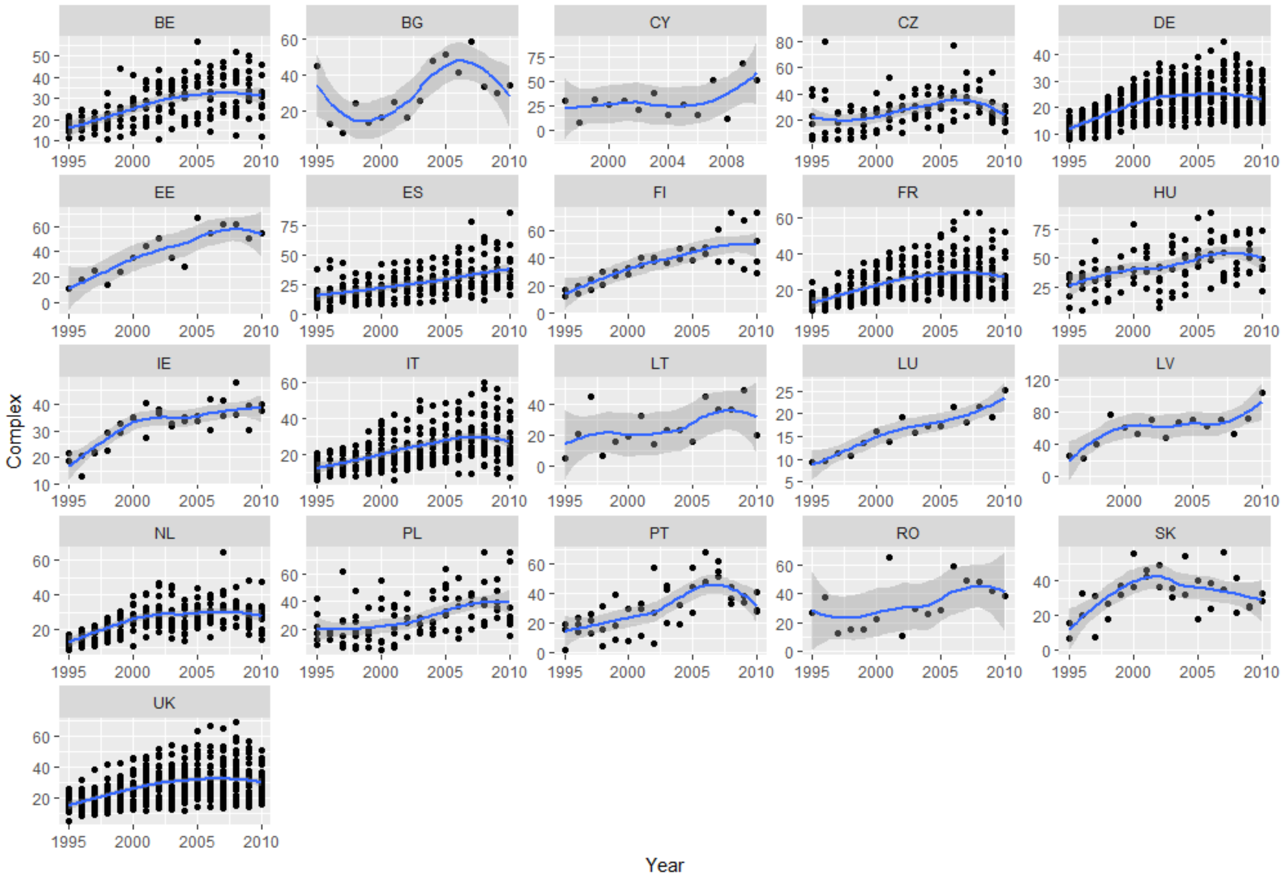


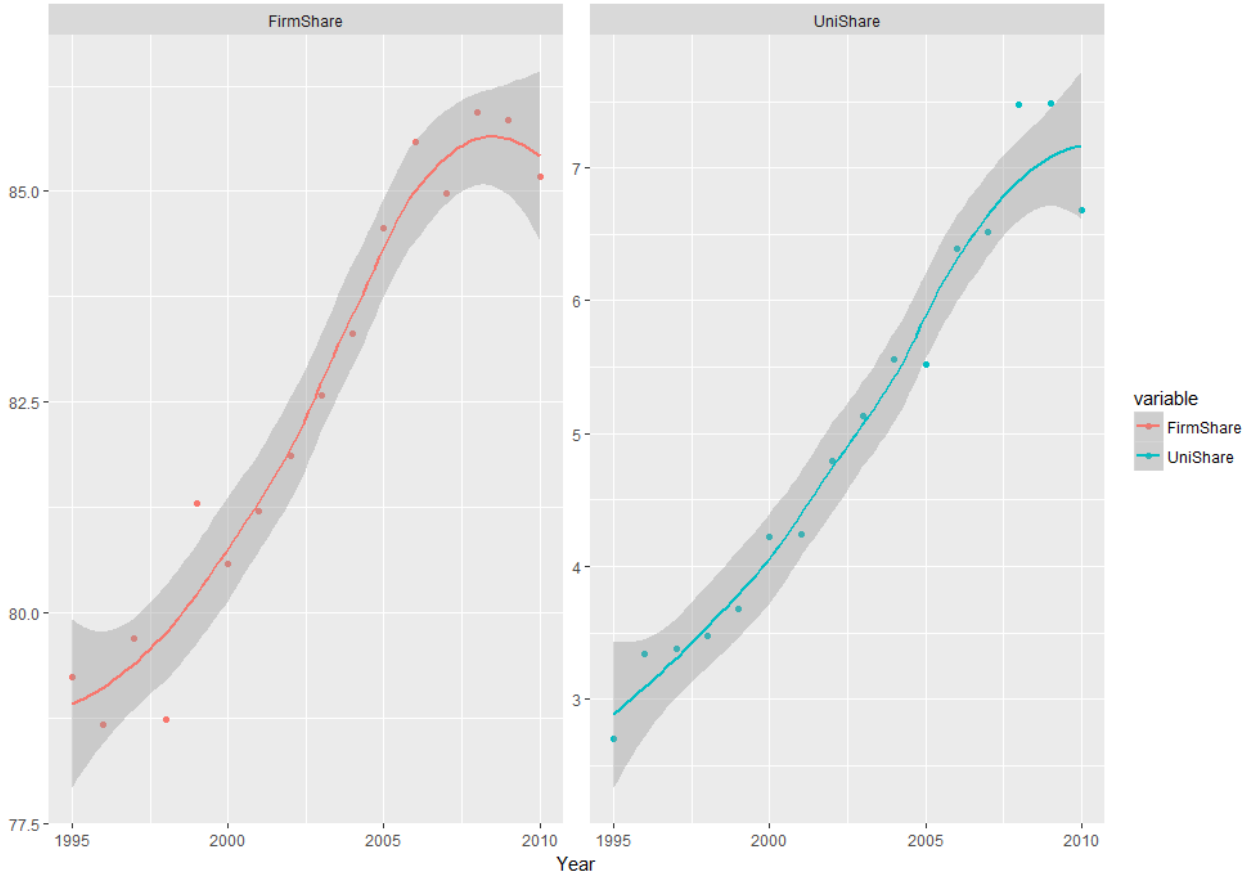
Figure 1 shows scatterplots for the trends in modular complexity in each country (Austria, Denmark and Croatia are excluded from the plot because they appear in the final dataset only for one year). Overall, the index increased across countries, but with heterogeneity both in the paces between countries and within them (NUTS2 regions). Moreover, approaching the end of the period, the trend seems to flatten.

3.1.1 The institutional factor

Two other variables appear in Equation (1): *FirmShare* and *UniShare*, i.e. how many applicants, compared to the regional total, are companies (versus individuals) and how many are universities or research institutes. Following Schmookler (1954, 1957), these two variables account for the penetration of formal institutions into inventive activity. They will control for the portion of teamwork accounted solely by the professionalization of production within firms and research labs.

Figure 2 plots the average trends of *FirmShare* and *UniShare*. It indicates that most applicants are companies and their dominance is actually increasing, but the penetration of research centres as invention commissioner is rising fast as well.

Figure 2*. Average institutional variables across time.*



3.1.2 Regional Knowledge Productivity

Even though there are studies investigating the causal effect of the composition of knowledge on economic productivity (e.g. see Quatraro, 2010; Antonelli, Krafft and Quatraro, 2010) and on knowledge production (see the above-mentioned literature based on the KPF) there is a lack of empirical investigation regarding knowledge productivity. However, it may be important to tackle the dynamics of the relationship between the knowledge production and its costs. As in every production process, the procedural organization is a crucial element of efficacy.

Given the lack of empirical research on the matter, the measure of regional knowledge productivity exploited in this paper is simply the ratio between the knowledge output and the cost of production. The denominator is approximated with the knowledge workforce – inventors – as much as in the

Cobb-Douglas production function. The numerator, instead, is subject to more discretional choices. In most basic version, it could be:

$$Prod_{i,t} = \frac{Patent\ Flow_{i,t}}{No. Inventors_{i,t}}$$

However, the count of new patent application is a very rough measure of performance. It assumes that every patent has the same economic value, which is not the case: only a small fraction of granted patents generate the greatest of economic value (Ejermo, 2009). Moreover, as Griliches (1994) reminds when investigating the R&D investments to invented patent ratio, the relationship between an invention and the patent application may not be fixed. It means that, even though we observe increasing (diminishing) growth rates of patent applications, it does not necessarily mean that inventive production is rising (falling) proportionally – accounting for patent value is a method to relax the assumption that the invention-patent relationship is fixed.

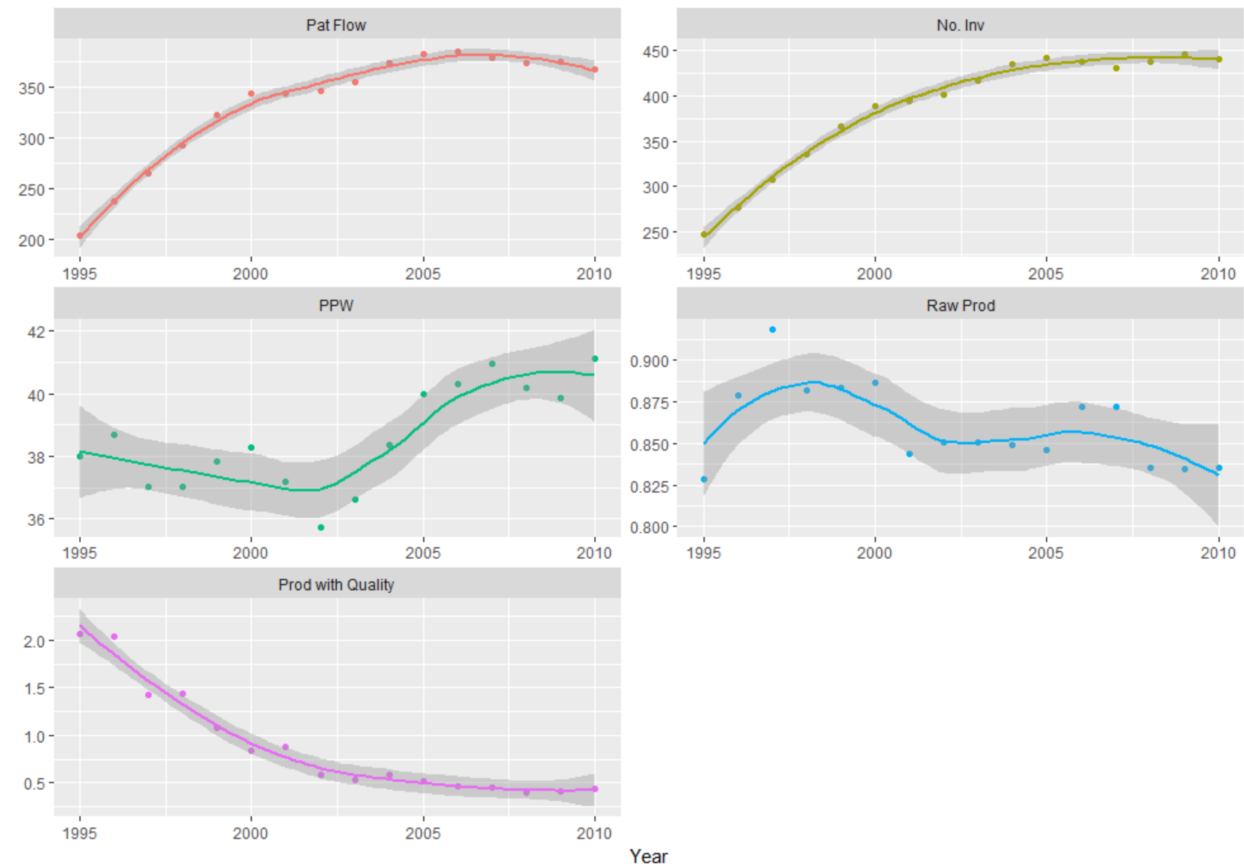
An expanding literature is debating the issue of how to measure qualitative aspects of invention from an economic perspective, and how to interpret each measurement methodology (see for e.g. Lanjouw and Schankerman, 1999; Sterzi, 2013; Squicciarini, Dernis and Criscuolo, 2013; Cassi and Plunket, 2014). The debate is far from being set. However, forward citations emerge as a good proxy of patent value (Hall, Jaffe and Trajtenberg, 2005; Sterzi, 2013).

Therefore, the dependent variable of Equation (2) will be a measure of patent production taking into account patent value through forward citations. There are different methodologies to implement this kind of measurement. The reference for this paper is Wohlrabe and Bornmann (2017), a methodological piece of work where a normalization procedure based on citation distribution's percentiles is theorized. Compared to the normalization of each patent citations count against the mean count by technological field and cohort, the percentiles-based method constitutes an advance in that it further controls for the skewness of the distribution. Intuitively, the methodology returns, for a given recipient – the region in this case – the share of patent applications that are in the upper (or lower) 50% of the citation distribution, adjusted by cohort and technology. The final measure of productivity takes the form of:

$$Prod_{i,t} = \frac{Above\ median\ Patent\ Flow_{i,t}}{No. Inventors_{i,t}}$$

Figure 3 plots the average trends for the two versions of *Prod*, computed with the raw count of patent (*Pat Flow* in the graph), or computed with the quality adjustment (*PPW* in the graph). The first thing to note is that the number of inventors and inventions follow very similar trends – a fast growth that is flattening in the end of the period. It suggests that inventors' productivity is at least not increasing. Indeed, average regional knowledge raw productivity depicts a downtrend curve over time. Figure 4 and 5 in the Appendix provide a break down by countries, which show that "smaller" knowledge producer (e.g. like Ireland, Luxemburg or Romania) are experience the steepest slowdown, whereas "bigger" countries (e.g. Germany, UK, Italy, Netherlands) experience more or less flat trends. Interestingly, the average share of patent applications belonging to the upper 50% of the distribution in terms of normalized citations has increased from 2002 onwards. However, descriptive evidence suggests that this increase in the capability to produce higher value inventions does not keep the pace of the increasing force employed.

Figure 3*. Average knowledge productivity and its components across time.*



3.1.3 Other controls

Other than the variables described in the previous paragraphs, the count of regional inventors enters Equation (1) to control for the fact that a bigger population of inventors increases the probabilities of collaborations, whereas the average quality of inventors serves as a proxy for regional HC, controlling for the knowledge population's skill composition in Equation (2).

In Equation (2), R&D expenditures per capita and the share of high tech employment over the total serve as inputs in the KPF. Unfortunately, EUROSTAT databases provide very scattered information about R&D and, to a minor extent, high tech employment. As a result, even though patent data may extend the analysis from the '80s to present day, the dataset for the empirical analysis shrinks to the 1995-2010 time window. Missing data for R&D and high tech employment are imputed with the *imputeTS* package of the R software environment, modelling each NUTS2 time-series with a structural algorithm (*na.kalman* command in R). Furthermore, NUTS2 regions belonging to the first or lower quintile of the patent production distribution are dropped in order to avoid biased results (Antonelli et al., 2016).

The final dataset results in 2808 observations of 196 NUTS2 belonging to the European Union, for 16 years from 1995 to 2010.

3.2 Econometric method: Mediation Analysis

Equation (1) and (2) are two autonomous but related regression models. The preferred method to test the hypotheses listed above is Causal Mediation Analysis, which is an inferential technique to estimate the extent to which a variable (M) mediates the effect of an explanatory (X) onto the dependent variable (Y). In particular, this methodology let the researchers appreciate both the direct effect of X on Y and the indirect effect of X, through M, on Y (Baron and Kenny, 1986). In the framework of this research, mediation analysis will quantify the extent at which the organization of knowledge production in teams acts as a digesting apparatus for the system to exploit effectively the complex knowledge embedded in the region in a recombination process.

According to Baron and Kenny (1986), every mediation analysis entails three steps. First, one has to test that the explanatory variable (*Complex* in this case) has a significant effect on the dependent (*Prod*). Second, the same must be true when regressing the explanatory (*Complex*) on the mediator (*TeamShare*). Finally, the effect of the explanatory (*Complex*) on the dependent (*Prod*) must disappear or significantly shrink in size when controlling for the mediator (*TeamShare*). In order to estimate the size and significance of the mediation effect, regressions have been run with the *mediation* package of the R software environment, which exploits the algorithms developed in Imai, Keele and Tingley (2010). The empirical models hence are:

Step1: $$Prod_{i,t} = \alpha_1 \, Complex_{i,t-1} + \boldsymbol{\alpha_2} \, Controls1_{i,t-1} + u_{i,t}$$

Step2: $$TeamShare_{i,t} = \beta_1 \, Complex_{i,t-1} + \boldsymbol{\beta_2} \, Controls2_{i,t-1} + \epsilon_{i,t}$$

Step3: $$Prod_{i,t} = \gamma_1 \, Complex_{i,t-1} + \gamma_2 \, TeamShare_{i,t-1} + \boldsymbol{\gamma_2} \, Controls1_{i,t-1} + \theta_{i,t}$$

Where *Controls1* is a matrix made of *R&Dpc*, *HighTech%* and *Inv HC*, whereas *Controls2* is a matrix containing *FirmShare*, *UniShare* and *No. Inv*. Every independent variable in each step is one-year lagged in order to reduce chances of spurious correlations.

First, each step of the mediation analysis implemented with OLS, regional dummies and time trends will be presented. However, panel data open up peculiar opportunities to explore the mediation effects (Selig and Preacher, 2009; Zhang et al., 2009). In particular, considering panel data as a particular kind of hierarchical data, where observations are nested into units or regions in our case, Random Effect estimators with varying intercepts may be applied in order to explore simultaneously Within and Between variation (Raudenbush, 2009; Bell and Jones, 2015). Following the recommendations in Zhang et al. (2009) a new set of results will be shown.

4. Results

The following tables will show the results of the mediation analysis. Specifically, Tables 1 to 3 report the main outputs retracing the hypothesis spanned above. Table 2 is already a robustness, in the sense, it follows the same mediation analysis model of Table 1, but implements the three steps with the RE Within-Between estimators. Table 1 and 2 are, in the end, identical. Tables 4 and 5 report the same kind of analysis but with different dependent (a raw patent count and the share of quality patents instead of *Prod*) or explanatory (a variable approximating the Burden of Knowledge hypothesis as Jones, 2009 suggested). Summing up, results confirm hypotheses. Modular knowledge complexity is indeed one of the causes of the specific organizational features of knowledge production, i.e. teamwork, which exerts a partial mediation effect respect to regional knowledge productivity, corrected by quality.

Tables 1, 2, 4 and 5 closely follow the instructions for mediation analysis provided by Baron and Kenny (1986) and Zhang et al. (2009). Each column corresponds to one of the three steps described in Section 3.2 (or a slight variation). Step 2 tests that the main explanatory variables is significantly related to the mediator. Steps 1 and 3 tests the ground for mediation. Two scenarios are possible in mediation analysis when introducing the mediator: either the main explanatory variable entirely loses significance and its size goes to zero, or it shrinks. In the former case, we are facing a total mediation; in the second, a partial mediation.

As for Table 1, Step 2 implements Equation (1) and confirms H1. Other than the positive and significant coefficient of *Complex*, only *FirmShare* is positive and significant, whereas *UniShare* is not significant. Step 3 confirms the ground for mediation. Indeed, from Step 1 to Step 3, the coefficient of *Complex* shrinks and loses significance, when introducing *TeamShare* in the picture, which is positive and significant. Technically, Table 1 tells a story of partial mediation, which is confirmed in Table 3, where the significance of the mediation effects is tested with a Quasi-Bayesian simulation approach (Imai et al., 2010). In Table 3, the Average Casual Mediation Effect (ACME) strictly represents the effect of *Complex* on *Prod* that is mediate by *TeamShare*. It is positive and significant and represents the 13% of the Total Effect. The Average Direct Effect (ADE) is instead, as the name suggests, the portion of the effect that directly goes from *Complex* to *Prod*. The ADE is positive and weakly significant, indicating poor precision of the sample mean estimate.

Zhang et al. (2009) point to the fact that, even in models where all the variables lay on the same measurement level, there may be confounding in the mediation estimates due to the conflation of within and between "group" variations, i.e. NUTS2 regions in this case. The meaning of a between effect variable cannot be generalized: it strongly depends on the empirical setting. In the case of this research, between effects of main explanatory variables (*Complex* and *TeamShare*) do not convey a contextual meaning, rather they circumscribe the cross-sectional relationship of variables of interest. Running the empirical models on the Within-Between version of Table 1 will provide the possibility to disambiguate the channel of mediation: if it depends on the change of the mediator and the explanatory, or, instead, if it depends on their cross-sectional differences.

A look at the within coefficients (DEV affix) in Table 2 confirms the results of Table 1. The only difference is the significance of *DEV No. Inv*, even though the point estimate is the same. *M Complex* appears positive and significant in every step, that is, it correlates positively with both *TeamShare* and *Prod*. Since the mediation analysis retrieves the very same results of Table 3, it is not reported but available upon request. This exercise was due to confirm that the channel of mediation is the variation of *TeamShare* as a response to the variation of *Complex*.

Whereas *R&Dpc* behaves as expected – an increasing input of investments on research activities exert a positive effect on productivity – the sign of *HighTech%* is unexpected. One explanation is that a mechanical bias is at work since the denominator of the dependent – the number of inventors – is quite correlated with *HighTech%* (Pearson Correlation coefficient around 0.49). If *HighTech%* gets dropped, results of both Table 2 and 3 are almost unchanged, if not that *Complex* is not any more significant in Step 3, and the significance of the ADE lowers further and the proportion of the effect accounted by the mediation channel increases at 16% (tables are available upon request). In other words, the mediation effect becomes total rather than partial, i.e. the effect of *Complex* on *Prod* goes completely through *TeamShare*.

***Table 1**. Results from LSDV regressions.*

	Dependent variable:		
	Prod Step 1	TeamShare Step 2	Prod Step 3
Complex (L1)	0.018^{***} (0.007)	0.106^{***} (0.026)	0.011^{*} (0.007)
TeamShare (L1)			0.017^{***} (0.005)
R&Dpc (L1)	0.002^{***} (0.001)		0.002^{***} (0.001)
HighTech% (L1)	-0.360^{***} (0.091)		-0.361^{***} (0.091)
Inv HC (L1)	0.049 (0.044)		0.044 (0.044)
FirmShare (L1)		0.079^{***} (0.017)	
UniShare (L1)		-0.004 (0.043)	
No. Inv (L1)		0.002 (0.001)	
Year			
NUTS2 FE	Yes	Yes	Yes
Observations	2,808	2,808	2,808
R^2	0.294	0.537	0.298
Adjusted R^2	0.240	0.502	0.243
Residual Std. Error	2.427	9.601	2.421
F Statistic	5.428^{***}	15.147^{***}	5.492^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Heteroskedastic Robust SE in parenthesis

Table 2. Results from Multilevel RE Regressions with Within-Between estimator.

	Dependent variable:		
	Prod Step 1	TeamShare Step 2	Prod Step 3
DEV Complex (L1)	0.018***	0.109***	0.011*
	(0.006)	(0.026)	(0.007)
M Complex	0.026**	0.439***	0.027**
	(0.012)	(0.074)	(0.013)
DEV TeamShare (L1)			0.018***
			(0.005)
M TeamShare			-0.003
			(0.012)
DEV R&Dpc (L1)	0.002***		0.002***
	(0.001)		(0.001)
DEV HighTech% (L1)	-0.365***		-0.366***
	(0.090)		(0.090)
DEV Inv HC (L1)	0.054		0.049
	(0.044)		(0.044)
DEV FirmShare (L1)		0.079***	
		(0.017)	
DEV UniShare (L1)		-0.005	
		(0.043)	
DEV No. Inv (L1)		0.002**	
		(0.001)	
Year	-0.164***	0.366***	-0.165***
	(0.015)	(0.055)	(0.015)
Constant	328.023***	-673.958***	330.338***
	(29.962)	(110.852)	(29.912)
NUTS2 Random Intercepts	Yes	Yes	Yes
Observations	2,808	2,808	2,808
Log Likelihood	-6,630.633	-10,558.200	-6,631.743
Akaike Inf. Crit.	13,279.270	21,134.400	13,285.490
Bayesian Inf. Crit.	13,332.730	21,187.860	13,350.830
Note:			*p<0.1; **p<0.05; ***p<0.01

***Table 3**. Causal Mediation Analysis Results*

Causal Mediation Analysis		
	Estimate	p-value
ACME	0.001893	<2e-16 ***
ADE	0.011540	0.076 .
Total Effect	0.013433	0.044 *
Prop. Mediated	0.134819	0.044 *

*p<0.1; **p<0.05; ***p<0.01
Quasi-Bayesian Confidence Intervals
Sample Size Used: 2808
Simulations: 1000

4.1 Robustness

The "burden of knowledge" hypothesis formalized the mainstream attitude towards the advancement of knowledge, grounded on the concept of cumulability. Inventors' productivity decreases and teams arise because knowledge is growing in size and inventors are becoming specialized. The linearity of the process may be broken solely by revolutions in the scientific paradigms (Jones, 2009). The aim of the present paper was to offer a more articulated vision of the knowledge dynamics and its consequences on systems' knowledge capacity. However, one may doubt that modular complexity is just a proxy for the linear process of cumulability. Therefore, in Table 4 the same set of Table 2 regressions are presented with the inclusion of a variable approximating the Burden of Knowledge (*BoK*), following the same methodology introduced by Jones (2009). The patent application knowledge burden is computed as the size of the backward citation tree of that patent. The average at the NUTS2 level is the *BoK* variable. Table 4 shows that the introduction of *BoK* does not exert any effect on *Complex* and *TeamShare* coefficients. Its within and between variation are significantly and positively related to *TeamShare*, but only its between variation correlates negatively with *Prod*. Such results confirm the burden of knowledge hypothesis but do not undermine the hypotheses here advanced.

The aim of Table 5 is, instead, to understand the specificity of the effects under consideration on knowledge production at large. In order to do so, Table 5 reports the results of Step 1 and Step 3 regressions whose dependent variable has been modified, switching from *Prod* to the raw count of patent applications per year (*Pat Flow*), in one case, and to the share of above median quality patent applications over the total (*PPW*) in the other. These are not productivity measures since they lack the "cost" of production, measured with the number of inventors. The results indicate that neither *Complex* nor DEV *TeamShare* exert an effect on *PatFlow*, but they do on *PPW* (see Step 1 bis and Step 3 bis of Table 5). Moreover, higher investments in R&D and higher shares of high-tech employment boost *PatFlow* but are not significant for *PPW*. Pooled together, these results seem to suggest that the developing of modular complexity affects the value of knowledge production rather than its quantity. In turn, quantity is more sensitive to changes in the pure economic production factors (R&D resources and the typology of the workforce) than value.

***Table 4**. Robustness with BoK.*

	Dependent variable:					
	Prod		TeamShare		Prod	
	Step 1	Step 1 bis	Step 2	Step 2 bis	Step 3	Step 3 bis
DEV Complex (L1)	0.018***	0.018***	0.109***	0.112***	0.011*	0.011*
	(0.006)	(0.006)	(0.026)	(0.026)	(0.007)	(0.007)
M Complex	0.026**	0.026**	0.439***	0.433***	0.027**	0.018
	(0.012)	(0.011)	(0.074)	(0.070)	(0.013)	(0.012)
DEV BoK (L1)		0.001		0.063**		-0.002
		(0.008)		(0.031)		(0.008)
M BoK		-0.094***		0.534***		-0.105***
		(0.019)		(0.118)		(0.020)
DEV TeamShare (L1)					0.018***	0.018***
					(0.005)	(0.005)
M TeamShare					-0.003	0.017
					(0.012)	(0.011)
DEV R&Dpc (L1)	0.002***	0.002***			0.002***	0.002***
	(0.001)	(0.001)			(0.001)	(0.001)
DEV HighTech% (L1)	-0.365***	-0.366***			-0.366***	-0.369***
	(0.090)	(0.090)			(0.090)	(0.090)
DEV inv HC (L1)	0.054	0.052			0.049	0.049
	(0.044)	(0.044)			(0.044)	(0.044)
DEV FirmShare (L1)			0.079***	0.077***		
			(0.017)	(0.017)		
DEV UniShare (L1)			-0.005	-0.011		
			(0.043)	(0.043)		
DEV No. Inv (L1)			0.002**	0.001		
			(0.001)	(0.001)		
Year	-0.164***	-0.164***	0.366***	0.335***	-0.165***	-0.163***
	(0.015)	(0.016)	(0.055)	(0.058)	(0.015)	(0.016)
Constant	328.02***	330.03***	-673.95***	-621.63***	330.33***	327.55***
	(29.962)	(31.287)	(110.852)	(116.976)	(29.912)	(31.226)
NUTS2 Random Intercepts	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,808	2,808	2,808	2,808	2,808	2,808
Log Likelihood	-6,630	-6,625	-10,558	-10,546	-6,631	-6,625
Note:						*p<0.1; **p<0.05; ***p<0.01

***Table 5**. Mediation analysis with Raw Patent Flow and Share of Quality Patents as the dependent variable*

	Dependent variable:				
	Pat Flow	PPW	TeamShare	Pat Flow	PPW
	Step 1	Step 1 bis	Step 2	Step 3	Step 3 bis
DEV Complex (L1)	0.319	0.079^{***}	0.109^{***}	0.398	0.043^{*}
	(0.274)	(0.024)	(0.026)	(0.283)	(0.025)
M Complex	-1.954	0.124^{*}	0.439^{***}	-6.014	0.030
	(4.427)	(0.066)	(0.074)	(4.850)	(0.070)
DEV TeamShare (L1)				-0.223	0.099^{***}
				(0.201)	(0.018)
M TeamShare				8.568^{**}	0.199^{***}
				(4.328)	(0.062)
DEV R&Dpc (L1)	0.263^{***}	-0.001		0.265^{***}	-0.002
	(0.024)	(0.002)		(0.024)	(0.002)
DEV HighTech% (L1)	12.147^{***}	-0.326		12.143^{***}	-0.336
	(3.808)	(0.341)		(3.808)	(0.339)
DEV inv HC (L1)	2.297	-0.135		2.360	-0.162
	(1.842)	(0.165)		(1.843)	(0.164)
DEV FirmShare (L1)			0.079^{***}		
			(0.017)		
DEV UniShare (L1)			-0.005		
			(0.043)		
DEV No. Inv (L1)			0.002^{**}		
			(0.001)		
Year	7.786^{***}	0.151^{***}	0.366^{***}	7.801^{***}	0.147^{***}
	(0.633)	(0.057)	(0.055)	(0.633)	(0.056)
Constant	$-15,219^{***}$	-267.855^{**}	-673.958^{***}	$-15,749^{***}$	-270.288^{**}
	(1,272.827)	(113.176)	(110.852)	(1,298.027)	(112.646)
NUTS2 Random Intercepts	Yes	Yes	Yes	Yes	Yes
Observations	2,808	2,808	2,808	2,808	2,808
Log Likelihood	-17,521.970	-10,418.620	-10,558.200	-17,517.710	-10,403.230
Note:					$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

5. Conclusions

The economics of knowledge has developed continuously in the last fifty years. However, it seldom discussed the organizational side of invention. The present research tried to tie together the economics of knowledge's inheritance regarding the characteristics of technological knowledge and the evidence that inventors are increasingly organized in teams. The argument nicely fits with an exploration of the determinants of regional knowledge productivity. The proposed model articulates teamwork as a consequence, or a strategic reaction, to the intrinsic characteristics of knowledge creation, defined as a collective, recombinant process. This strategic reaction has been framed as an effective strategy to cope with the difficulties of the recombinant process, hence a positive effect on knowledge productivity was expected.

Empirical results confirm the intuition indeed as befitting. In a mediation analysis framework, teamwork emerges as a partial mediator of the effect of modular knowledge complexity on territorial knowledge productivity. The characteristics of knowledge are surely not the sole determinants of the organization of knowledge production, still, they occupy a relevant role. Teamwork, as consequence of modular knowledge complexity, exerts a positive rather than negative effect on knowledge productivity. Furthermore, it seems that the higher benefits of team working emerge on the value of knowledge production rather than on the quantity, and no negative effects are detected on quantity. This result should encourage future research to go beyond a KPF whose only output is quantity, exploring, instead, various measures to evaluate the output of knowledge production.

Interactions among inventors, across territories and institutions, should not be seen as a negative side effect of the inevitable evolution of technological knowledge and science. Preferably, they are the natural consequence of an intrinsic characteristic of knowledge creation. Nonetheless, they are very effective in patent value enhancement. Therefore, the collective nature of technological knowledge should be sustained and reinforced instead of avoided.

The empirical investigation has some limitations. In the first place, a more accurate study on the dynamics of technological knowledge should cover a longer time window. Such possibility is constrained by harsh issues of data availability for what concerns other-than-patent sources. Secondly, the characteristics of knowledge that may affect teamwork and productivity may be more numerous than the sole modular complexity, which is, however, particularly relevant. Many other paths might be explored, taking advantage of the expanding list of knowledge composition indexing that are being created. Similarly, other measurements of productivity may be investigated. However, the author hopes that this study served to raise some interest in the organizational side of inventive activity.

References

Acs, Z.J., Anselin, L. and Varga, A. (2002) Patents and innovation counts as measures of regional production of new knowledge. *Research policy*, **31**, 1069–1085.

Antonelli, C. (2000) Collective Knowledge Communication and Innovation: The Evidence of Technological Districts. *Regional Studies*, **34**, 535–547.

Antonelli, C. (2008) Pecuniary knowledge externalities: The convergence of directed technological change and the emergence of innovation systems. *Industrial and Corporate Change*, **17**, 1049–1070.

Antonelli, C. (ed). (2011) *Handbook on the Economic Complexity of Technological Change*. Edward Elgar, Cheltenham, UK.

Antonelli, C., Crespi, F., Mongeau Ospina, C.A. and Scellato, G. (2016) Knowledge composition, Jacobs externalities and innovation performance in European regions. *Regional Studies*, 1–13.

Antonelli, C., Krafft, J. and Quatraro, F. (2010) Recombinant knowledge and growth: The case of ICTs. *Structural Change and Economic Dynamics*, **21**, 50–69.

Antonelli, C., Patrucco, P.P. and Quatraro, F. (2008) The governance of localized knowledge externalities. *International Review of Applied Economics*, **22**, 479–498.

Antonelli, C. and Scellato, G. (2013) Complexity and technological change: knowledge interactions and firm level total factor productivity. *Journal of Evolutionary Economics*, **23**, 77–96.

Arrow, K. (1962) Economic welfare and the allocation of resources for invention. *The rate and direction of inventive activity: Economic and social factors* pp. 609–626. Princeton University Press.

Arthur, W.B. (2009) *The Nature of Technology: What It Is and How It Evolves*, 1. Free Press pbk. ed. Free Press, New York.

Asheim, B.T. and Coenen, L. (2005) Knowledge bases and regional innovation systems: Comparing Nordic clusters. *Research Policy*, **34**, 1173–1190.

Audretsch, D.B. and Feldman, M.P. (1996) R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, **86**, 630–640.

Balland, P.-A. and Rigby, D. (2017) The Geography of Complex Knowledge. *Economic Geography*, **93**, 1–23.

Baron, R.M. and Kenny, D.A. (1986) The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, **51**, 1173–1182.

Bell, A. and Jones, K. (2015) Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods*, **3**, 133–153.

Bercovitz, J. and Feldman, M. (2011) The mechanisms of collaboration in inventive teams: Composition, social networks, and geography. *Research Policy*, **40**, 81–93.

Bloom, N., Jones, C.I., Van Reenen, J. and Webb, M. (2017) Are ideas getting harder to find? *NBER Working Paper No. 23782*.

Boschma, R. and Frenken, K. (2010) The Spatial Evolution of Innovation Networks: A Proximity Perspective. *The handbook of evolutionary economic geography* (eds R. Boschma), & R. Martin), pp. 120–135. Cheltenham: Edward Elgar.

Cassi, L. and Plunket, A. (2014) Proximity, network formation and inventive performance: in search of the proximity paradox. *The Annals of Regional Science*, **53**, 395–422.

Cassi, L. and Zirulia, L. (2012) Friends and rivals: modelling the social relations of inventors. *The Rimini Centre for Economic Analysis Working Paper Series 12-39*.

Coffano, M. and Tarasconi, G. (2014) CRIOS-Patstat database: sources, contents and access rules. *Center for Research on Innovation, Organization and Strategy, CRIOS working paper*, **1**.

Cohen, W.M. and Levinthal, D.A. (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, **35**, 128.

Crescenzi, R., Nathan, M. and Rodríguez-Pose, A. (2016) Do inventors talk to strangers? On proximity and collaborative knowledge creation. *Research Policy*, **45**, 177–194.

David, P.A. (1992) Knowledge, Property, and the System Dynamics of Technological Change. *The World Bank Economic Review*, **6**, 215–248.

Ejermo, O. (2009) Regional Innovation Measured by Patent Data—Does Quality Matter? *Industry and Innovation*, **16**, 141–165.

Fleming, L. (2001) Recombinant Uncertainty in Technological Search. *Management Science*, **47**, 117–132.

Fleming, L. and Sorenson, O. (2001) Technology as a complex adaptive system: evidence from patent data. *Research policy*, **30**, 1019–1039.

Fleming, L. and Sorenson, O. (2004) Science as a map in technological search. *Strategic Management Journal*, **25**, 909–928.

Gehringer, A. (2011) Pecuniary knowledge externalities and innovation: intersectoral linkages and their effects beyond technological spillovers. *Economics of Innovation and New Technology*, **20**, 495–515.

Graf, H. (2012) Inventor networks in emerging key technologies: information technology vs. semiconductors. *Journal of Evolutionary Economics*, **22**, 459–480.

Griliches, Z. (1979) Issues in Assessing the Contribution of Research and Development to Productivity Growth. *The Bell Journal of Economics*, **10**, 92.

Griliches, Z. (1990) Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, **28**, 1661–1707.

Griliches, Z. (1994) Productivity, R&D and the Data Constraint. *American Economic Review*, **84**, 1–23.

Hall, B.H., Jaffe, A.B. and Trajtenberg, M. (2005) Market value and patent citations. *The RAND Journal of Economics*, **36**, 16–38.

Harrison, D.A. and Klein, K.J. (2007) WHAT'S THE DIFFERENCE? DIVERSITY CONSTRUCTS AS SEPARATION, VARIETY, OR DISPARITY IN ORGANIZATIONS. *Academy of Management Review*, **32**, 1199–1228.

Häussler, C. and Sauermann, H. (2014) The Anatomy of Teams: Division of Labor and Allocation of Credit in Collaborative Knowledge Production. *SSRN Electronic Journal*.

Imai, K., Keele, L. and Tingley, D. (2010) A general approach to causal mediation analysis. *Psychological Methods*, **15**, 309–334.

Jacobs, J. (1969) *The Economy of Cities*. Vintage Books, New York.

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

Jones, B.F. (2009) The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? *The Review of Economic Studies*, **76**, 283–317.

Lane, D., Malerba, F., Maxfield, R. and Orsenigo, L. (1996) Choice and action. *Journal of Evolutionary Economics*, **6**, 43–76.

Lane, D. and Maxfield, R. (1996) Strategy under complexity: Fostering generative relationships. *Long Range Planning*, **29**, 215–231.

Lane, D.A. and Maxfield, R. (1997) Foresight, complexity and strategy. *The Economy as a Complex Evolving System 2* p. Addison-Wesley.

Lane, D.A. and Maxfield, R.R. (2005) Ontological uncertainty and innovation. *Journal of Evolutionary Economics*, **15**, 3–50.

Lanjouw, J.O. and Schankerman, M. (1999) The quality of ideas: measuring innovation with multiple indicators. *NBER Working Paper No. 7345*.

Lee, Y.-N., Walsh, J.P. and Wang, J. (2015) Creativity in scientific teams: Unpacking novelty and impact. *Research Policy*, **44**, 684–697.

Melero, E. and Palomeras, N. (2015) The Renaissance Man is not dead! The role of generalists in teams of inventors. *Research Policy*, **44**, 154–167.

Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V. and van den Oord, A. (2007) Optimal cognitive distance and absorptive capacity. *Research Policy*, **36**, 1016–1034.

Pakes, A. and Griliches, Z. (1998) Patents and R&D at the Firm Level: A First Look. *R & D and productivity: the econometric evidence* pp. 55–72. University of Chicago Press, Chicago.

Patrucco, P.P. (2008) The economics of collective knowledge and technological communication. *The Journal of Technology Transfer*, **33**, 579–599.

Powell, W.W. and Grodal, S. (2005) Networks of Innovators. *The Oxford Handbook of Innovation* p. Oxford University Press.

Quatraro, F. (2010) Knowledge coherence, variety and economic growth: Manufacturing evidence from Italian regions. *Research Policy*, **39**, 1289–1302.

Raudenbush, S.W. (2009) Adaptive centering with random effects: An alternative to the fixed effects model for studying time-varying treatments in school settings. *Education*, **4**, 468–491.

Reagans, R. and Zuckerman, E.W. (2001) Networks, Diversity, and Productivity: The Social Capital of Corporate R&D Teams. *Organization Science*, **12**, 502–517.

Romer, P. (1990) Endogenous technological change. *Journal of Political Economy*, S71-102.

Schettino, F., Sterlacchini, A. and Venturini, F. (2013) Inventive productivity and patent quality: Evidence from Italian inventors. *Journal of Policy Modeling*, **35**, 1043–1056.

Schmookler, J. (1954) The Level of Inventive Activity. *The Review of Economics and Statistics*, **36**, 183.

Schmookler, J. (1957) Inventors Past and Present. *The Review of Economics and Statistics*, **39**, 321.

Selig, J.P. and Preacher, K.J. (2009) Mediation Models for Longitudinal Data in Developmental Research. *Research in Human Development*, **6**, 144–164.

Singh, J. and Fleming, L. (2010) Lone Inventors as Sources of Breakthroughs: Myth or Reality? *Management Science*, **56**, 41–56.

Sorenson, O. (2005) Social networks, informational complexity and industrial geography. *The role of labour mobility and informal networks for knowledge transfer* pp. 79–96. Springer.

Sorenson, O., Rivkin, J.W. and Fleming, L. (2006) Complexity, networks and knowledge flow. *Research Policy*, **35**, 994–1017.

Squicciarini, M., Dernis, H. and Criscuolo, C. (2013) *Measuring Patent Quality*. OECD Science, Technology and Industry Working Papers.

Sterzi, V. (2013) Patent quality and ownership: An analysis of UK faculty patenting. *Research Policy*, **42**, 564–576.

Stewart, G.L. (2006) A Meta-Analytic Review of Relationships Between Team Design Features and Team Performance. *Journal of Management*, **32**, 29–55.

Stiglitz, J.E. (2000) The contributions of the economics of information to twentieth century economics. *The Quarterly Journal of Economics*, **115**, 1441–1478.

Taylor, A. and Greve, H.R. (2006) Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal*, **49**, 723–740.

Weitzman, M.L. (1998) Recombinant growth. *The Quarterly Journal of Economics*, **113**, 331–360.

Wohlrabe, K. and Bornmann, L. (2017) Normalization of citation impact in economics. *MPRA Paper No. 80384*.

Wutchy, S., Jones, B.F. and Uzzi, B. (2007) The Increasing Dominance of Teams in Production of Knowledge. *Science*, **316**, 1036–1039.

Zhang, Z., Zyphur, M.J. and Preacher, K.J. (2009) Testing Multilevel Mediation Using Hierarchical Linear Models: Problems and Solutions. *Organizational Research Methods*, **12**, 695–719.

Appendix

***Figure 4**. Prod with Quality - break down by nat.*

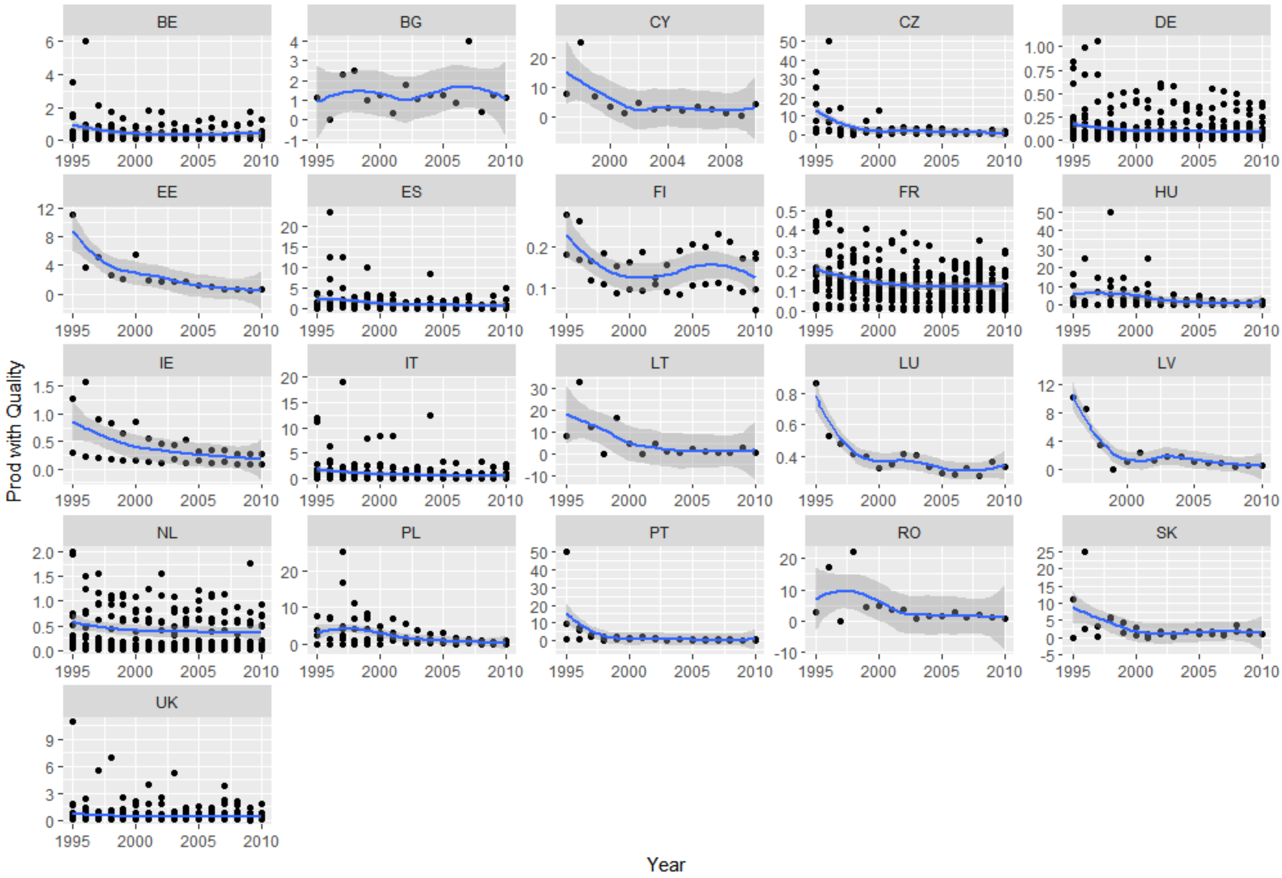
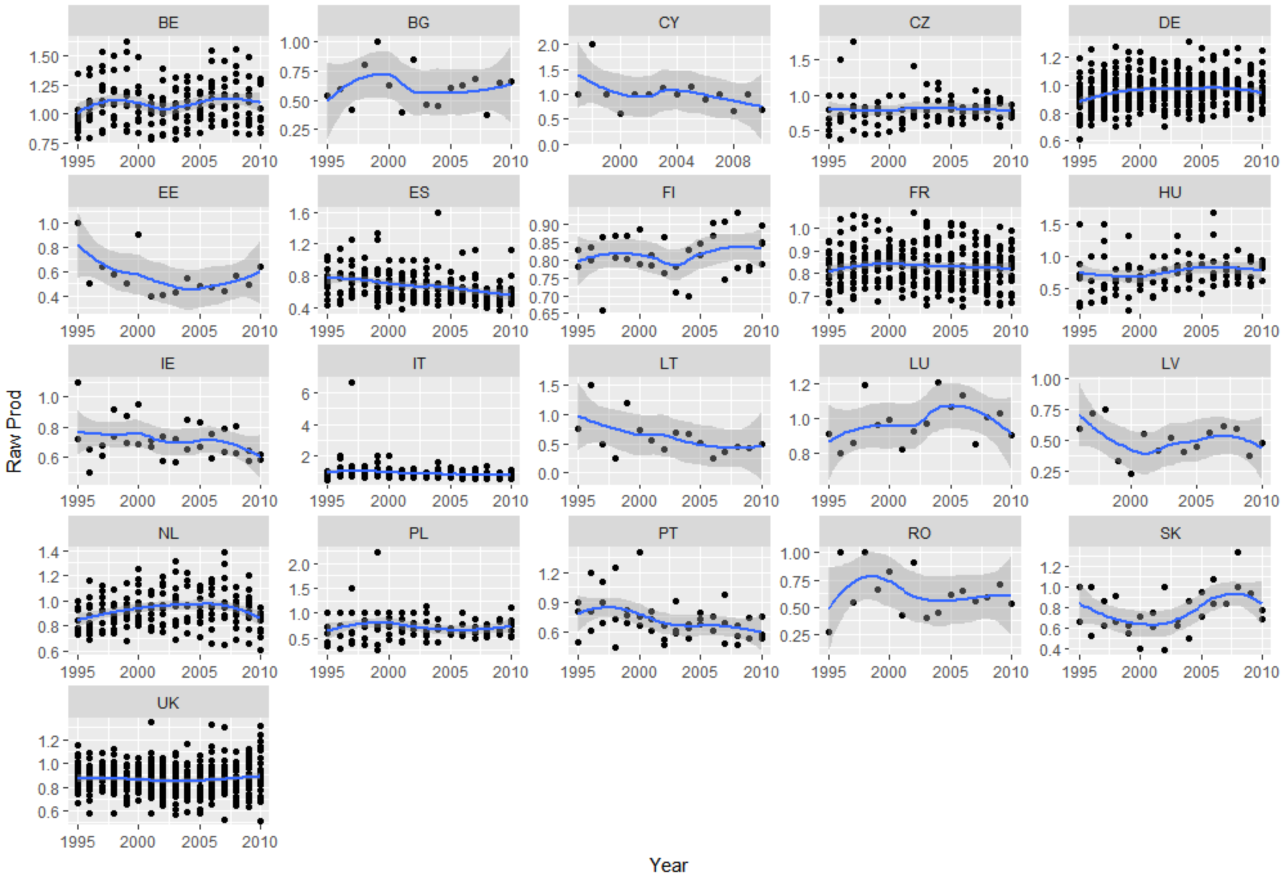


***Figure 5**. Raw Prod - break down by nat.*



The Knowledge dynamics of inventors' performance: A multilevel analysis of knowledge externalities

Matteo Tubiana[a,b]

Ernest Miguélez[c,d] (CNRS, GREThA at University of Bordeaux and AQR-IREA, University of Barcelona),

Rosina Moreno[d] (AQR-IREA, University of Barcelona)

[a] *Department of Economics and Statistics Cognetti de Martiis, University of Torino*

[b] *BRICK, Collegio Carlo Alberto*

[c] CNRS, GREThA at University of Bordeaux

[d] *AQR-IREA, University of Barcelona*

Abstract

The present paper exploits a large EPO patent dataset to explore the impact of knowledge externalities on individual inventors' performance. The first objective of the research consists in highlighting the importance to take into account the multilevel socio-economic structure each inventor is embedded, in order to understand where they source knowledge and how valuable these sources are. Therefore, we make use of a multilevel approach and simultaneously investigate the effect of the collaboration network, the firm and the metropolitan region's knowledge capacity on inventors' performance. In doing so, we bridge two traditions, namely the economics of knowledge externalities and the literature exploiting the theoretical background of labour economics to explain knowledge flows, currently developing on parallel strands. Results suggest that inventors take advantage of the knowledge pools at each layer they are embedded in. Territorial knowledge externalities are stronger than network externalities. Interestingly, there is a substantial complementarity between the network and firm knowledge bundles, whereas firm and territorial knowledge appear as a substitute.

1. Introduction

The role knowledge externalities play for the productivity, competitiveness and growth of firms, regions and countries is still a debated topic, despite the vast amount of studies that have flourished in recent years (for a review of these studies, see Antonelli et al., 2016). Even though the relevance of knowledge externalities exploitation both in technological and knowledge production is settled since the pivotal contributions of Griliches (1979), Jaffe (1986) and Crepon, Duguet and Mairesse (1998), the mechanisms governing these dynamics are still to be understood. In this rich and heterogeneous quest, we contribute in a twofold way. First, we extend the research strand about knowledge generation at the individual level, which is increasingly considered as the fundamental level of analysis for exploring knowledge creation mechanisms (Fleming, 2001), but still under-investigated. In particular, as a novelty in the literature, we explore the effect of the pool of regional knowledge on individuals' performance directly. Specifically, we concentrate on inventors, a class of highly skilled, highly educated knowledge workers who are behind the production of technological innovations spurring economic growth and well-being. Second, we add on the few attempts in the economics of innovation and knowledge literature to account for the multilevel structure of the empirical settings (Rothaermel and Hess, 2007; Raspe and van Oort, 2011; Grillitsch, Martin and Srholec, 2016). Indeed, in most social settings, individuals are embedded in a variety of different layers, affecting their choices and capabilities (Raudenbush and Bryk, 2002). Accounting for this hierarchical structure may better assign to each level the respective importance in directing the phenomenon under investigation.

When implementing these objectives, we will try to match two parallel traditions dealing with knowledge externalities. The first is the economics of knowledge and its localization approach (Griliches, 1979; Asheim and Coenen, 2005). The basics of this tradition are straightforward: knowledge is a peculiar good, in between private and public features; for this reason, it is not fully appropriable and it spills out of the owner/producer. This first insight developed into a vibrant literature addressing the costs and benefits of knowledge thereon referred to as externalities leakages (see Antonelli, 2016 for a treatise of the possible trade-offs). In deepening the understanding of what affects knowledge to flow among individuals, this literature soon advocated for the territorial boundaries of such a process, giving birth to the concept of Regional Innovation Systems (Asheim and Coenen, 2005, 2006). It is within this strand of literature that the concept of *stock of knowledge* or *knowledge base* developed, as measures of the knowledge potential of economic agents. Individuals, e.g. entrepreneurs, can partly exploit the knowledge of other individuals merely because they are co-located in space, in a bond of institutions, supply chains and repetitive transactions and interactions (topic deepened by the Knowledge Spillover Theory of Entrepreneurship, see (Audretsch et al., 2012; Audretsch, Lehemann and Hinger, 2015). Such a space constraint gives the chance to build territorial indicators of knowledge that reflect individual endowments by definition, but, in theory, are not a simple sum of individual addends. Knowledge creation is, indeed, a complex phenomenon taking the shape of a *collective process* (Antonelli, 2000).

The second tradition is the one focusing on collaboration networks and mobility (Breschi and Lissoni, 2001, 2009). Contributions referring to this tradition concentrate on the micro-level analysis of individuals' interactions, trying to account for the role of different kinds of links as vehicles of knowledge flows. What matters most in this literature is the stress on the transmission *channel*. Indeed, it brought evidence of the significant role of inventors/scientists mobility across territories and institutions in explaining the variation in knowledge flows patterns measured as patent/paper citations. Furthermore, it suggested that networks for business and networks for knowledge flows do not necessarily coincide, meaning that different resources flow on different channels. In this literature, the very concept of spillover is criticized along with a debunking effort, which seems to aim at laying the groundwork for a microeconomics of knowledge externalities.

In this paper, we set our interest at the micro-level taking into consideration inventor performance, while keeping in mind the achievement of a systemic, multilevel approach to knowledge creation. Indeed, our understanding of the factors driving and influencing inventors' productivity is still modest (Giuri et al., 2007). However, the topic deserves paramount attention because it is within individuals that knowledge is created and among individuals that it is exchanged (Grant, 1996, p. 199; Fleming, 2001; Fleming and Szigety, 2006). In the last decade, attention towards individual inventors and their activities grew, bringing to the light a series of studies investigating the individual-level determinants of inventive activity (notable examples are Hoisl, 2007; Hussinger, 2012; Zwick et al., 2017). However, apart from a few studies by Fleming and Sorenson (Fleming and Sorenson, 2001, 2004) there are not attempts to account for individual inventor performance with the toolkit of the economics of knowledge literature. Therefore, we aim at analysing the direct effect of the pool of regional knowledge at various socio-economic layers, on inventors' performance.

In the next Section, we will review the literature presented above. It will provide the theoretical background to outline, in Section 3, our approach and contribution. Section 4 will present the empirical model, whereas Section 5 will describe the data building process and the final dataset. Finally, Section 6 will outline the preliminary results and conclusions will follow.

2. Literature review

The first economist to cast attention on the role of individuals as engine of knowledge dynamics was Joseph Schumpeter, clarifying that invention and innovation are two different moments. The first moment is that of individual creativity, whereas the second one concerns selection, diffusion and creation of wealth, and is much more 'systemic' (Schumpeter, 1939; Fleming and Szigety, 2006). Similarly, the innovative dynamics of a system can be divided into a technological production routine, and a knowledge production one, where the latter enters as a key factor in the former. The series of prominent contributions by Zvi Griliches established the methodologies to investigate these routines separately, namely the Technology Production Function (TPF) (Griliches, 1979) and Knowledge Production Function (KPF) (Jaffe, 1986). Establishing the TPF and the KPF enabled the appreciation of the role of knowledge as the hidden factor boosting firms' productivity thanks to the virtuous presence of externalities or spillovers, i.e. knowledge pieces that can be used by others than the creator at lower than equilibrium-cost (Griliches, 1995, 1998; Pakes and Griliches, 1998). At the very earth of this approach, there are (at least) two pillars: the "special" characteristics of knowledge as an economic good (Arrow, 1962) and its recombinant nature (Weitzman, 1998). Knowledge as a non-rival and only partially appropriable good motivates a theory of the existence of knowledge externalities. The fact that knowledge creation happens through recombination of existing knowledge bundles leads to appreciating the interactive and collective nature of knowledge, whose generation is therefore bounded within the social and geographical limit of interactions between individuals. Soon, knowledge evolution has been appreciated as a cumulative, path-dependent, and interactive process (Dosi, 1982; Nelson and Winter, 2004). Therefore, the amount and quality of the knowledge produced in a system determine the extension and composition of the knowledge externalities embedded within its boundaries (Boschma, Balland and Kogler, 2015). In turn, internal characteristics of firms embedded into the systems determine their ability to absorb, metabolize and put in production external knowledge – knowledge externalities have a cost (Cohen and Levinthal, 1990; Antonelli, 2008).

By virtue of knowledge collective and interactive nature and of its sticky and tacit components (Cowan, David and Foray, 2000) knowledge dissemination strongly decays with space. Consequently, the knowledge localization explains the propensity for innovative activities to cluster geographically (Jaffe, Trajtenberg and Henderson, 1993; Audretsch and Feldman, 1996; Audretsch and Stephan, 1999), and the spatial heterogeneity in quantity and quality of technological production (Hidalgo and Hausmann, 2009; Balland and Rigby, 2017). The consequence of the increasing attention upon the bounding and rooted aspects of knowledge dynamics induced a dedicated series of empirical studies on so-called Regional Innovation Systems (Asheim and Coenen, 2005) and the interplays between various forms of proximity (Boschma, 2005; Balland, Boschma and Frenken, 2015). The coordinated governance of the territorial knowledge potential emerged as a specific issue (Antonelli, Patrucco and Quatraro, 2008).

On a parallel, more recent strand, some authors cast scepticism on the theory of knowledge externalities. The problematic aspect of such a theory carried out at an aggregated level is that it treats generation and appropriation of externalities/spillovers as a 'black box', whereas, instead, a multiplicity of forces are at stake (Agrawal, Cockburn and McHale, 2006; Rodríguez-Pose and Crescenzi, 2008). Miguélez and Moreno (2013) perfectly synthesises these positions: "As Zucker, Darby and Armstrong (1998) or Breschi and Lissoni (2009) put it, in the absence of large levels of local labour mobility of super-skilled labour and research networks of formal collaboration, informal linkages and serendipitous encounters explain only a relatively minor part of the localization of knowledge flows. Thus, knowledge flows might be a powerful agglomeration force and might basically occur at the regional level, but not in the form of spillovers, rather, through well-regulated knowledge exchanges deliberated on a market basis (Breschi and Lissoni, 2001)." Breschi and Lissoni's series of papers focused on the foundational empirical demonstration of the existence of localized knowledge spillovers – Jaffe et al., 1993 paper – where, for the first time, the paper trails of knowledge have been identified exploiting patent citation data. The fundamental challenge to Jaffe's work was that looking at inventors as knowledge carriers, and tracking their mobility, much variance in citations patterns was explained.

The limit of this approach seems to consist of the inadequate attention paid to exploring the context into which mobility takes place. It is clear in fact that mobility within knowledge-rich contexts is likely to yield far more results than mobility in knowledge-poor context. For the same token, this literature does not fully explore the direction of mobility: mobility from knowledge-poor context to knowledge-rich ones is likely to yield better results than mobility from knowledge rich context to knowledge poor ones. A recent contribution by Fernandez-Zubieta and colleagues exploring a related issue, i.e. the effects of mobility on academic carriers, has implemented this distinction (Fernández-Zubieta, Geuna and Lawson, 2016). We aim at elaborating a "contextual" approach that tries to combine the analysis of the effects of the context along the lines of the Griliches-Jaffe tradition together with the attention towards the specific channels of knowledge diffusion suggested by Breschi and Lissoni.

Onto this track, a new literature focusing on inventor networks flourished (Singh, 2005). In particular, massive attention has been dedicated to inventor's mobility, primarily labour-related (Almeida and Kogut, 1999; Breschi and Lissoni, 2009) but also geography and technology related (Latham et al., 2011). Indeed, the primary aim of these research efforts, primarily exploiting the precious availability of patent data, has been to highlight the channels through which knowledge disseminates. An obvious consequence of the interest in individual-level dynamics has been a renewed attention on the inventor as the *locus* of knowledge creation, a bit forgotten by the main strands of the economics of knowledge. Two are the directions this literature has taken lately: one exploits the experimental and theoretical toolkit provided by the research on labour/geographical/technological mobility and networks (Hoisl, 2007; Palomeras and Melero, 2010; Nakajima, Tamura and Hanaki, 2010; Latham et al., 2011; Hussinger, 2012; Miguélez and Moreno, 2013). The other, instead, promotes individual surveys investigating psychological, educational and subjective characteristics of individual inventors (Mariani and Romanelli, 2007; Schettino, Sterlacchini and Venturini, 2013; Bell et al., 2017; Zwick et al., 2017). In this new wave of studies centred on inventors, the reference to the original themes of the economics of knowledge has been neglected. The literature on team composition and performance stands out as a partial exception. This strand of research investigates the complex interactions taking place between different typologies of knowledge backgrounds – i.e. diversity, generality and specialization – when they come together into a team (Taylor and Greve, 2006; Wuchty, Jones and Uzzi, 2007; Singh and Fleming, 2010; Graf, 2012; Melero and Palomeras, 2015).

3. A Multilevel Spillover Approach

The effort of opening the 'black box' of knowledge externalities has been a significant step forward in our understanding of *how* knowledge disseminates, stating distinctly the role of networks and institutions into the coordination of the knowledge dynamics (Cowan and Jonard, 2003, 2004). The theoretical and empirical system-level analysis within the framework of the economics of knowledge has also provided the necessary tools to understand how knowledge is created, evolves and transmutes into technological change (Antonelli and Colombelli, 2017). However, these traditions are developing along two parallel courses. Hence, the primary focus of our research is that of letting them touch on the ground of the studies about invention and inventors.

We see the inventor as the ground zero of creativity. Exploiting the largely available and corroborated data-source of patent documents, we set our unit of analysis at the inventor level. No creative action takes place into the solitude. On the contrary, every individual is embedded into a multiplicity of social layers. We focus on three of them, which, in our view, synthesize at best the complicated context most inventors operate in. The first layer is the network of job relationships each inventor builds around himself during his activity. As widely reported above, the inventor-network has profound influences on the inventor's choices. The second layer is the institutional dominion par excellence: the firm where the inventor is employed. The workplace is not only the main environment for the inventor to interact with other carriers of knowledge, but it is also an important driver of research trajectories. The third and last layer is the geographical space where the inventor operates. This layer is encompassing the other two, but also comprehending other relevant collective events and environments, we cannot directly account for.

At each of these three layers, a multitude of forces may take place, affecting inventor's choices and performance. We are interested in the role of some typical metrics of the economics of knowledge, oriented to quantify and qualify the magnitude and kind of knowledge embedded in a repository. In this paper, we will focus on the stock of knowledge at each layer. In our perspective, the stock of knowledge is not a measure of tangible assets at disposal in the knowledge production. Instead, it is an index of a knowledge potential embedded in the repositories (the network, the firm or the location).

We plan to test to what extent the stock of knowledge at the geographical level – the level at which knowledge externalities have been quested for – sustains the inventor's creative effort when the stocks of the other two layers (the firm and the network) are controlled for. From the literature focusing on networks and mobility's perspective, there should be little evidence of any significant impact of the territorial level once the channels of knowledge dissemination are controlled for. Instead, institutional and formal channels (labour-related mechanisms) could not be enough to account for all the creative knowledge potential of a territory. We will investigate the direct impact of both the network and the institutional and territorial environments, with specific attention to the interactions between them. The rationale is that a fertile pool of knowledge may be metabolized and digested differently by inventors equipped with different network and institutional knowledge potential. This kind of reasoning aligns with the absorptive capacity literature (Cohen and Levinthal, 1990).

4. Econometric strategy

Both the strand of literature investigating the consequences of networking and mobility on individual inventors and the stream focusing instead on individual characteristics, are ego-centred, i.e. they deal with one level of analysis only: the individual. The peculiarity of our approach, instead, is that we want to look at the different layers simultaneously building up the inventor's creative environment. Such a pursuit entails methodological carefulness. There is a long tradition in the economics of education addressing the issue of hierarchical settings, i.e. settings where individuals are nested into groups at many layers (Raudenbush, 2009). For example, pupils belonging to the same school or/and to the same neighbourhood; or, inventors belonging to the same firm or/and the same MR. When the hierarchical structure of the data is ignored and only one layer is analysed, two implicit underlying assumptions are made: 1) that the salient heterogeneity takes place only within that layer and that other layers are more or less homogeneous, and 2) that the layer analysed is independent of the others (Rothaermel and Hess, 2007). In some settings, such assumptions may be undesirable or inappropriate. There are a number of possible approaches to the issue, which, in more familiar econometrics terms is referred as 'clustered data', but two are the most famous: clustering the standard errors in a FE regression settings, or Multilevel Analysis (Raudenbush and Bryk, 2002; Fazio and Piacentino, 2010; Cameron and Miller, 2015). Even though Multilevel Analysis has seen only a few applications in regional economics (Fazio and Piacentino, 2010; but only Raspe and van Oort, 2011; López-Bazo and Motellón, 2017 in the subfield of economics of innovation and knowledge, at the very best of our knowledge) recent pieces of works suggest its superiority over FE regressions in a number of settings (Raudenbush, 2009; Bell and Jones, 2015; Bell, Fairbrother and Jones, 2016).

Traditionally, the big advantage assigned to the FE estimator is that it eliminates by definition group-invariant variables and their interactions with lower-level variables (Clarke et al., 2010). In so doing, any possible correlation between covariates and the errors due to un-observed group-invariant characteristics is avoided. In a longitudinal setting, the serial auto-correlation of lowest level variables can be controlled for with clustered or adequately modelled autoregressive SE. However, researchers may be interested in direct effects of higher-level variables and cross-level effects. MA accommodates this need. Indeed, the basics of the MA approach is that the individual unobserved heterogeneity is accounted for with an individual specific-component that is not fixed, instead comes from a random distribution. Therefore, group(time)-invariant variables can enter into the model. It is important to mention that MA handles both perfectly nested and cross-classified hierarchical structure. However, one critical point of the MA approach is that the unobserved heterogeneity is not eliminated, meaning that, if the model is not perfectly specified, the omitted variables bias threatens causal interpretation. Both Raudenbush (2009) and Bell and Jones (2015) suggest a robust version of MA, where variables are demeaned as in the Mundlack formulation of the FE estimator. When more than one group fixed effect is needed, sequential demeaning is allowed in balanced panels, whereas it is not feasible in unbalanced settings. It is, therefore, problematic to control for multiple group fixed effect at different levels in MA, even though it is theoretically possible. We are still working on this issue.

4.1. The model

In the following section, we will present the preliminary results for a four-way FE regression estimated with OLS.

$$AppXinv_{i,f,m,t} = MRstock_{m,t-1} + FIRMstock_{f,t-1} + NETstock_{i,t-1} + Controls_{i,f,m,t-1} + \delta_i + \delta_f + \delta_m + \delta_t + \varepsilon_{i,f,m,t}$$

Where *i* is the inventor, *f* the firm or establishment, *m* the MR, *t* stands for time and the δs are a set of FE. The three main explanatory variables account for the knowledge potential of the multilevel structure the inventor is embedded in, respectively the MR, the firm and the network of past collaborators. In this specification of the model, the multilevel structure of the data will be accounted with a full set of interactions of main variables of interest across levels and cluster-robust SE (Cameron and Miller, 2015).

In the next section, preliminary results will be spanned out. In our regression models, we decided to exploit the establishment level rather than the firm one (see section 5.2). Indeed, even though assuming the existence of establishments (such as laboratories or branches) from inventor's addresses may lead to some misinterpretation (inventors may commute or work in hosted institutions), we believe the benefits stemming from production localization are exceeding potential threats and, as a matter of fact, knowledge production is localized. With the *ESTABstock* we grasp the territorial dimension in the firm's externality reach. Nonetheless, we control for firm-specific, rather than establishment-specific, FE.

5. Data

As anticipated above, we extensively exploit patent data to build our primary variables of interest. Even though patent documents represent only a product-oriented subset of possible knowledge production, they represent a unique opportunity of observing the moment of creativity at wide, across individuals, territories and time. Methodologically, patent documents make it possible to build longitudinal datasets, whose potential in terms of inference is substantial. Moreover, this is the only source we can exploit to observe the multiple layers we are interested in: the collaboration network and its evolution throughout time and the inventor's engagement with one or more firms and the territory he belongs to.

We match two different patent databases in order to retrieve all the necessary information about our three layers of interest: the CRIOS Patent Database (2014) (Coffano and Tarasconi, 2014) and the HAN Database (2016) from REGPAT. Both databases have the crucial feature of being the output of a process of name-disambiguation: inventors' names in CRIOS, applicants' names in HAN. Out of this matched dataset, we build our main variables of interest: the number of patent applications per inventor-year (*AppXinv*, dependent variable), and the stock of patent application per inventor's network (*NETstock*, five year window), applicant-year (*FIRMstock* or *ESTABstock,* see section 5.2) and territory-year (*MRstock*, explanatory variables). As for the territory, we concentrate on those areas where the collective process of recombinant knowledge creation is more likely to occur: metropolitan regions (Boschma et al., 2015; Balland and Rigby, 2017). For this purpose, we exploit the recently developed EUROSTAT classification of Metropolitan Regions (MR) as the unit of reference.

Each patent is assigned to a repository (the network, the applicant, which mostly is a firm rather than an individual, and the MR) with a whole count. This means that, for example, if a patent application is assigned to more than one applicant, the stock count of each of these applicants increases of one unit rather than half – as it would be for fractional counts. One peculiar characteristic of knowledge, i.e. knowledge indivisibility, supports this approach, which is free of problematic assumptions about the allocation of knowledge creation effort and product among producers. After the assignation to each repository, the patent stock is discounted every year with a 15% depreciation factor (the so-called Permanent Inventory Method).

5.1. Collaboration Network

The definition of the collaboration network needs to be set. The inventor's network patent stock is the sum of the individuals' patent history who collaborated with the focal inventor within a 5-year window. Other definitions may be adopted, but it seems reasonable to suppose that a past collaboration remains an active source of knowledge for a 5-year period at most (Breschi and Lenzi, 2016).

5.2. Firm and Establishment

We explore two different definitions of applicant/firm, namely the worldwide applicant and the establishment. In so doing, we try to better account for the localized nature of knowledge production, i.e. its embeddedness into a territory. For each applicant/firm, we compute its knowledge activity irrespectively of the territory where that activity has taken place, or, on the contrary, we define the tuple "firm/applicant ID – Metropolitan Region" as the visible sign of a localized establishment of the firm. In patent documents, indeed, only the headquarter address is reported. For example, IBM will have two (or more) knowledge stock variables associated with: one indicating its knowledge potential as a worldwide brand, the other, varying across the different territories where IBM operated, measuring the local knowledge pool stemming from IBM knowledge activities in that territory – say, Barcelona Metropolitan Region.

5.3. Metropolitan Region (MR)

EUROSTAT defines MR as "NUTS 3 regions or a combination of NUTS 3 regions which represent all agglomerations of at least 250 000 inhabitants. These agglomerations were identified using the Urban Audit's Functional Urban Area (FUA)". In turn, FUA identifies a city of > 250 000 inhabitants plus its commuting zone. We adopt this classification, but we add some areas excluded by the original MR definition, which emerged as relevant according to patent production rates (e.g.

Cambridge Area). Our criterion was to retrieve back those MRs whose yearly patent production belonged to the top quartile of the patent distribution.

5.4. Controls

We want to assure that the variables referred to the stock of patents at different levels only measure the layer's knowledge capacity and its externalities dynamics, rather than the intensity of innovativeness. With this aim, we compute a set of patent-based control variables for productivity, measured as the average inventor's productivity at each layer. In order to smooth temporal disturbances, we compute the inventor's productivity on a time window between *t* and *t-4*. Then, as for the stock variables, the network's productivity (*NETprod*) is computed on a 5-year window from *t-1* backwards, whereas the firm level variables (*FIRMprod* and *ESTABprod*) are computed on a 3-year window from *t-1* backwards in order to minimize missing values in the lag variable (very few firms invent more than once in consecutive years). MR average productivity (*MRprod*) enters the regressions as the respective value for each MR at *t-1*.

Evidence of the importance of multinational firms in affecting firms and territorial productivity and knowledge capacity is growing (Iammarino and McCann, 2013) so much that Crescenzi, Gagliardi and Iammarino (2015) states that "MNEs are amongst the main 'creators' of new technology [...] since they represent the largest source of technology generation, transfer and diffusion in the world economy". Therefore, we control if a firm is a multinational with a dichotomous dummy variable (*Multin*). Even though we aim at controlling for other relevant socio-economic regional variables, most of them are not available for a long time window and at our territorial unit of analysis, without incurring in a large number of missing values. Specifically, the Cambridge Econometrics (CE) database partially provides NUTS3 level data (that can be translated to MR) on GVA, population and employment, which we use to compute gross value-added per capita (*GVA PC*), population density (*Pop Dens*) and an index of employment specialization (*HH index*). Even though we can impute some of the missing values, CE does not contain data at all for Switzerland, which is, instead, an important provider of inventors in our final dataset. We show regressions with CE controls in the robustness checks. As for R&D expenditure, which is a standard input of the KPF, EUROSTAT data are profoundly incomplete, so much that imputation is not advisable. We are still working on a solution.

5.5. Matching

The inventor-applicant-MR matching process generated multiple ambiguous assignations, e.g. more than one MR or applicant for inventor-year. In order to operate with unique assignation for each inventor-year, we set up some decision rules for the disambiguation algorithms we use: the one for the applicant ID and for the MR ID. The rationale behind these algorithms is continuity, i.e. we want to detect when mobility patterns of inventors across firms and MR are too frequent to be realistic, and we assign more weight to long-lasting ties in case of plausible ambiguous assignations (Hoisl, 2007; Nakajima et al., 2010).

Our final dataset results in a strongly unbalanced panel of

- 243320 multiple inventors (inventors that applied for patents more than once)
- 58665 applicants (mainly firms)
- 282 MR
- over a total of 30 years covered, from 1980 to 2010.

5.6. Descriptive evidence

Figures in the Appendix provide some descriptive statistics. Figure 1 shows that even though the vast majority of inventors apply for one patent a year at most, there are significant groups producing more than one application a year (left panel). Similarly, most inventors appear only twice in the dataset, but the number of multiple inventors appearing more than twice is not negligible. Figures 2 to 5 display the trends of knowledge stock across layers. Overall, the stocks of knowledge are increasing but heterogeneously across the different units of analysis. Among these figures, Figure 3 shows that there are many inventors without any past collaboration. Figures 6 to 8, instead, plot some boxplots for the relationship between inventor productivity (*appXinv*) and one knowledge stock variable for each layer. An overall positive relationship is observed, even though it is more evident for *NETstock*, decreasing in intensity for *MRstock* and more ambiguous for *ESTABstock* (it is very similar in the case of firms). Given the high skewness of *appXinv* distribution, boxplots obviously show outliers. However, the percentage of outliers on the total number of observations for each variable is low and uninfluential on the results (not shown but available upon request).

6. Results

***Table 1**. Main results.*

4-ways FE OLS regression with clustered SE by MR							
	Dependent variable:						
	Log of application per Inventor-Year						
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
MR Stock (lag 1)	0.043***	0.035***	0.033***	0.033**	0.032**	0.032**	0.031**
	(0.013)	(0.012)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)
Establishment Stock (lag 1)		0.008***	0.006***	0.018***	0.015***	0.018***	0.014***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Network Stock (lag 1, 5-years window)			0.006**	0.012***	-0.002	0.007***	-0.006***
			(0.002)	(0.005)	(0.003)	(0.002)	(0.002)
Establishment Avg Productivity (lag 1, 3-years window)				-0.082***	-0.080***	-0.084***	-0.081***
				(0.004)	(0.005)	(0.005)	(0.004)
MR Avg Productivity (lag 1, 3-years window)				0.050	0.053	0.052	0.054
				(0.044)	(0.046)	(0.044)	(0.046)
Network Avg Productivity (lag 1, 5-years window)				-0.015**	-0.010*		
				(0.007)	(0.006)		
Multinational Firm (T/F)				0.029***	0.029***	0.029***	0.029***
				(0.003)	(0.003)	(0.003)	(0.004)
*Network Stock * Establishment Stock (lag 1)*					0.003***		0.003***
					(0.001)		(0.001)
*Network Stock * MR Stock (lag 1)*					-0.002		-0.001
					(0.001)		(0.001)
*Establishment Stock (lag 1) * MR Stock (lag 1)*					-0.004***		-0.004***
					(0.001)		(0.001)
Observations	766,414	766,359	766,359	766,281	766,281	766,281	766,281
R^2	0.485	0.485	0.485	0.487	0.488	0.487	0.488
Adjusted R^2	0.193	0.193	0.194	0.197	0.197	0.196	0.197
Residual Std. Error	0.334	0.334	0.334	0.334	0.333	0.334	0.333
Note:	*$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$*						

Table 2. *By technological macro-area.*

4-ways FE OLS regression with clustered SE by MR, technological area (OST7) sub-samples

	Dependent variable:						
	Log of application per Inventor-Year						
	Electrical eng.	Instruments	Chemicals	Pharma/ Bio	Indust. processes	Mech. eng.	Consumer goods
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MR Stock (lag 1)	0.010	0.010	0.044***	0.056**	0.042***	0.027	0.039*
	(0.019)	(0.021)	(0.015)	(0.025)	(0.015)	(0.019)	(0.021)
Establishment Stock (lag 1)	0.019***	0.018***	0.022***	0.025***	0.021***	0.016***	0.021***
	(0.003)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)
Network Stock (lag 1)	0.010**	0.016***	0.026***	0.035***	0.024***	0.015**	0.028***
	(0.005)	(0.005)	(0.005)	(0.008)	(0.005)	(0.006)	(0.007)
Establishment Avg Productivity (lag 1)	-0.093***	-0.090***	-0.082***	-0.091***	-0.091***	-0.083***	-0.105***
	(0.007)	(0.007)	(0.007)	(0.010)	(0.006)	(0.005)	(0.008)
MR Avg Productivity (lag 1)	0.137*	0.117*	0.095	0.095	0.074	0.044	0.147**
	(0.078)	(0.063)	(0.062)	(0.085)	(0.058)	(0.060)	(0.073)
Network Avg Productivity (lag 1)	-0.007	-0.015	-0.028***	-0.054***	-0.026***	-0.016	-0.036***
	(0.009)	(0.010)	(0.009)	(0.014)	(0.009)	(0.012)	(0.013)
Multinational Firm (T/F)	0.031***	0.037***	0.039***	0.044***	0.032***	0.028***	0.030***
	(0.007)	(0.005)	(0.006)	(0.008)	(0.004)	(0.005)	(0.008)
Observations	282,203	290,951	297,532	171,513	303,185	268,603	133,534
R^2	0.485	0.483	0.476	0.497	0.481	0.488	0.503
Adjusted R^2	0.216	0.208	0.241	0.253	0.226	0.210	0.223
Residual Std. Error	0.357	0.358	0.363	0.375	0.356	0.342	0.356
Note:					$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$		

Given the high number of observations and levels we aim at simultaneously controlling for, we are still working on the results. In particular, at the very best of our knowledge, there is no prior work trying to apply MA to unbalanced longitudinal data qualified by a 3-levels hierarchical cross-classified structure. At this moment, we still have not solved the puzzle.

Preliminary results of the 4-way FE regressions are shown in Table 1. Models 1 to 4 progressively plugs in explanatory variables. In Model 5, instead, interaction terms enter the regression. Since the standalone coefficients of interacted variables must be read as their effect when the variables they are interacting are set at zero, we decided to centre *MRstock* only on its mean, because, differently from *ESTABstock* and *NETstock*, it only rarely reaches zero. Models 6 and 7 present the same regressions without *NETprod*. Indeed, as the correlation matrix in Table 1 shows, the network average productivity and knowledge stock are highly collinear. Although none of our variables performs with a VIF higher than 10 (the usual threshold for multicollinearity detection) we prefer to run regressions both with and without the dubious variable.

Models 1 to 3 suggest that inventors take advantage of the knowledge pools at each layer they are embedded in. Simultaneous consideration of the three relevant knowledge flows level let us better gauge the respective coefficients than previous research. In terms of magnitude, MR knowledge externalities are still much stronger than network externalities; *MRstock* regression coefficient shrinks but still keeps its lead. Model 4, instead, suggests that inventors seem to suffer competition. The negative and significant signs of controls regarding the establishment and network productivity are a signal that peers innovativeness tends to inhibit inventors capacity to produce new knowledge, instead of providing a stimulating work environment. Interestingly, in Model 5, complementarity between network and establishment knowledge pools emerges, whereas establishment and MR knowledge pools appear as a substitute and no significant relationship is detected between MR and network knowledge stocks. Across all models, working in an establishment owned by a multinational firm affects inventors' productivity positively.

This preliminary evidence seems to suggest that each of the three knowledge externalities sources that we highlighted as relevant exert a positive and independent influence on inventors' creativity. The inventor's institutional environment and the web of past collaborations display a strong synergy, suggesting that the inventor's ability to exploit efficiently the external knowledge coming from his network also depends on the knowledge capacity of his work environment. Past collaborators may overlap with current workplace colleagues, but only partially: the inventor's network may stretch well beyond firm and regional boundaries. On the contrary, the workplace and the territorial settings where inventors reside appear as substitute channels for knowledge flows (the interaction term is negative). Indeed, when *ESTABstock* is zero, *MRstock* keeps its positive influence on inventor productivity, and the reverse is true as well. However, when both of them are present and rising longitudinally, the total benefit on productivity from a rise in externalities is not equal to the sum of the two increases in respective stocks, but significantly lower. More investigation is needed to understand this offset better. Somewhat surprisingly, Model 7 reports that when *ESTABstock* is zero and *MRstock* is on its mean, an increasing *NETstock* produces negative effects on inventors productivity, when not controlling for *NETprod*. Such unexpected result may partially be imputable to a strong synergy between the establishment and the network knowledge stocks, i.e. valuable knowledge collected from past collaborators interactions is productive only when the inventor's creative potential is plunged into a knowledge intense work environment.

5.1 Technological differences

Preliminary results confirm the hypothesis concerning the existence of different channels/levels where knowledge is circulating on. In order to grasp some specificity of each of these knowledge externalities mechanisms, we split the sample by technological areas according to inventors affinity. We exploit the OST reclassification of IPC codes into 7 macro-categories. Since providing a unique inventor-OST7 assignation is passive, to a large extent, to a subjectivity fallacy in the formulation of the matching procedure, we assign inventors to more than one sub-sample if they invented in many OST7. Table 2 shows the results. Indeed, there is a substantial degree of heterogeneity across technological macro-areas. It provides valuable information about the specific peculiarities of each knowledge channel, which we interpret and classify according to two axes: the form of the knowledge exchanged, in the fashion set by Polanyi (1958) counterpoising tacit and codified knowledge, and the mode of the exchange, which happens through a formal rather than informal interaction.

The results from Table 2 tell a story where inventors in engineering technological areas do not take advantage of territorial knowledge externalities, whereas those inventing in Chemicals, Pharmaceuticals and Processes do. One possible interpretation is that salient technological knowledge in engineering disciplines is mostly tacit, thus sticky with inventors' expertise. Tacit knowledge circulates through experience-sharing and repetitive interactions, which is not the mode of interactions at MR level. Tacit and sticky knowledge transfer, instead, requires that kind of secure and repetitive environment taking place within collaborations usually framed by contracts. MR externalities emerge through a complexity of informal and occasional encounters continually taking place along time, indeed they qualify a territorial endowment. These informal and widespread knowledge exchange occasions generate a particularly suitable *milieu* for codified knowledge and information about knowledge to flow. Chemicals, Pharmaceutical and Biotech, in particular, are fields where knowledge codification is both crucial and advanced, so that MR externalities may take place along with other levels. The establishment level lays in between the two and synthetizes them: both formal and informal interactions take place in a secure environment where collaborations are framed within incomplete contracts.

5.2 Robustness: firm level variables and CE controls

In the Appendix, Tables 4 and 5 report the robustness checks. First, we substituted the establishment level variables with the firm level ones. Results are very similar to those of Table 2. In Table 5 CE controls for GVA per capita (economic performance), population density (agglomeration economies) and employment HH index (industrial specialization) are added, but none of them turns statistically significant and results are unchanged.

7. Conclusions

Understanding the mechanisms of knowledge diffusion is a relevant topic in the economics of innovation since knowledge creation is at the very base of the innovative dynamics. This research contributes to enriching such understanding, appreciating the complexity of the social structure where inventors are embedded in. We provide a novel contribution in many respects. First, we analyse individual inventors' capacity to create new knowledge with a large, longitudinal dataset. In so doing, we apply the precious inheritance of the economics of knowledge, mainly dealing with regions and firms, to the individuals. Second, we exploit the new EUROSTAT classification of the European territories by Metropolitan Regions in order to target more efficiently than before the actual locus of knowledge production. Third, we account simultaneously for what we believe are the three most fundamental levels where knowledge might flow. Combing those contributions together, we are able to state that, even after accounting for the knowledge potential delivered by the network of collaborators, and the knowledge capacity of the firm where inventors do invent, territorial knowledge externalities stand up as a significant and sizeable force enhancing knowledge production. The establishment knowledge capacity emerges as a pivotal point: it not only exerts a positive standalone effect on inventors' productivity, but it also proves crucial for the effectiveness of the network knowledge stock and appears as an alternative source of knowledge respect to the territorial environment.

The empirical investigation leaves space for improvements. First, the multilevel structure of our data (and reality) might be better appreciated with ad-hoc hierarchical regression models. However, addressing altogether the 4-level (time, inventor, firm, MR) structure and the unobserved heterogeneity at the individual level in a longitudinal framework needs a tailored solution still to come up. Second, our results concerning the sub-samples by technological area point the direction for future research, in order to appreciate the heterogeneity of knowledge production not only across levels but across fields as well. Third, even though the canonical socio-economic controls drawn from the CE database do not add anything to the results, we still believe R&D expenditure is a missing information. Fourth, our results may hint the existence of a mediation effect of the network knowledge stock respect to higher level knowledge stocks, an aspect that we plan to deepen in future research.

References

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

Almeida, P. and Kogut, B. (1999) Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Management Science*, **45**, 905–917.

Antonelli, C. (2000) Collective Knowledge Communication and Innovation: The Evidence of Technological Districts. *Regional Studies*, **34**, 535–547.

Antonelli, C. (2008) Pecuniary knowledge externalities: The convergence of directed technological change and the emergence of innovation systems. *Industrial and Corporate Change*, **17**, 1049–1070.

Antonelli, C. (2016) Digital knowledge generation and the appropriability trade-off. *Telecommunications Policy*.

Antonelli, C. and Colombelli, A. (2017) The locus of knowledge externalities and the cost of knowledge. *Regional Studies*, **51**, 1151–1164.

Antonelli, C., Crespi, F., Mongeau Ospina, C.A. and Scellato, G. (2016) Knowledge composition, Jacobs externalities and innovation performance in European regions. *Regional Studies*, 1–13.

Antonelli, C., Patrucco, P.P. and Quatraro, F. (2008) The governance of localized knowledge externalities. *International Review of Applied Economics*, **22**, 479–498.

Arrow, K. (1962) Economic welfare and the allocation of resources for invention. *The rate and direction of inventive activity: Economic and social factors* pp. 609–626. Princeton University Press.

Asheim, B.T. and Coenen, L. (2005) Knowledge bases and regional innovation systems: Comparing Nordic clusters. *Research Policy*, **34**, 1173–1190.

Asheim, B.T. and Coenen, L. (2006) Contextualising regional innovation systems in a globalising learning economy: On knowledge bases and institutional frameworks. *The Journal of Technology Transfer*, **31**, 163–173.

Audretsch, D.B., Falck, O., Feldman, M.P. and Heblich, S. (2012) Local Entrepreneurship in Context. *Regional Studies*, **46**, 379–389.

Audretsch, D.B. and Feldman, M.P. (1996) R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, **86**, 630–640.

Audretsch, D., B., Lehemann, E.E. and Hinger, J. (2015) From knowledge to innovation: the role of knowledge spillover entrepreneurship. *Routledge Handbook of the Economics of Knowledge*, Routledge p. Antonelli C. and Link A. N., New York.

Audretsch, D.B. and Stephan, P.E. (1999) Knowledge spillovers in biotechnology: sources and incentives. *Journal of Evolutionary Economics*, **9**, 97–107.

Balland, P.-A., Boschma, R. and Frenken, K. (2015) Proximity and Innovation: From Statics to Dynamics. *Regional Studies*, **49**, 907–920.

Balland, P.-A. and Rigby, D. (2017) The Geography of Complex Knowledge. *Economic Geography*, **93**, 1–23.

Bell, A., Chetty, R., Jaravel, X., Petkova, N. and Van Reenen, J. (2017) Who Becomes an Inventor in America? The Importance of Exposure to Innovation. *NBER Working Paper No. 24062*.

Bell, A., Fairbrother, M. and Jones, K. (2016) Fixed and Random effects: making an informed choice. *Under Review*.

Bell, A. and Jones, K. (2015) Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. *Political Science Research and Methods*, **3**, 133–153.

Boschma, R. (2005) Proximity and Innovation: A Critical Assessment. *Regional Studies*, **39**, 61–74.

Boschma, R., Balland, P.-A. and Kogler, D.F. (2015) Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. *Industrial and Corporate Change*, **24**, 223–250.

Breschi, S. and Lenzi, C. (2016) Co-invention networks and inventive productivity in US cities. *Journal of Urban Economics*, **92**, 66–75.

Breschi, S. and Lissoni, F. (2001) Knowledge Spillovers and Local Innovation Systems: A Critical Survey. *Industrial and Corporate Change*, **10**, 975–1005.

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

Cameron, A.C. and Miller, D.L. (2015) A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, **50**, 317–372.

Clarke, P., Crawford, C., Steele, F. and Vignoles, A.F. (2010) The choice between fixed and random effects models: some considerations for educational research. *The Centre for Market and Public Organisation 10/240*.

Coffano, M. and Tarasconi, G. (2014) CRIOS-Patstat database: sources, contents and access rules. *Center for Research on Innovation, Organization and Strategy, CRIOS working paper*, **1**.

Cohen, W.M. and Levinthal, D.A. (1990) Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, **35**, 128.

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

Cowan, R. and Jonard, N. (2003) The dynamics of collective invention. *Journal of Economic Behavior & Organization*, **52**, 513–532.

Cowan, R. and Jonard, N. (2004) Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, **28**, 1557–1575.

Crépon, B., Duguet, E. and Mairesse, J. (1998) Research, Innovation And Productivi: An Econometric Analysis At The Firm Level. *Economics of Innovation and New Technology*, **7**, 115–158.

Crescenzi, R., Gagliardi, L. and Iammarino, S. (2015) Foreign multinationals and domestic innovation: Intra-industry effects and firm heterogeneity. *Research Policy*, **44**, 596–609.

Dosi, G. (1982) Technological paradigms and technological trajectories. *Research Policy*, **11**, 147–162.

Fazio, G. and Piacentino, D. (2010) A Spatial Multilevel Analysis of Italian SMEs' Productivity. *Spatial Economic Analysis*, **5**, 299–316.

Fernández-Zubieta, A., Geuna, A. and Lawson, C. (2016) Productivity pay-offs from academic mobility: should I stay or should I go? *Industrial and Corporate Change*, **25**, 91–114.

Fleming, L. (2001) Recombinant Uncertainty in Technological Search. *Management Science*, **47**, 117–132.

Fleming, L. and Sorenson, O. (2001) Technology as a complex adaptive system: evidence from patent data. *Research policy*, **30**, 1019–1039.

Fleming, L. and Sorenson, O. (2004) Science as a map in technological search. *Strategic Management Journal*, **25**, 909–928.

Fleming, L. and Szigety, M. (2006) Exploring the Tail of Creativity: An Evolutionary Model of Breakthrough Invention. *Advances in Strategic Management* pp. 335–359. Emerald (MCB UP), Bingley.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., et al. (2007) Inventors and invention processes in Europe: Results from the PatVal-EU survey. *Research Policy*, **36**, 1107–1127.

Graf, H. (2012) Inventor networks in emerging key technologies: information technology vs. semiconductors. *Journal of Evolutionary Economics*, **22**, 459–480.

Grant, R.M. (1996) Toward a knowledge-based theory of the firm: Knowledge-based Theory of the Firm. *Strategic Management Journal*, **17**, 109–122.

Griliches, Z. (1979) Issues in Assessing the Contribution of Research and Development to Productivity Growth. *The Bell Journal of Economics*, **10**, 92.

Griliches, Z. (1995) R&D and productivity: econometric results and measurement issues. *Handbook of the economics of innovation and new technology*, Blackwell pp. 52–89. Stoneman, P., Oxford.

Griliches, Z. (1998) *R & D and Productivity: The Econometric Evidence*. University of Chicago Press, Chicago.

Grillitsch, M., Martin, R. and Srholec, M. (2016) Knowledge Base Combinations and Innovation Performance in Swedish Regions. *Economic Geography*, 1–22.

Hidalgo, C.A. and Hausmann, R. (2009) The building blocks of economic complexity. *proceedings of the national academy of sciences*, **106**, 10570–10575.

Hoisl, K. (2007) Tracing mobile inventors—The causality between inventor mobility and inventor productivity. *Research Policy*, **36**, 619–636.

Hussinger, K. (2012) Absorptive capacity and post-acquisition inventor productivity. *The Journal of Technology Transfer*, **37**, 490–507.

Iammarino, S. and McCann, P. (2013) *Multinationals and Economic Geography: Location, Technology and Innovation*. Edward Elgar, Cheltenham.

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

Jaffe, A.B., Trajtenberg, M. and Henderson, R. (1993) Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, **108**, 577–598.

Latham, W., Bas, C.L., Hassane, J.R.B. and Volodin, D. (2011) Interregional mobility, productivity and the value of patents for prolific inventors in France, Germany and the UK. *International Journal of Foresight and Innovation Policy*, **7**, 92.

López-Bazo, E. and Motellón, E. (2017) Innovation, heterogeneous firms and the region: evidence from Spain. *Regional Studies*, 1–15.

Mariani, M. and Romanelli, M. (2007) “Stacking” and “picking” inventions: The patenting behavior of European inventors. *Research Policy*, **36**, 1128–1142.

Melero, E. and Palomeras, N. (2015) The Renaissance Man is not dead! The role of generalists in teams of inventors. *Research Policy*, **44**, 154–167.

Miguélez, E. and Moreno, R. (2013) Skilled labour mobility, networks and knowledge creation in regions: a panel data approach. *The Annals of Regional Science*, **51**, 191–212.

Nakajima, R., Tamura, R. and Hanaki, N. (2010) The effect of collaboration network on inventors’ job match, productivity and tenure. *Labour Economics*, **17**, 723–734.

Nelson, R.R. and Winter, S.G. (2004) *An Evolutionary Theory of Economic Change*, digitally reprinted. The Belknap Press of Harvard Univ. Press, Cambridge, Mass.

Pakes, A. and Griliches, Z. (1998) Patents and R&D at the Firm Level: A First Look. *R & D and productivity: the econometric evidence* pp. 55–72. University of Chicago Press, Chicago.

Palomeras, N. and Melero, E. (2010) Markets for Inventors: Learning-by-Hiring as a Driver of Mobility. *Management Science*, **56**, 881–895.

Polanyi, M. (1958) *Personal Knowledge: Towards a Post-Critical Philosophy*, Nachdr. Univ. of Chicago Press, Chicago.

Raspe, O. and van Oort, F.G. (2011) Firm heterogeneity, productivity and spatially bounded knowledge externalities. *Socio-Economic Planning Sciences*, **45**, 38–47.

Raudenbush, S.W. (2009) Adaptive centering with random effects: An alternative to the fixed effects model for studying time-varying treatments in school settings. *Education*, **4**, 468–491.

Raudenbush, S.W. and Bryk, A.S. (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods*, 2nd ed. Sage Publications, Thousand Oaks.

Rodríguez-Pose, A. and Crescenzi, R. (2008) Research and Development, Spillovers, Innovation Systems, and the Genesis of Regional Growth in Europe. *Regional Studies*, **42**, 51–67.

Rothaermel, F.T. and Hess, A.M. (2007) Building Dynamic Capabilities: Innovation Driven by Individual-, Firm-, and Network-Level Effects. *Organization Science*, **18**, 898–921.

Schettino, F., Sterlacchini, A. and Venturini, F. (2013) Inventive productivity and patent quality: Evidence from Italian inventors. *Journal of Policy Modeling*, **35**, 1043–1056.

Schmoch, U., Laville, F., Patel, P. and Frietsch, R. (2003) Linking technology areas to industrial sectors. *Final Report to the European Commission, DG Research*.

Schumpeter, J.A. (1939) *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*, 1. ed., [reprint]. Bartleby's Books [u.a.], Chevy Chase, Md.

Singh, J. (2005) Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science*, **51**, 756–770.

Singh, J. and Fleming, L. (2010) Lone Inventors as Sources of Breakthroughs: Myth or Reality? *Management Science*, **56**, 41–56.

Taylor, A. and Greve, H.R. (2006) Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal*, **49**, 723–740.

Weitzman, M.L. (1998) Recombinant growth. *The Quarterly Journal of Economics*, **113**, 331–360.

Wuchty, S., Jones, B.F. and Uzzi, B. (2007) The Increasing Dominance of Teams in Production of Knowledge. *Science*, **316**, 1036–1039.

Zucker, L.G., Darby, M.R. and Armstrong, J. (1998) Geographically localized knowledge: spillovers or markets? *Economic Inquiry*, **36**, 65–86.

Zwick, T., Frosch, K., Hoisl, K. and Harhoff, D. (2017) The power of individual-level drivers of inventive performance. *Research Policy*, **46**, 121–137.

Appendix

***Figure 1**. Even though the vast majority of inventors apply for one patent a year, there are significant groups producing >1 application a year (left panel). Similarly, most inventors appear only twice in the dataset, but the number of multiple inventors appearing >2 is not negligible.*

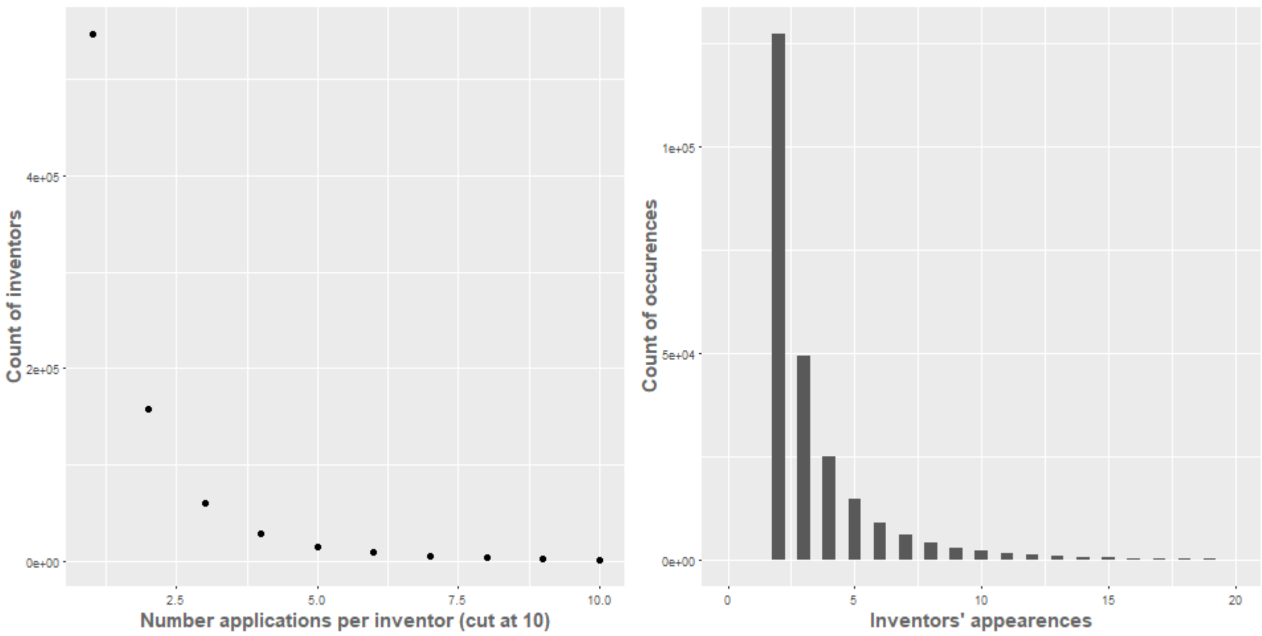


***Figure 2**. Mean Firm and Establishment Knowledge Stock across years and MRs. The overall Pearson Correlation Coefficient is 0.85, indicating that the two measures are indeed strongly related, but still they measure slightly different aspects*

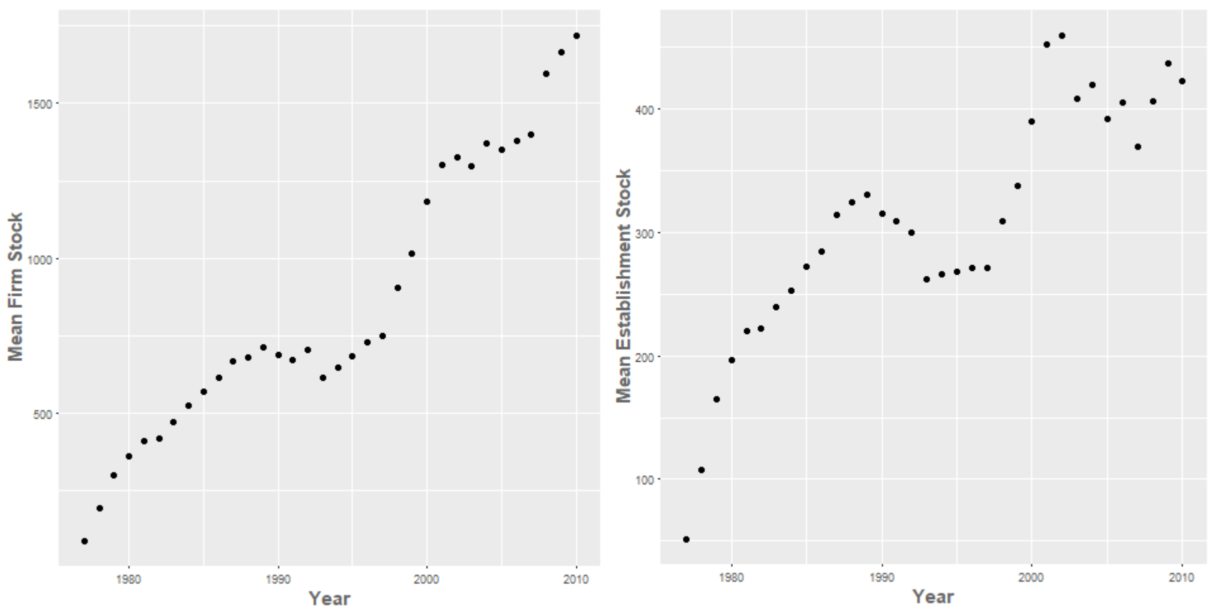


Figure 3*. Mean of the Network Stock and Network Size, 1st and 3rd quartile*

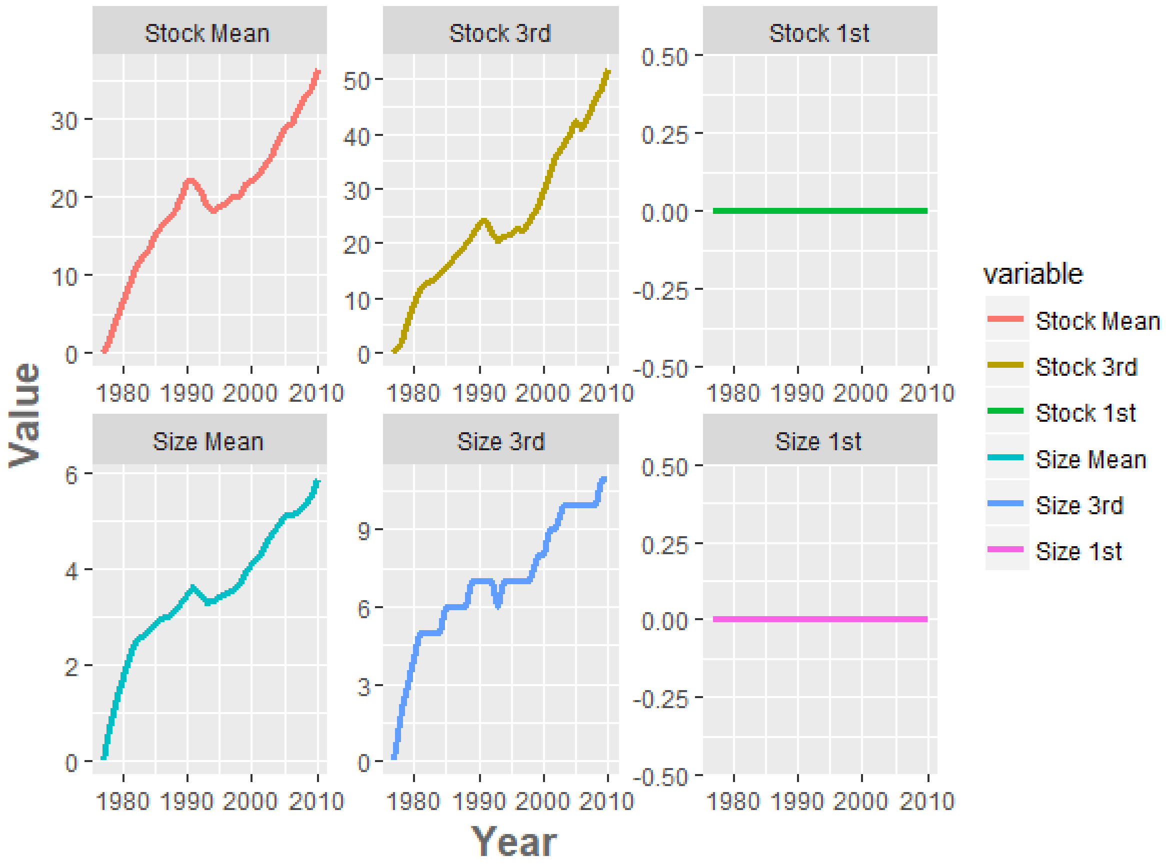


Figure 4*. Mean Establishment Stock across Countries.*

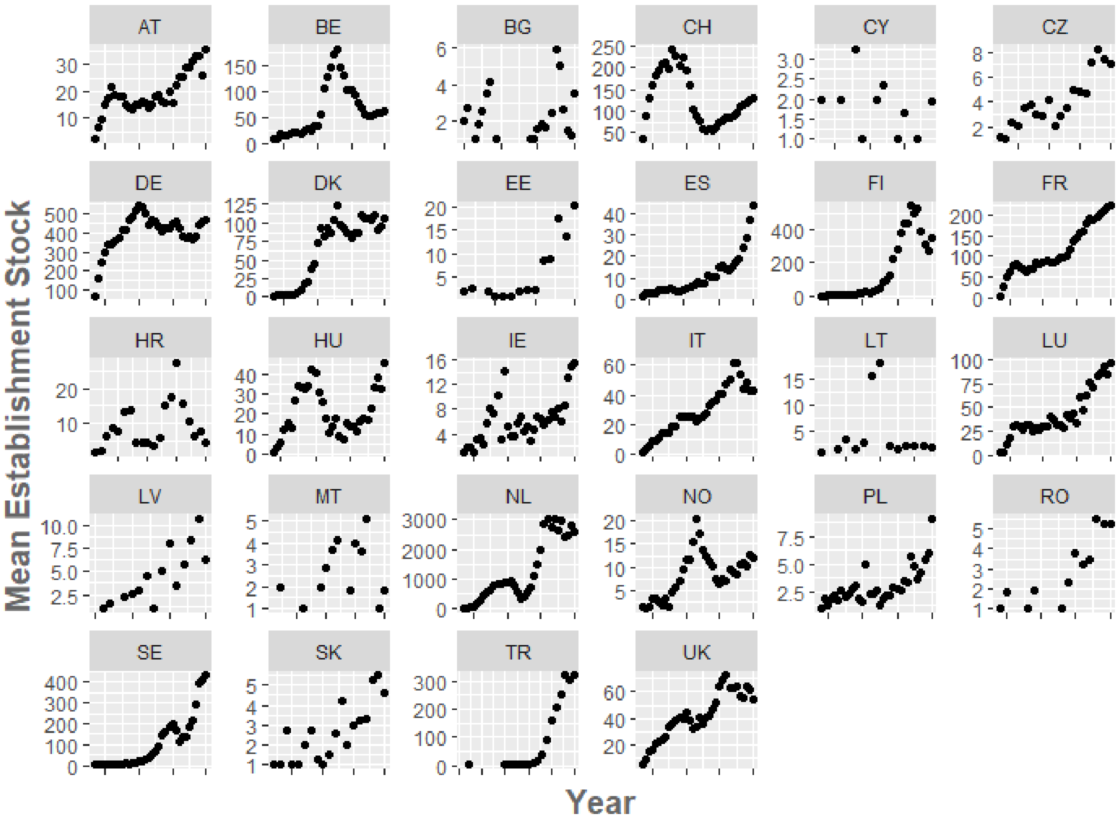


Figure 5*. Mean MR Stock across Countries.*

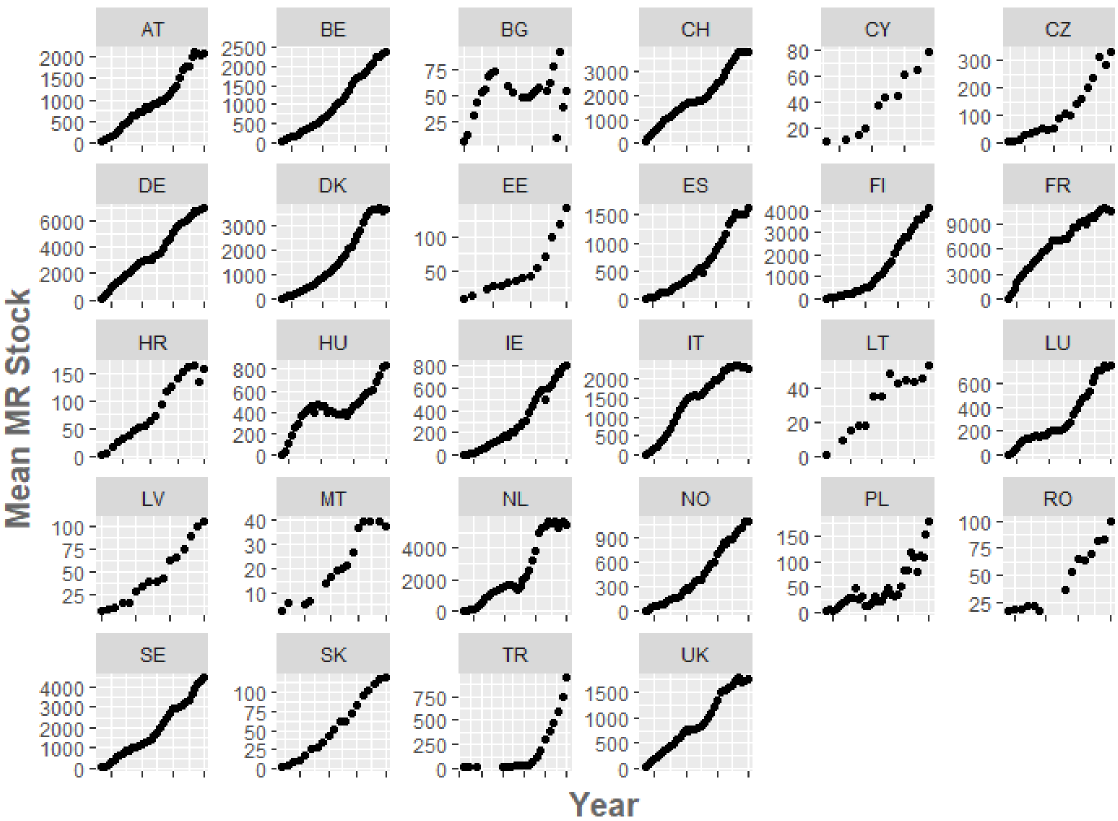


Figure 6. *Boxplot (1).*

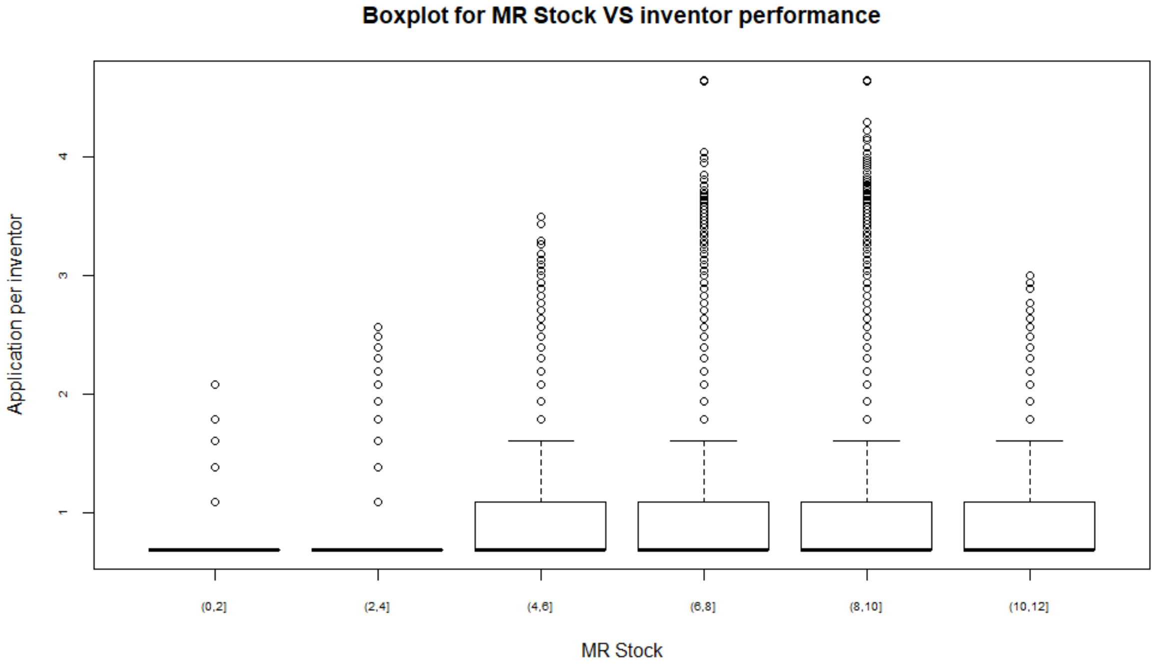


Figure 7. *Boxplot (2).*

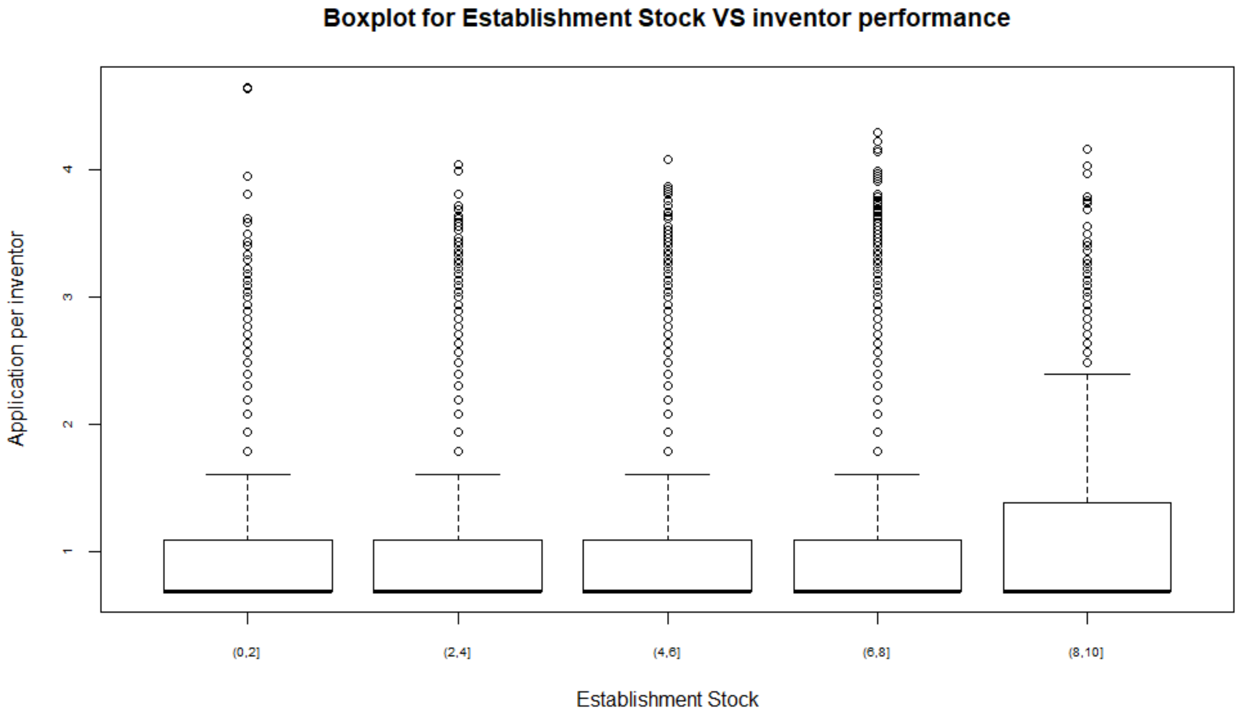


***Figure 8**. Boxplot (3).*

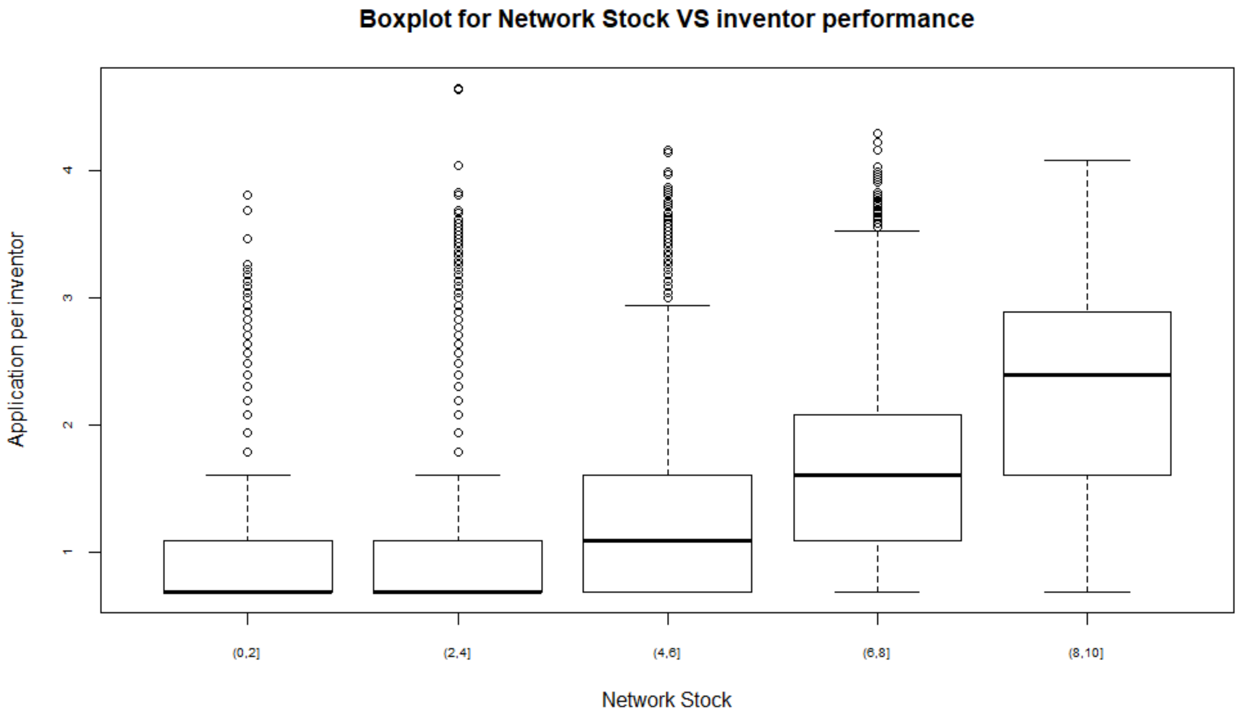


***Table 3.** Descriptive statistics.*

All variables in log

Statistic	N	Mean	St. Dev.	Min	Max
Application x inventor	777,807	0.916	0.371	0.693	4.644
MR Stock (lag 1)	777,807	7.693	1.287	0.693	10.057
Establishment Stock (lag 1)	765,555	3.563	2.200	0.693	8.828
Firm Stock (lag 1)	766,414	4.649	2.507	0.693	9.556
Network Stock (lag 1, 5-years window)	777,807	1.519	1.657	0.000	8.640
Establishment Avg Productivity (lag 1, 3-years window)	777,807	0.808	0.368	0.000	4.244
Firm Avg Productivity (lag 1, 3-years window)	777,807	0.865	0.333	0.000	5.221
Network Avg Productivity (lag 1, 5-years window)	777,807	0.593	0.564	0.000	5.260
MR Avg Productivity (lag 1)	777,807	0.911	0.087	0.693	1.694

Figure 9. *Correlation Matrix (pairwise complete observations). Variables are in log.*

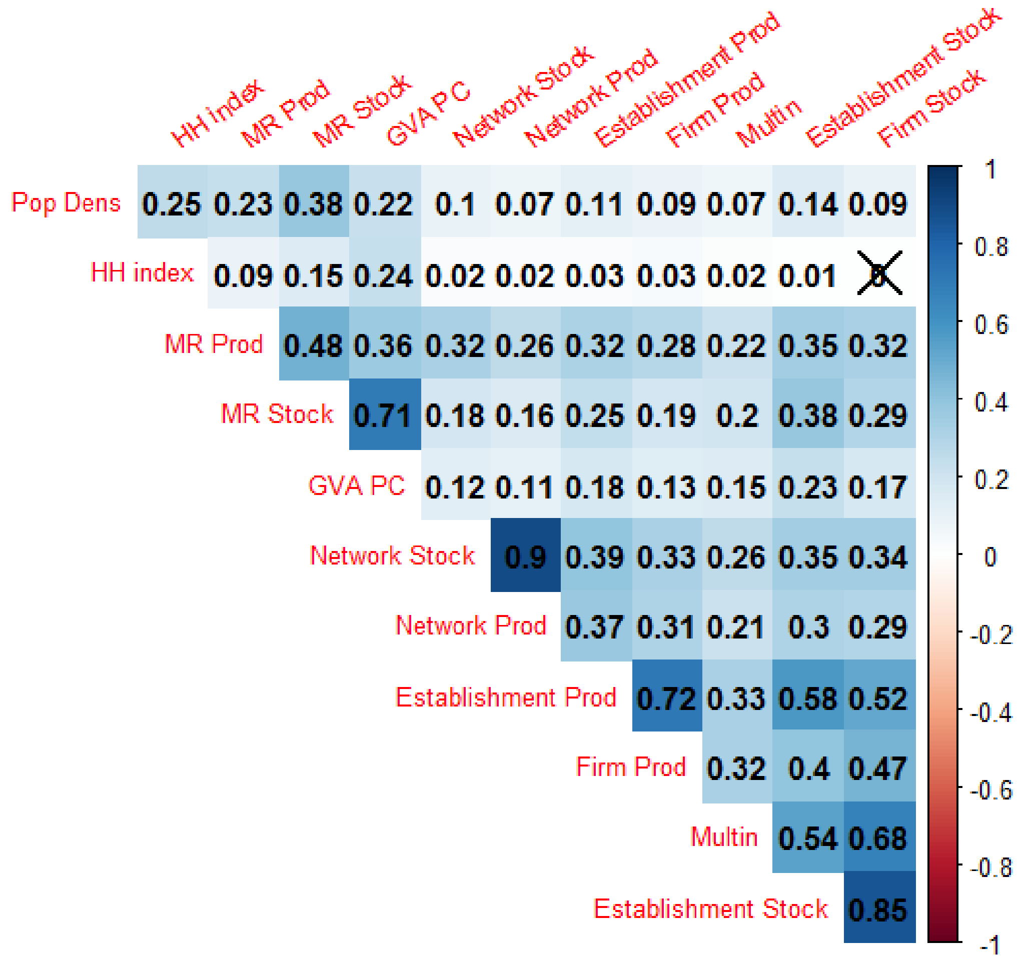


***Table 4**. Robustness check (1)*

4-ways FE OLS regression with clustered SE by MR. Robustness check with FIRMstock

	Dependent variable:			
	Log of application per Inventor-Year			
	(1)	(2)	(3)	(4)
MR Stock (lag 1)	0.037***	0.035***	0.041***	0.040***
	(0.012)	(0.012)	(0.012)	(0.012)
Firm Stock (lag 1)	0.016***	0.016***	0.014***	0.014***
	(0.003)	(0.003)	(0.003)	(0.003)
Network Stock (lag 1, 5-years window)	0.013***	0.007***	-0.009***	-0.014***
	(0.005)	(0.002)	(0.003)	(0.003)
Firm Avg Productivity (lag 1, 3-years window)	-0.079***	-0.081***	-0.071***	-0.072***
	(0.005)	(0.006)	(0.005)	(0.005)
MR Avg Productivity (lag 1, 3-years window)	0.045	0.048	0.041	0.042
	(0.044)	(0.044)	(0.046)	(0.046)
Network Avg Productivity (lag 1, 5-years window)	-0.017**		-0.011*	
	(0.007)		(0.006)	
Multinational Firm (T/F)	0.029***	0.029***	0.029***	0.029***
	(0.003)	(0.003)	(0.004)	(0.004)
Network Stock * Firm Stock (lag 1)			0.003***	0.004***
			(0.001)	(0.001)
Network Stock * MR Stock (lag 1)			-0.001	-0.001
			(0.001)	(0.001)
Firm Stock (lag 1) * MR Stock (lag 1)			-0.002**	-0.002**
			(0.001)	(0.001)
Observations	766,336	766,336	766,336	766,336
R^2	0.487	0.487	0.487	0.487
Adjusted R^2	0.196	0.196	0.196	0.196
Residual Std. Error	0.334	0.334	0.334	0.334
Note:	*p<0.1; **p<0.05; ***p<0.01			

Table 5. *Robustness check (2)*

4-ways FE OLS regression with clustered SE by MR. Robustness check with CE controls

	Dependent variable:	
	Log of application per Inventor-Year	
	(1)	(2)
MR Stock (lag 1)	0.031**	0.030**
	(0.014)	(0.014)
Establishment Stock (lag 1)	0.018***	0.015***
	(0.002)	(0.002)
Network Stock (lag 1, 5-years window)	0.012**	-0.003
	(0.005)	(0.003)
Establishment Avg Productivity (lag 1, 3-years window)	-0.083***	-0.080***
	(0.004)	(0.005)
MR Avg Productivity (lag 1, 3-years window)	0.047	0.050
	(0.046)	(0.048)
Network Avg Productivity (lag 1, 5-years window)	-0.014*	-0.009
	(0.008)	(0.006)
Multinational Firm (T/F)	0.029***	0.029***
	(0.004)	(0.004)
GVA PC (lag 1)	0.006	0.003
	(0.043)	(0.044)
Population density (lag 1)	-0.234	-0.112
	(0.335)	(0.342)
Employment Specialization (lag 1)	-0.246	-0.241
	(0.170)	(0.177)
Network Stock * Establishment Stock (lag 1)		0.003***
		(0.001)
Network Stock * MR Stock (lag 1)		-0.002
		(0.001)
Establishment Stock (lag 1) * MR Stock (lag 1)		-0.004***
		(0.001)
Observations	725,417	725,417
R^2	0.488	0.488
Adjusted R^2	0.198	0.198
Residual Std. Error	0.334	0.334
Note:		*p<0.1; **p<0.05; ***p<0.01

***Table 6**. OST30 and OST7. Source: Schmoch et al. 2003 (Schmoch et al., 2003)*

OST30-code	OST30-name	OST7-code	OST7-name
1	Electrical engineering	1	Electrical engineering; Electronics
2	Audiovisual technology	1	Electrical engineering; Electronics
3	Telecommunications	1	Electrical engineering; Electronics
4	Information technology	1	Electrical engineering; Electronics
5	Semiconductors	1	Electrical engineering; Electronics
6	Optics	2	Instruments
7	Technologies for Control/Measures/Analysis	2	Instruments
8	Medical engineering	2	Instruments
9	Nuclear technology	2	Instruments
10	Organic chemistry	3	Chemicals; Materials
11	Macromolecular chemistry	3	Chemicals; Materials
12	Basic chemistry	3	Chemicals; Materials
13	Surface technology	3	Chemicals; Materials
14	Materials; Metallurgy	3	Chemicals; Materials
15	Biotechnologies	4	Pharmaceuticals; Biotechnology
16	Pharmaceuticals; Cosmetics	4	Pharmaceuticals; Biotechnology
17	Agricultural and food products	4	Pharmaceuticals; Biotechnology
18	Technical processes (chemical, physical, mechanical)	5	Industrial processes
19	Handling; Printing	5	Industrial processes
20	Materials processing, textile, glass, paper	5	Industrial processes
21	Environmental technologies	5	Industrial processes
22	Agricultural and food apparatuses	5	Industrial processes
23	Machine tools	6	Mechanical eng.; Machines; Transport
24	Engines; Pumps; Turbines	6	Mechanical eng.; Machines; Transport
25	Thermal processes	6	Mechanical eng.; Machines; Transport
26	Mechanical elements	6	Mechanical eng.; Machines; Transport
27	Transport technology	6	Mechanical eng.; Machines; Transport
28	Space technology; Weapons	6	Mechanical eng.; Machines; Transport
29	Consumer goods	7	Consumer goods; Civil engineering
30	Civil engineering	7	Consumer goods; Civil engineering

Social Capital, the Knowledge Filter and innovative start-ups: an Italian evidence

Matteo Tubiana[a,b]

Alessandra Colombelli[b,c]

[a] *Department of Economics and Statistics Cognetti de Martiis, University of Torino*

[b] *BRICK, Collegio Carlo Alberto*

[c] *DIGEP, Politecnico di Torino*

Abstract

According to the Knowledge Filter and Knowledge Absorptive Capacity Theory of Entrepreneurship, the appropriation and exploitation of knowledge externalities by potential entrepreneurs is hindered by social, institutional and individual constraints. We argue that the territorial endowment of Social Capital works as a moderator of these constraints, sometimes favouring, sometimes impeding the effective exploitation of knowledge-related resources, which in turn are the main driver of innovative start-up creation. However, Social Capital is an ambiguous notion in its definition and a problematic item to measure. Building on the sociological tradition, we adopt the instrumental definition of Social Capital, and exploit the ISTAT AVQ survey to measure its multiple dimensions, taking advantage of Principal Component Analysis. New data about innovative Italian start-ups from 2009 to 2015 made available by the Italian Government allow us to evidence, indeed, a positive moderation effect of the intensity of friendly relationship and a negative moderation effect of political participation on patterns of start-up creation at the NUTS3 regional level in Italy.

1. Introduction

A large body of the economic literature brought evidence of the importance of entrepreneurship for economic growth (Audretsch, 1995). Particularly, small nascent firms (start-ups) evidenced as an important introducer of radical innovations (Aghion and Howitt, 1992; Audretsch, Keilbach and Lehemann, 2006). The Knowledge Spillover Theory of Entrepreneurship (henceforth KSTE) is emerging the main theory of the link between the presence of knowledge externalities and their effective exploitation by entrepreneurs, hence explaining various phenomena such as the localization of both knowledge and entrepreneurship (Audretsch, 1995; Audretsch, Lehemann and Hinger, 2015). Recently, the KSTE has been expanded in order to refine the understanding of the context's role in shaping entrepreneurial knowledge appropriation and exploitation (Acs and Plummer, 2005; Qian and Acs, 2013; Colombelli and Quatraro, 2018). With this paper, we queue on this research line, providing a further extension of the KSTE through its implementation with Social Capital theory (henceforth SC).

The literature has already explored, to some extent, the connection between SC, or social interactions in a broader sense, and entrepreneurship (Davidsson and Honig, 2003; Liao and Welsch, 2005; Stuart and Sorenson, 2005; Chiesi, 2007). However, we are not aware of any study accounting for the role of SC in knowledge diffusion, thus sustaining knowledge-driven entrepreneurship. Within the KSTE framework, we argue that the territorial SC endowment of Italian provinces acts as a moderator for diffusion and exploitation of relevant resources for knowledge-driven start-ups.

In order to exploit the concept of SC we must face two difficulties. The first concerns the ambiguity of the notion of SC. We align to some works in the literature advocating for a recovery of the instrumental, structural meaning, against a wider sense embracing civicness and trust (Trigilia, 2001; Durlauf and Fafchampes, 2004). Nonetheless, SC remains a multidimensional concept. Once the theoretical definition and its relationship with the KSTE is settled, the second difficulty regards SC measurement. Indeed, it is not possible to observe directly SC and to measure it neatly. A big variety of proxy have been used in the literature. We opt for an application of Principal Component Analysis to a set of raw survey variables, drawn from ISTAT AVQ[1], ending up to equip us with a set of indexes approximating different mechanisms of SC formation (Sabatini, 2009; Righi, 2013).

[1] I dati utilizzati nel presente lavoro sono di fonte Istat e relativi alle rilevazioni Multiscopo "Aspetti della Vita Quotidiana" dal 2008 al 2015. Le elaborazioni sono state condotte presso il Laboratorio

Using data about Italian NUTS3 regions, we find that SC may both facilitate and hinder knowledge-related resources effectiveness respect to start-up creation, depending on the empirical dimension measured. Indeed, we explore the multiple facets of SC, after circumscribing its definition to the structural and relational dimensions. We find that the intensity of friendly relationship generates a conducive environment for technological knowledge, human capital and the presence of technology incubators for start-ups' positive effects, whereas political participation appears as hindering them.

In Sections 2 and 3 the relevant reference literature is summarized, whereas Section 4 develops a link between them. Section 5 spans our data collection and the construction of the relevant variables, eventually presenting the empirical model. Finally, Section 6 collects the results and Section 7 concludes.

2. KSTE, the knowledge filter and the entrepreneurial absorptive capacity

A large amount of literature has addressed the relevance of knowledge spillovers to regional development. Within this area of investigation, a first attempt to link knowledge spillovers and new firm formation was made by Audretsch (1995), Audretsch and Lehmann (2005), who have articulated the Knowledge Spillover Theory of Entrepreneurship (henceforth KSTE). This theory moves from a critique to endogenous growth models (Lucas, 1988; Romer, 1990) and adds to the theories of entrepreneurship that have focused on the ability of individuals to recognize opportunities (Venkataraman, 1997; Shane and Venkataraman, 2000) by linking knowledge spillovers and the entrepreneurship theory.

The Knowledge Spillover Theory of Entrepreneurship (KSTE) argues that knowledge generated by incumbents is not necessarily commercialised by these organisations, generating entrepreneurial opportunities for new firms. Incumbent organizations are often unable or unwilling to fully appropriate and commercialize new knowledge and ideas generated within their research laboratories as they lack capabilities or do not want to take the risks of introducing radically new

per l'Analisi dei Dati ELEmentari dell'Istat di Torino e nel rispetto della normativa in materia di tutela del segreto statistico e di protezione dei dati personali. I risultati e le opinioni espresse sono di esclusiva responsabilità dell'autore, non costituiscono statistica ufficiale e non impegnano in alcun modo l'Istat.

technologies onto the market, and prefer to focus on small improvements in their existing products and processes. As a consequence, an opportunity to start a new firm is generated in order to exploit and commercialize that knowledge and those ideas. In this context, the start-up of a new firm is a mechanism in which knowledge spillovers from the organization can create opportunities that the new firm can exploit. The KSTE thus suggests that the start-up of a new firm is an endogenous response to opportunities that have been generated, but not fully exploited, by incumbents. The KSTE is based on the idea, developed by Arrow, (1962) that knowledge, unlike traditional production factors, is characterized by non-excludability and non-exhaustibility. This implies that knowledge is not fully appropriable and may spill over from the organization that produces it to a new organization Griliches (1992). An important implication of the KSTE is that contexts characterized by greater amounts of knowledge generate more entrepreneurial opportunities.

Later, in the attempt to understand how new firms are created as a response to knowledge spillovers, Acs et al. (2004), Acs and Plummer (2005), Audretsch et al. (2006) and Braunerhjelm et al. (2010), have proposed the concept of the knowledge filter. According to this approach, the spillover of knowledge is not automatic, as conceptualized in the endogenous growth theory (Romer 1990) but is instead hindered by the so-called 'knowledge filter'. In this view, the presence of unused knowledge pools can be considered as a necessary, but not sufficient, condition for the actual exploitation of such entrepreneurial opportunities. The transformation of knowledge stocks into economically useful knowledge requires the presence of enabling conditions at the local level. These conditions pertain to the existence of supporting institutions, knowledge intermediaries, regulatory frameworks and appropriate financial markets. The absence of these conditions could create a barrier – the knowledge filter - that hinders the transformation of knowledge into economic knowledge à la Arrow. New firms may serve as a conduit for knowledge spillovers, insofar as the features of local contexts allow them to penetrate the knowledge filter. This concept encompasses the basic characteristics of knowledge set forth by Arrow (1962), although it is broader in scope. According to Audretsch (2007): it is the outcome of "the characteristics of knowledge distinguishing it from information, a high degree of uncertainty combined with non-trivial asymmetries, combined with a broad spectrum of institutions, rules and regulations" (Audretsch, 2007: p.67). Therefore, the knowledge filter generates a gap between the creation of knowledge and its commercialization through the establishment of new ventures.

A further step in the literature has been made by Qian and Acs (2013) who has proposed the absorptive capacity theory of entrepreneurship. The absorptive capacity theory adds to the KSTE that the extent to which the market value of new knowledge is discovered and exploited depends on the capability of entrepreneurs to recognize such opportunities. This theory advances the KSTE by introducing absorptive capacity as a critical determinant of knowledge-based entrepreneurial activity. Entrepreneurial absorptive capacity is defined as the ability of an entrepreneur to understand new knowledge, recognize its value, and subsequently commercialize it by creating a firm (Qian and Acs, 2013). The extent to which the market value of new knowledge is discovered and exploited depends on the capability of entrepreneurs to recognize such opportunities.

In short, the KSTE-related literature discussed so far implies that regional variations in the availability of knowledge are associated with differential rates of new firm formation. However, a local abundance of knowledge does not necessarily lead to its commercialization through new ventures. Regional variations, in terms of supporting institutions, regulations and entrepreneurial absorptive capacity, may, in fact, create a filter that affects the likelihood of prospective entrepreneurs actually succeeding in exploiting the market opportunities provided by unexploited knowledge.

In this paper, we maintain that the regional social capital can be regarded as a factor that may contribute to mitigate or magnify the gap between the creation of new knowledge and its commercialization by newly born firms.

3. Social Capital Theory

There is not a clear and unique definition of SC. The economic literature dealing with the concept is vast and heterogeneous and encompass different facets of SC concept, like norms, trust and altruism, cooperative behaviour, family and informal ties, business networks and networks exploitation.

Leaving apart the classical foundations (Tosini, 2005), the explicit history of the concept traces back to Coleman (1990), Loury (1976, 1987) and Bourdieu (1986). With due differences among them, their definitions of SC is instrumental and pertains strictly to the societal structure of individuals' relationships[2]. As Lin (2002) states, “Bourdieu, Coleman [..], all share the understanding that social capital consists of resources embedded in social relations and social structure, which can be mobilised when an actor wishes to increase the likelihood of success in a purposive action.” On this vein, one of the most renowned works on networks effects on economically relevant issues is that of Granovetter on networks and job search (Granovetter, 1974).

Some years later, Putnam published two of the most prominent work on the matter in economics (Putnam, Leonardi and Nanetti, 1994; Putnam, 2000). In these works, the authors enlarge the concept of SC considerably, defined as “networks, norms and trust – that enable participants to act together more effectively to pursue shared objective” (Putnam, 2000). Similarly, Fukuyama (1996) stresses the role of communitarian norms of trust, cooperation and honesty in fostering prosperity because they smooth frictions in economic interactions. In these definition provided by political scientists, rather than sociologists, SC loses its instrumental flavour in favour of a more systemic, civic trait (Guiso, Sapienza and Zingales, 2010).

There is a divide between sociologists and political scientists’ approaches to SC, which is often underestimated. According to the formers, SC facilitates the actions of individuals, whereas for the latter it regards institutions at large (Hardin, 1999; Westlund and Bolton, 2003).

The Social Capital Initiative[3] follows the political science path defining SC as a two-facets entity: structural SC entailing information sharing, collective action and decision-making, and cognitive SC that regards norms, trust, beliefs, etc. The first being relatively more objective and measurable than the second (Grootaert and van Bastelaer, 2001; Grootaert, Van Bastelaer and World Bank, 2002). However, some harsh critics have been moved towards the SC account (e.g. see Portes, 1998 and Solow, 2000). Trigilia (2001) and Durlauf and Fafchampes (2004) emerges for clarity. The former is a sociological book where the authors spend much effort criticizing Putnam’s operationalization of SC, advocating for a purely relational, instrumental definition. The former, instead, is a massive

[2] (Bourdieu, 1986): “The aggregate of the actual or potential resources which are linked to possession of a durable network of more or less institutionalised relationship of mutual acquaintance and recognition – or in other words to membership in a group – which provides each of its member with the backing of the collectivity-owed capital, a “credential” which entitles them to credit, in the various senses of the word”. Coleman (1990): “a variety of entities having two characteristics in common: They all consist of some aspect of a social structure, and they facilitate certain actions of individuals who are within the structure.”

[3] The Social Capital Initiative has been funded by the Government of Denmark in 1996 and held by the World Bank, whose aim was to “assess the impact of social capital on the effectiveness of development projects and to contribute to the development of indicators for monitoring social capital and methodologies for measuring its impact” (Grootaert and van Bastelaer, 2001).

review of the economic empirical literature on the matter, which is, according to the author, mostly unable to identify empirically the effect of SC on various dependent variables at the various level of aggregation. The background of the authors are very distant, but they similarly sustain that confounding trust and networks misleads statistical conclusions and make theories vague. Most importantly, Durlauf and Fafchampes (2004) provides an economic definition of the consequences of SC in welfare terms and mentions three channels of influence: “First, [SC] may resolve a coordination failure in an economy that has multiple Pareto-ranked equilibria. Second, it may alter individual incentives so as to replace the decentralized equilibrium with a superior one. Third, it may affect the technology of social exchange, for instance by opening new avenues for the circulation of information.” And again: “The literature has identified a number of channels by which social capital improves efficiency. Most of these channels fall under one or a combination of the following three categories: information sharing, group identity, and explicit coordination.”

Nooteboom (2006) tries to define SC respective to the positioning of the three pillars upon which the political scientists’ definition relies (networks of relations, trust and norms, and institutions): “My proposal now is to define social capital as a source of reliance intermediate between institutions (macro) and relationship-internal features (micro) [..]. Perhaps this helps to make the concept of social capital less expansive and vague, and more determinate, and to clarify its difference with and its relation to institutions, trust and what happens within relationships. Recall that social capital lies in voluntary and informal intermediaries, i.e. not externally imposed and not supported by formal authority and official sanctions, which would make them part of institutions.” In 1998, Nahapiet and Ghoshal proposed to extend the structural and cognitive dimensions with a relational one. Structural and relational forms differ in that the former is a quantitative measure of the network shape and extension, whereas the latter defines as “those assets created and leveraged through relationships”, i.e. the qualitative measure of a tie in the network.

Reordering the inputs provided up to now, SC emerges surely as a multifaceted concept. For clarification purpose, we elaborated a summary table that tries to bring together different theoretical contributes to the theory of SC. The table, in the end, appears to juxtapose nicely.

***Table 1**. Graphical summary of SC definitions. Content inspired from Chiesi (2007), Nooteboom (2006), Durlauf and Fafchamps (2004)*

		Micro	*Meso*	*Macro*
		Network	Club good	Public good
Forms	Structural	Info and Know Sharing		
	Relational		Group Identity	
	Cognitive			Explicit Coordination

Each row represents one of the three forms or dimensions of SC as often defined in the literature. Each column instead refers to a different level of analysis of the concept and its socio-economic nature. Blank cells do not mean that a given form does not pertain to a given level of analysis at all. We gave labels to those intersections named in the literature. The dimensions of SC intertwine in complex ways (see Liao and Welsch, 2005 and Schenkel, D'Souza and Matthews, 2012); the partitioning is purely artificial but useful for analytical purposes.

The basic skeleton of SC is the structure of the system's network, which is micro-funded because the individuals' relationships are the channels conveying SC effectiveness. The core business of this SC form is information and knowledge sharing. Such a business is crucial once one recognizes that information asymmetries and transaction costs induce market failure (Akerlof, 1970; Stiglitz, 2002). The information gain is typically involuntary: it appears to be an externality of the network (Coleman, 1990).

The relational dimension mirrors the quality of the relationships, which ranges from the typology – familiar, friendship, business, etc. – to the timing – intensity, regularity, etc. It is analytically useful because enlightens the role of SC in affecting preferences and choices through group identity. The power of group identity expresses within the circumscribed network of insiders. In this case, SC does not flow freely through the relationships within the system, rather it sticks within the group as a typical club good.

Lastly, the cognitive form of social capital pertains to norms and general attitudes sharing, and trust, all of these smoothing interactions and social exchanges, but also sustaining a respectful exploitation of common resources (Nahapiet and Ghoshal, 1998). For example, Owen-Smith and Powell (2004) found that when a network is dominated by public research organizations, instead of private institutes, informational sharing is stronger and patenting activity of single organizations is higher. The "explicit coordination" label underlines that this public good dimension operates directly on the management of territorial dynamics, rather than implicitly and indirectly as in the case of individuals interactions.

We follow the line traced by Trigilia (2001), Durlauf and Fafchampes (2004) and Nooteboom (2006) making clear that the cognitive dimension, usually approximated with measures of trust and compliance to civic norms, overloads the theoretical definition of SC. We adopt the original, instrumental definition of SC, which, in the summary Table 1, encompasses the structural and relational dimensions. Similarly, Guiso et al. (2010) dissected the SC definition and circumscribed the set of principles typically assigned to the cognitive dimension, defining them "civic capital". Another argument in favour of our choice is that, as Nooteboom (2006) suggests, trust and compliance fall under the discussion upon the role of institutions in the system functioning set by North (1990). Civic capital or institutions are an interesting and relevant topic to investigate. However, we suggest that they are confounding the SC definition, which is instead rather clear according to the sociological, instrumental line of thought.

4. Social Capital as knowledge moderator

We suggest that SC is a knowledge moderator. The knowledge filter and the entrepreneurial absorptive capacity concepts help to appreciate the dynamics of such a facilitating effect. Recalling the review of section 2, it emerges that the knowledge filter deploys its hindering effect at the system level, whereas the entrepreneurial absorptive capacity regards the individual entrepreneur directly. SC may affect both levels. Stuart and Sorenson (2005) define entrepreneurship as involving two tasks: the identification of promising opportunities and the mobilization of resources, i.e. gaining access to them. Once the structural and relational dimensions of SC are appreciated, together with the channels of their economic effects as described in Table 1, SC appears to affect positively both tasks. Indeed, social ties are the conduits for information to flow (identification of opportunities) and they prompt the circulation of the entrepreneurially relevant financial capital, skilled labour and knowledge.

Liao and Welsch (2005), drawing on the antecedent literature, distinguish between “ability” and “propensity” to enterprise. They argue that structural SC positively affects the ability to enterprise, whereas relational SC mainly influences the potential entrepreneur’s propensity. From the specific point of view of innovative, knowledge-intensive start-ups, SC as a resources mobilizer facilitates individuals to penetrate the knowledge filter, enhancing their ability to enterprise, whereas SC as a conduit for information and knowledge flows empowers individuals’ entrepreneurial absorptive capacity (“individuals ability to understand and absorb new technologies”, Qian and Acs, 2013). The resources mobilized through SC are not exclusively related to technological knowledge, however we focus on the specific SC – technology-related resources relation that is most important for innovative start-ups.

The KSTE is a fundamental advance in our understanding of a very problematic tile of the economic growth puzzle: the mechanisms through which newly generated knowledge fosters innovative activities and, consequentially, economic growth. The developing of the entrepreneurial absorptive capacity signals the need to expand further the KSTE to refine better the knowledge generation, diffusion, appropriation and exploitation processes at the micro-level (Qian, Acs and Stough, 2012; Qian and Acs, 2013; Qian and Jung, 2017). In particular, a fertile research avenue is the investigation of the relationship between the entrepreneur and its context, both from a technological (Colombelli and Quatraro, 2018) and a societal point of view (what this research aims at).

The inclusion of SC theory within the KSTE framework may provide fertile insights. For example, it provides a further basis to explain the territorial embeddedness of entrepreneurship (Audretsch et al., 2012). Knowledge spillovers have been shown to be strongly related to networks of collaborations and interactions among knowledge workers (see e.g. Breschi and Lissoni, 2001, 2009). Ferrary and Granovetter (2009) made clear that a territory endowed with valuable entrepreneurial resources becomes an enduring engine of breakthrough (e.g. Silicon Valley) only when these entrepreneurial resources are embedded in a complex system of interactions. This complex system of interactions, i.e. SC, is not easily transferable. It follows that entrepreneurial activity concentrates in space.

In the empirical sections of the paper, we will test the hypothesis that SC is indeed a moderator of the relationship between knowledge and innovative start-up formation. We will explore its potential to moderate the effect of the relevant technological resources for entrepreneurship highlighted in the KSTE literature: the stock of previous technological knowledge, human capital availability, plus an institutional variable – technology incubators presence. Even though we expect a globally positive effect of the SC territorial endowment, the literature calls for caution. Given its non-linear, multifaceted nature, SC may exert both a enhancing and an inhibiting effect, depending on the dimension considered (Westlund and Bolton, 2003). In this regard, the ability/propensity duality appears as a good candidate to frame specific, dimensional effects. Indeed, for a potential entrepreneur to start up a firm, resources availability is necessary but not sufficient; prospective scenarios must also motivate or enhance the propensity to enterprise. If the extension and strength of the network is expected to exert an exclusively positive moderation effect, its typology, which may augment the club good side of SC, can, instead, generate bonding effects that, for innovative knowledge-intensive firms are detrimental because hinder resources circulation and limit prospective scenarios.

5. Empirical Analysis

In the following subsections, we will describe our data sources and the composition of the final dataset. If the data about start-up creations are made available and easy to use by the Italian Chamber of Commerce, data concerning SC needs instead a more dedicated processing. Finally we will discuss the knowledge-related resource for entrepreneurship as suggested by the KSTE theory and some control variable for the estimation. The last paragraph will describe the empirical strategy.

5.1 The Italian Framework and the Italian Data (Thesis Chapter, The Italian case)

The empirical exercise of this research deals with Italian data because of two reasons: first, there are not many empirical studies on Social Capital and entrepreneurship in Italy, which is a country in need of stronger firms' development towards innovative outcomes. Second, starting from 2012 precious data about start-ups have been collected, providing a good opportunity to investigate entrepreneurship at the territorial level.

At the end of 2012 the Italian Ministry of Economic Development approved a Law Decree on "Further urgent measures for Italy's economic growth", providing for specific measures which are aimed at promoting the creation and development of innovative start-ups. This was the first time the Italian legislation took this kind of companies into consideration. The law recognizes that start-ups are important for the promotion of sustainable growth, technological development and employment, in particular youth employment, and aims at developing an environment that foster the creation of entrepreneurial opportunities, innovation and social mobility; strengthen the universities and businesses links; attract investments and talented people from abroad. Under this law, at the end of 2016 more than 4735 innovative start-ups registered at the Chambers of Commerce in Italy.

In order to be included in the register of "innovative start-ups" and to benefit from governmental incentives, a new company needs to fulfil some requirements. In particular, according to the Law Decree definition, a start-up is a corporation that has a turnover lower than 5 million euros, has been operational for less than 48 months and whose ownership is direct and in hand of physical subject for at least 51%. Most importantly, its social aim is the development of innovative products or services with a high technological content.

In order to satisfy this latter requirement and classify as innovative, the start-up needs to fulfil at least one out of three criteria. Either 15% of its costs are R&D activities related, or at least one third of the team is made up of highly qualified members; or, finally, the enterprise is the holder, depositary or licensee of a registered patent or the owner of a program for original registered computers.

All the companies included in the register of "innovative start-ups" benefit from the support measures provided by the Law Decree. For example, the possibility to use start-up's specific flexible employment contracts; to remunerate their team members and the providers of external services with stock options and work for equity, respectively; and to access to incentives for the employment of highly qualified personnel. Moreover, the Law Decree introduces a "fail fast" procedure with the aim to give the entrepreneur the chance to start a new business project as soon as possible.

In addition, the Italian Government, in the attempt to stimulate entrepreneurial activities, provides some specific measures and incentives for incubators or accelerators that fulfil specific requirements concerning the start-up's physical structures, management, facilities and its track record and also aims at increasing the resources available for venture capital.

The Chamber of Commerce provides the data about registered start-ups and incubators, with many registry information like year and territory of foundation. Thanks to this freely available source we can compute the number of new start-ups (*Startup*) for every year *t* and every territorial province *i*, corresponding to the EUROSTAT NUTS3 level, as well as the number of established incubators at the same levels (*Incub*). We exploit innovative start-ups data from 2009 to 2015.

5.2 Principal Component Analysis on Social Capital

The ISTAT AVQ survey is part of a series of multi-purpose surveys on families conducted by ISTAT since 1993 on a yearly basis. Its aim is monitoring aspects of citizens' habits and judgments on social and institutional life that should be meaningful for their quality of life. Exploring variables of SC is not a primary objective of the survey. However, following previous works of Righi (2013), we extrapolate from available items those better approximating our interests. At the very best of our knowledge, this is the only (partially freely) available source of NUTS3 level data on SC for Italy[4].

The crucial issue about SC measurement is that rarely we observe SC directly. Ideally, we would want a precise measure for the structural and relational dimensions. However, such measurements are not viable. Researchers need to approximate SC measurement with proxy variables, as much as innovation economists exploit patents data to measure innovative activities. In the literature, there is a clear divide between those researches using one single measure to address SC (Crescenzi, Gagliardi and Percoco, 2013) and those applying data-reduction methodologies to generate various indexes, altogether representing various aspects of the SC endowment (Sabatini, 2009; Righi, 2013). We queue on the second line. We build on the hypothesis that social participation generates SC. Antoci, Sabatini and Sodini (2013) sustain – following Coleman's tradition – that structural SC is a by-product of social interactions. SC is an endowment generated both by individual investments in social relations and, primarily, by sparse social interactions. Similarly, Sabatini (2009) states that SC "may be seen as an improvement in the technology of production of relational goods" and as a "by-product" of social participation (relational goods are those peculiar "goods" that can be enjoyed only if shared and whose consumption and production is simultaneous, see Gui and Sugden, 2010).

[4] Measurement is a special issue in SC research. Many useful surveys are available but none of them displays important characteristics altogether: vast territorial reach, in order to permit aggregate analysis; a design planned to measure separately the different dimensions of SC (clearly related to the lack of a consensual definition); a panel dimension. Different studies often exploit different survey designs or different census variables to measure the same dimensions, such that comparative literature review is not straightforward. Even in face of these complications, we believe that the sociality dimension is crucial for understanding economic issues, even though research conditions are sub-optimal. Carrying on with the effort will induce, we hope, better survey to be designed, data to be collected and methodologies to be drafted.

Individuals may participate to social life in various ways. Following Righi (2013), we selected eight that appeared to be the most suitable for investigating SC. They fall into four categories: social participation, participation in particularistic associations, political participation and friendship relations. Unlike Righi, (2013) we left out measures of generalized trust and civicness, as motivated in Section 2. The unit of analysis of this research is the NUTS3 territory, whereas ISTAT AVQ is an individual-level survey. Therefore, individual answers are aggregated into provinces. However, the ISTAT AVQ survey is designed to be representative at the NUTS2 level. In order to exploit this valuable data source, we build territorial weights and correct for disturbances in population composition according to age and sex[5].

In this research, we exploit PCA on these eight ISTAT AVQ variables in order to reduce the dimensionality of measurements and produce a set of SC indexes measuring the categories of social participation. Altogether, these indexes retrieve a picture of the territorial SC endowment[6]. We run a PCA analysis on the full set of eight SC variables for the whole sample. The first 4 Principal Components mirror the thematic subgroups outlined above: social participation, participation in particularistic associations (e.g. unions), political participation and friendship relations, even though only the first two Principal Components pass the Kaiser criteria (eigenvalue bigger than 1) (see Figures 2, 3 and 4 in the Appendix). Therefore, we extract the first two PCs (PCs are linear combination with PCA weights of all the variables included in the analysis) indicating respectively Social Participation (*SocPart*) and Political Participation (*PolPart*), and keep the raw variables that represent the better the third and fourth components: volunteering in unions (*volsi*) and frequency of amical relationships (*amici567*).

[5] Following the "Nota metodologica" available online at http://www.istat.it/it/archivio/129916, we apply a simplified version of their weighting methodology. Each observation, before the aggregation at NUTS3 level, is weighted according to the different compositional characteristics of the sample compared to the population, according to six categories of age and sex.

[6] PCA is a dimension-reducing technique, which serves to approximate a component that is not directly observed but correlated with some observed variables. The causality arrows goes from the observed variables to the index (the extracted Principal Component) approximating the unobserved dimension.

***Table 2**. SC variables description and division in thematic areas as a consequence of PCA.*

	NOME	DESCRIZIONE
PC1: Social Participation	volon	volunteering activities for voluntary organizations
	atgra	volunteering activities for non-voluntary organizations
	finas	financing organizations
PC2: Political Participation	volpa	volunteering activities for a political party
	comiz	participation to political assemblies
	finpa	financing political parties
Amical relationships	amici567	frequency of meetings with friends during free time (high frequency)
Particularistic Participation	volsi	volunteering activities for an union

5.3 Knowledge-related resources variables

Other than the number of new start-ups, and the set of SC principal components, we compute three knowledge-related resource variables approximating the amount of knowledge entrepreneurial opportunities in the province. One crucial resource for innovative firms' insurgence is the technological knowledge pool, which is radically embedded in the territory because of its tacitness and stickiness (Antonelli, 2000; Acs et al., 2009). The literature traditionally adopts the local expenditure for research and development (R&D) as a proxy. Unfortunately, there are no available data concerning R&D expenditure at the NUTS3 level in Italy. Therefore, we adopt the local knowledge stock (*Patent Stock*) calculated with the permanent inventory method on patent applications, with an obsolescence rate of 15% (Hall, Jaffe and Trajtenberg, 2005). Moreover, we calculated the number of technology incubators (*Incubator Stock*) in each province. Technology incubators are supposed to represent a key resource to the creation of new innovative firms and provide the conditions for successful undertakings and increase the survival likelihood (Colombo and Delmastro, 2002). Incubators' mission is to act as intermediaries in the market for finance, technology and advisory. They are a knowledge-related resource because they shield and redistribute competence and knowledge to those that have access to their services. Technology incubators are key participants to the technological knowledge recombination process (Weitzman, 1998). Finally, we control for the number of university graduates for province of residence (*HC*), inspired from the literature suggesting the pivotal role of human capital in entrepreneurial decision (Davidsson and Honig, 2003; Schenkel et al., 2012).

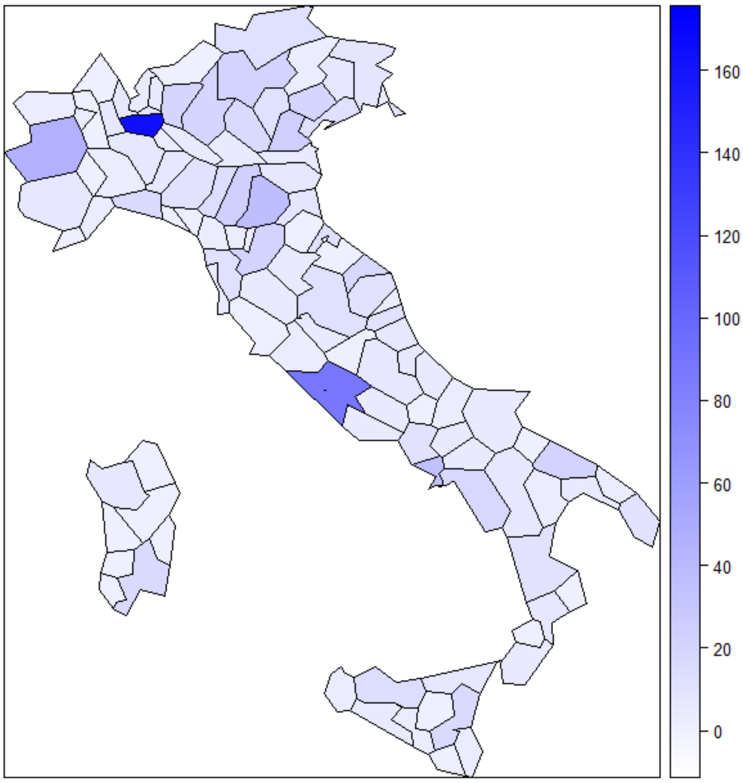
5.4 Controls

According to previous studies in this stream of literature, new firm formation may be triggered by pull factors such as high innovative potential (Acs and Audretsch, 1989; Geroski, 1995). Therefore, we control for the effects of agglomeration economies estimated by population density at the NUTS3 (*Density*). A complementary measure of prospective economic benefits is also represented by the distance (*Distance*) of each province from the administrative chief town of the NUTS2 region (Baptista and Mendonça, 2010; Bonaccorsi et al., 2013). Moreover, agglomeration economies can also stem from the presence of other firms in the same place, which ensures to some extent the availability of local markets for intermediate goods. In this direction, we added as a control variable the firm density (*Firm Density*) calculated as the ratio between the number of registered firms and the land use area. Another aspect of agglomeration economies regards their industrial composition, which signals the availability of pecuniary externalities. We account for this aspect with a Hirschman-Herfindahl index of specialization of employment broken down by industries (*HH*). In order to control for territorial size effects, regional workforce (*Workforce*) was included into the model. As the creation of new firms can be the outcome of an "escape from unemployment" strategy, we control for the unemployment rate at the local NUTS3 level (*Unempl rate*). In order to control for demand effects, we also included per capita value added (*GVA pc*). Lastly, in order to address the issue of geographical differences in terms of new firm creation, we plug in dummy variables at the NUTS2 level, plus we account for time trends with year dummies.

5.5 The dataset

Our final dataset consists of a slightly unbalanced panel of 110 provinces for a time-window spanning from 2009 to 2015. Data covers NUTS3 provinces start-ups creation, SC indexes and the set of controls described above. Table 3 reports some descriptive statistics of these variables. Note that *Startup* shows a SD much higher that its mean and a wide range between the minimum and maximum values, suggesting a skewed distribution (as suggested by the QQ plot, Figure 8 in the Appendix). Figure 1 plots the distribution of *Startup* across territories in the whole period.

***Figure 1**. Start-up count by province from 2009-2015*



The biplot in Figure 2 reports the position of each NUTS3 province in the two dimensional space traced by the PCs. Each point's size is due to that point contribution to the PC , whereas the colour and shape signal the macro-area it belongs to (ITC is Northwest, ITH is Northeast, ITI is the Centre, ITF the South and ITG the Islands). The same colour patterns are used in Figure 2. We think the most interesting thing emerging from these two pictures is that breaking down SC endowment let us overcome a rigid classification of territories by macro cultural areas. Indeed, even though points of the same colour are denser in certain areas of the graph, suggesting that some degree of cohesion is at work within macro-areas, they follow a sparse pattern. For example, red points, belonging to Northwest macro-area (ITC), distribute across the space, from high-negative to medium-positive values in the first PC (Dim1, that is *SocPart*) and quite the same respect to the second PC (Dim2, that is *PolPart*), even though they are denser close to the centre and first two quadrants. Similarly, the pairs plot does not indicate the presence of any coloured cluster. The takeaway is that, with high territorial granularity and heterogeneous description of SC, much more territorial heterogeneity (i.e. information) discloses. This is both an opportunity and a challenge for empirical analysis.

Table 3. Descriptive Statistics

All variables (except Startup) lagged 1 year

	N	Mean	St. Dev.	Min	Max
socPart	762	-0.06	1.65	-4.42	6.53
polPart	762	-0.16	1.26	-4.27	6.68
volsi	762	-0.05	1.00	-1.29	8.99
amici567	762	-0.14	0.99	-5.03	3.73
Log Patent Stock	776	4.81	1.54	0.00	8.65
Log Incubator Stock	776	0.17	0.36	0.00	2.20
Log HC	741	0.004	0.004	0.0000	0.02
Log GVA per capita	770	0.02	0.01	0.01	0.05
Log Unemployment Rate	764	2.29	0.46	1.06	3.36
Log Firm Density	734	0.08	0.01	0.06	0.16
Log Workforce	764	0.35	0.04	0.24	0.59
Log Distance	770	3.47	1.69	0.00	5.06
Log Population Density	770	5.15	0.80	3.46	7.88
HH index	770	0.23	0.02	0.19	0.29
Startup	776	7.42	20.33	0	320

***Figure 2**. Biplot of PCA results. Elements are provinces, their shape and colour signal macro-areas, their size the extent of their contribution to the PCA dimension.*

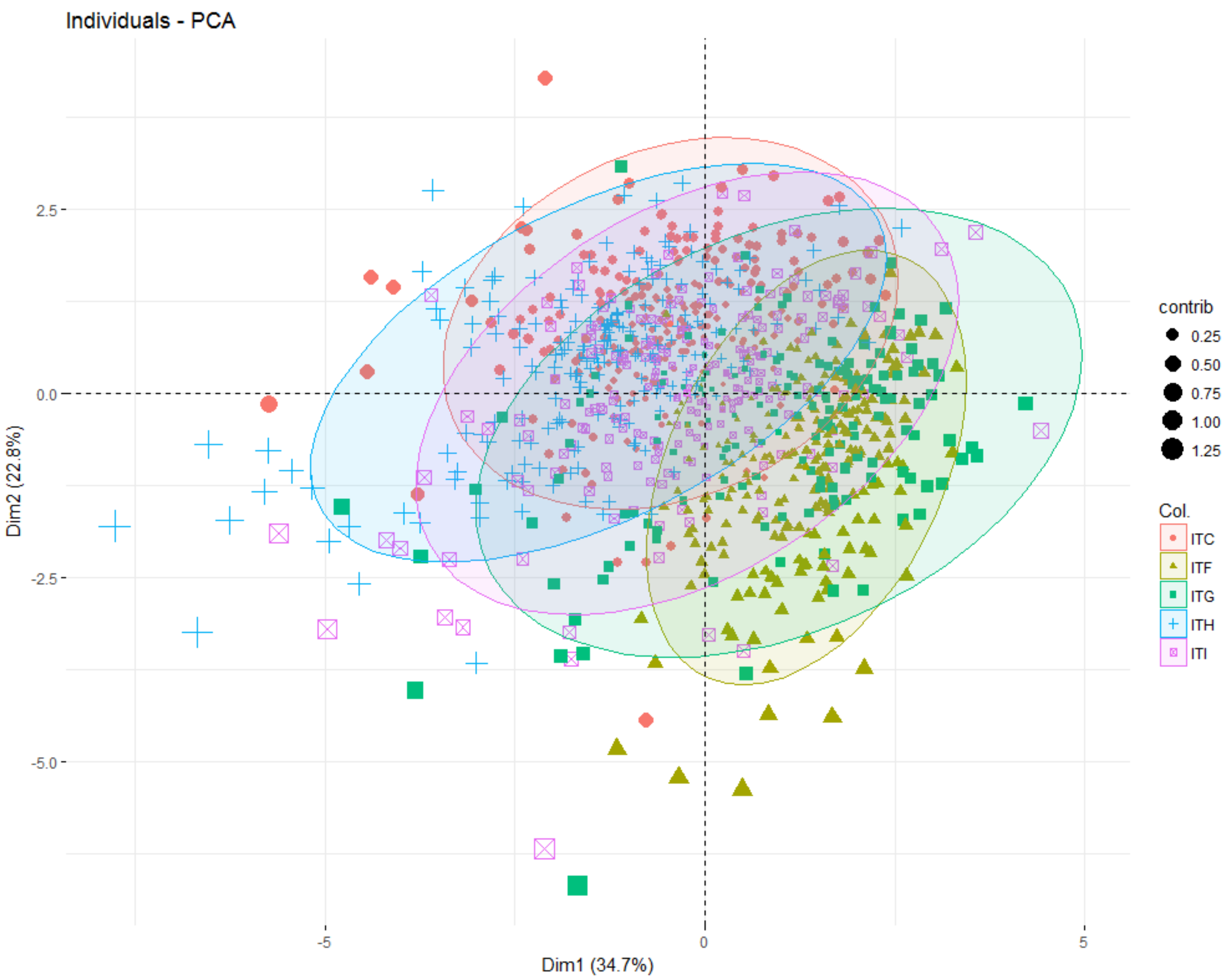
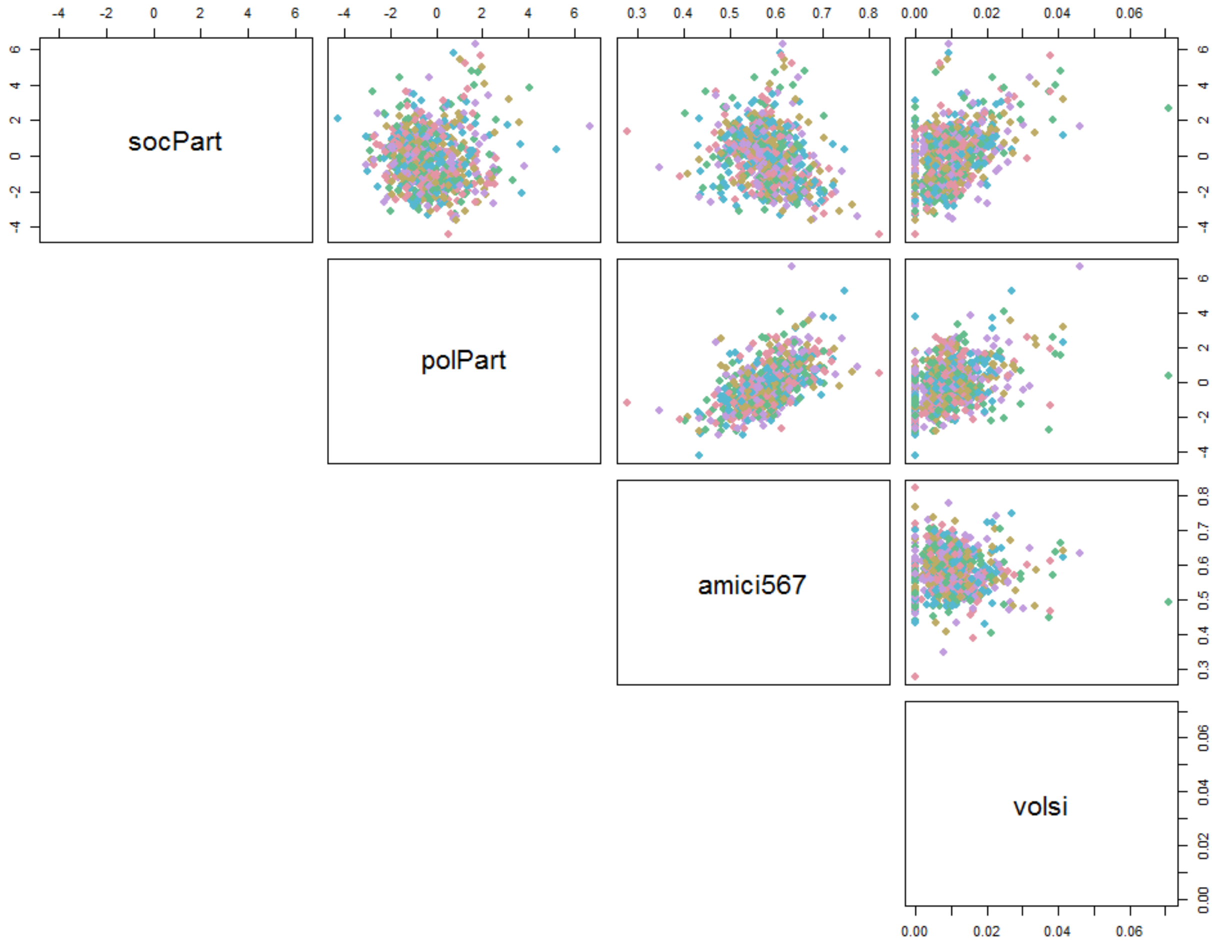


***Figure 3**. Pairs Plot of SC dimensions.*



5.6 The econometric model

As spelled out in the theoretical section, we look for a test of the "facilitating" effect SC should exert on the appropriation chances of knowledge factors by potential entrepreneurs. In econometric terms, we test for a moderating effect of SC dimensions respect to the knowledge related variables.

The baseline specification of the econometric model for investigating the effect of SC on innovative start-ups is the following:

$$Startup_{p,t} = SocPart_{p,t-1} + PolPart_{p,t-1} + Amici_{p,t-1} + Volsi_{p,t-1} + Controls_{p,t-1} + \delta_r + \delta_t + \varepsilon_{p,t}$$

Where *p* is the NUTS3 province, *t* is time, *r* the NUTS2 region and the δs are dummies. *Controls* is the vector of controls described above, meant to address the local characteristics of the territory under analysis. Explanatory variables enter the regressions at 1-year lag because of two reasons: i) in order to address potential issues of reverse causality, even though the effect of a rising rate of start-ups creation in a territory on the societal structure measured as volunteering activities and amical habits, if any, should be reasonably low; ii) because the sociological literature suggests that SC is a resource slow to build, hence it might need time to exert its effect. However, our theoretical reasoning points us to give more credit to an indirect, moderating effect of the SC dimensions. The final model specification that will be tested hence is the following:

$$Startup_{\ p,t} = KR_{\ p,t-1} * SC_{\ p,t-1} + Controls_{\ p,t-1} + \ \delta_{\ r} + \ \delta_{\ t} + \ \varepsilon_{p,t}$$

Where *KF* is a matrix of Knowledge-related Resources variables (*Stock*, *HC* and *Incubator*) and *SC* is a matrix of Social Capital dimensions (*SocPart*, *PolPart*, *Amici* and *Volsi*).

This equation can be estimated using the Negative Binomial (NB) estimator. Indeed, the discrete and non-negative nature of the dependent variable suggests the adoption of estimation techniques for "count data" models (Hausman, Hall and Griliches, 1984). Out of these models, the Poisson regression assumes that the dependent variable follows a Poisson distribution. Furthermore, it assumes the equality between conditional variance and conditional mean in the distribution of the dependent variable. When this condition is not met, as it is in the present case, the NB class of models is used, which permits over-dispersion. Therefore, the estimation model takes the following form.

$$E[(Startup_{\ p,t}\ |X_{p,t}, \varepsilon_{p,t})] = exp(X_{p,t-1} + C_{p,t-1} + \ \delta_{\ r} + \ \delta_{\ t} + \ \varepsilon_{p,t})$$

Where $X_{i,t-1}$ is the vector of our variables of interest. This model is formally estimated using Maximum Likelihood.

NB regression do not perform well under excess of zeros in the dependent variable (Gurmu and Trivedi, 1996; Dalrymple, Hudson and Ford, 2003). Within the category of "count data regression models", there are two alternative options: Zero-Inflated (ZI) models and Hurdle models. The crucial difference between the two estimation methodologies is that ZI models suppose two different but related data generation processes for the zeros, as in a two-step decision procedure (e.g. when you are allowed to not buy anything after your decision to buy something). Hurdle models, instead, fetch separately the two processes, claiming no assumptions on the generation of zeros. Indeed, Hurdle models are a combination of a binary regression (logit or probit) – in order to asses whether the count is 0 or non-zero – and, if the hurdle is crossed, i.e. the outcome is strictly positive, a truncated at-zero count regression (Poisson or NB). Since we do not have specific motivations to suggest the existence of a sub-process determining zero in our dependent variable, we prefer to exploit the potential of Hurdle models as robustness check. Unfortunately, the ML algorithm hade issues in convergence when interactions were plugged in, hence the main results will be shown as the output of a NB regression. However, robustness tests show no relevant differences between the NB regression without interactions and the Hurdle regression on the same model.

6. Results and discussion

The results of the regression analysis are shown in Table 4. Separate regressions have been run for each set of interactions between SC dimensions (*socPart*, *polPart*, *amici*, *volsi*) and one technological variable at a time (*Patent Stock*, *Incubator Stock* and *HC*, in column 1, 2 and 3 respectively) in order to avoid multicollinearity (checked with a VIF smaller than 10). The interaction coefficients and p-values indicate that only *amici567* and *polPart* exerts a significant moderation respect to every knowledge resource. However, *amici567* exerts a positive effect, whereas *polPart* is negative. Instead, *socPart* and *volsi* are not significant in any of the three regression sets.

Since interpretation of standalone estimates of interacted variables is sometimes cumbersome, Table 5 reports regression outputs of models without interactions, thus investigating the direct effect of each variable of interest on start-up creation. In these models, *Patent Stock*, *HC* and *Firm Density* are significant and positive, whereas *Distance* is significant and negative. It is important to note that columns 3 and 4 do not distance substantially, suggesting the reliability of NB estimates.

6.1 Discussion

Table 5 partially confirms literature's results. *Patent Stock* and *HC* are positive and significant, indicating that nascent innovative start-ups are reactive to both the technological knowledge capacity of a province and its skills endowment (Davidsson and Honig, 2003; Schenkel et al., 2012; Colombelli, 2016). Technology incubators, instead, are not a significant resource for provincial patterns of start-up creation in our sample, contrary to other evidence in the literature (Colombo and Delmastro, 2002). The only other two variables consistently significant in Table 4 and 5 are

Distance and *Firm Density*, signalling the existence of agglomeration effects stemming from pecuniary externalities (Scitovsky, 1954; Gehringer, 2011). *GVA pc* is slightly significant only in the moderation analysis. The empirical evidence seems to suggest that knowledge-related resources are of primary importance over economic factors at large for innovative start-ups creation.

Commenting specifically Table 4, our theoretical conjecture of SC as a knowledge "facilitator" is partially confirmed and partially dismissed. Indeed, SC emerges as a moderator on the effect of knowledge-related resources respect to innovative start-up provincial patterns. Breaking down the SC measure along various measurement dimensions proves a good choice since it let us appreciate some heterogeneity within the sphere of structural/relational SC. Within this sphere, the intensity of friendly relationships appears as an effective conduit for knowledge-related resources, amplifying their final effect on start-up creation. This result is in line with the centrality of individual networks in a variety of socio-economically relevant activities (Granovetter, 1974; Ioannides and Loury, 2004; Jackson, 2014). We expected social participation to be significant as well. As outlined in the theoretical section, social participation should generate SC in that it enhances openness hence the potential for new ties and exchanges to settle. However, in our dataset it is not relevant, which may due to two reasons. Either the variables building our *socPart* PC do not really capture the territorial openness, or the kind of real or potential ties developing from social participation are not suitable for a specific endowment like knowledge-related resources, or information about these resources, to flow.

The negative and significant effect of political participation was not expected but not totally surprising. The "bad" side of SC is customary in the literature, which underlines that SC is not "good" or "bad" per se, rather it differentiates its effect depending on the context of application (Westlund and Bolton, 2003). In our specific context, there are two possible explanations. Higher political participation may measures citizens' malcontent for their social conditions, hence the inadequacy of territorial social infrastructures to sustain private initiatives in uncertain ventures. Another rationale is that it measures a bonding type of SC that induces groups closure via identity-reinforcement mechanisms, whereas resources circulation is a prerequisite for innovative start-up propensity (Putnam, 2000; Tura and Harmaakorpi, 2005; Kwon and Adler, 2014). In other words, the participation to particularistic associations may stretch the SC measurement toward the relational dimension, where SC qualifies more as a club good, whereas active social participation at large and friendly relationships may go in the direction of the structural dimension (see Table 1 in Section 3).

***Table 4**. Moderating effects investigation. Negative Binomial Regressions with SC-Technological resources interactions by Technological resources (Centred). Every explanatory variable is lagged 1 year. Clustered SE at NUTS3. Year and NUTS2 FE. SC variables not shown for sake of space.*

	Dependent variable:		
	startup		
	(1)	(2)	(3)
Log Patent Stock	0.688*** (0.062)	0.673*** (0.063)	0.692*** (0.069)
Log Incubator Stock	0.150 (0.127)	0.117 (0.198)	0.107 (0.124)
Log HC	0.417*** (0.098)	0.385*** (0.101)	0.417*** (0.132)
Log GVA per capita	0.325* (0.178)	0.363* (0.187)	0.353* (0.191)
Log Unemployment Rate	0.132 (0.189)	0.112 (0.193)	0.121 (0.191)
Log Firm Density	0.127*** (0.048)	0.124** (0.051)	0.117** (0.050)
Log Workforce	-0.034* (0.020)	-0.030 (0.021)	-0.035 (0.022)
Log Distance	-0.067** (0.027)	-0.062** (0.028)	-0.065** (0.027)
Log Population Density	0.115 (0.096)	0.110 (0.103)	0.102 (0.106)
HH index	-2.505 (3.425)	-0.568 (3.434)	-1.832 (3.585)
socPart * Patent Stock	0.018 (0.020)		
polPart * Patent Stock	-0.061*** (0.022)		
amici567 * Patent Stock	0.104*** (0.033)		
volsi * Patent Stock	0.042 (0.028)		
socPart * Incubator Stock		0.069 (0.068)	
polPart * Incubator Stock		-0.184* (0.109)	
amici567 * Incubator Stock		0.250** (0.103)	
volsi * Incubator Stock		0.123 (0.087)	
socPart * HC			-0.017 (0.076)
polPart * HC			-0.192** (0.085)
amici567 * HC			0.258** (0.114)
volsi * HC			0.130 (0.083)
Observations	717	717	717
Log Likelihood	-1,457.696	-1,460.845	-1,460.541
theta	7.210*** (1.038)	6.854*** (0.959)	6.892*** (0.968)
Akaike Inf. Crit.	2,997.391	3,003.691	3,003.081
Note:			*p<0.1; **p<0.05; ***p<0.01

***Table 5**. Negative Binomial and Hurdle regressions (Year and NUTS2 FE, Clustered SE at NUTS3) with only knowledge-related and control variables. Every variable is lagged 1 year.*

	Dependent variable:		
	startup		
	negative binomial		*hurdle*
	(1)	(2)	(3)
Log Patent Stock	0.855*** (0.067)	0.678*** (0.065)	0.647*** (0.055)
Log Incubator Stock	0.375*** (0.124)	0.133 (0.133)	0.164 (0.110)
Log HC	0.543*** (0.130)	0.389*** (0.102)	0.378*** (0.079)
Log GVA per capita		0.284 (0.190)	0.212 (0.161)
Log Unemployment Rate		0.098 (0.187)	-0.117 (0.168)
Log Firm Density		0.122** (0.050)	0.148*** (0.047)
Log Workforce		-0.031 (0.022)	-0.035* (0.019)
Log Distance		-0.065** (0.027)	-0.070*** (0.025)
Log Population Density		0.121 (0.098)	0.113 (0.084)
HH index		-1.242 (3.210)	-3.792 (2.775)
Observations	740	717	717
Log Likelihood	-1,513.954	-1,463.507	-1,411.020
theta	4.828*** (0.590)	6.611*** (0.906)	
Akaike Inf. Crit.	3,095.908	3,009.014	
Note:			*p<0.1; **p<0.05; ***p<0.01

7. Conclusions

The interest upon innovative new small firms has been rising in the last decade since they have been recognized as fundamental carriers of novelties. A fundamental step to assess the role of new entrepreneurs in economic growth is the KSTE, grafting the economics of knowledge externalities onto entrepreneurship studies. With the further implementation of the knowledge filter and entrepreneurial absorptive knowledge capacity theories, we now have a vast toolkit to appreciate, and eventually understand, the dynamics governing territorial knowledge exploitation by enterprising individual and the intermediating role of institutions in knowledge dissemination.

With the present study, we enlarged further the understanding of territorial innovative entrepreneurial dynamics, enriching the KSTE from the contextual, societal perspective. SC theory offers a theoretically developed and fertile background to track the effect of interactions between individuals at large in a society on the ability and propensity towards entrepreneurship. Our results suggest that, indeed, knowledge-related resources diffusion and appropriation is influenced by the kind of social structure the territory is made of. The investigation of the channels through which information and knowledge externalities diffuse is important once the collective nature of technological knowledge creation, which lays at the base of innovative entrepreneurship, is appreciated (Antonelli, 2000). SC theory let us investigate in depth the heterogeneous characteristics of the collective channels for relevant resources mobilization.

More investigation is needed to better qualify the interaction between the dimensions of the social structure and the relevant resources for innovative start-ups to proliferate, hence spurring economic growth. Future research may augment the present analysis with compositional indexes of the technological stock, or investigating the effect of external economic shocks on the capability of the social structure to mediate the relevant knowledge-related resources. In any way, a deeper and richer understanding of the knowledge and informational externalities dynamics of diffusion, appropriation and transformation in economic value will make only good for the economics of innovation.

References

Acs, Z.J. and Audretsch, D.B. (1989) Small-Firm Entry in US Manufacturing. *Economica*, **56**, 255.

Acs, Z.J., Audretsch, D.B., Braunerhjelm, P. and Carlsson, B. (2004) The Missing Link: The Knowledge Filter and Entrepreneurship in Endogenous Growth. *CEPR Discussion Papers 4783*.

Acs, Z.J., Braunerhjelm, P., Audretsch, D.B. and Carlsson, B. (2009) The knowledge spillover theory of entrepreneurship. *Small Business Economics*, **32**, 15–30.

Acs, Z.J. and Plummer, L.A. (2005) Penetrating the ``knowledge filter'' in regional economies. *The Annals of Regional Science*, **39**, 439–456.

Aghion, P. and Howitt, P. (1992) A Model of Growth Through Creative Destruction. *Econometrica*, **60**, 323.

Akerlof, G. (1970) The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, **84**, 488–500.

Antoci, A., Sabatini, F. and Sodini, M. (2013) Economic Growth, Technological Progress and Social Capital: The Inverted U Hypothesis: Economic Growth, Technological Progress and Social Capital. *Metroeconomica*, **64**, 401–431.

Antonelli, C. (2000) Collective Knowledge Communication and Innovation: The Evidence of Technological Districts. *Regional Studies*, **34**, 535–547.

Arrow, K. (1962) Economic welfare and the allocation of resources for invention. *The rate and direction of inventive activity: Economic and social factors* pp. 609–626. Princeton University Press.

Audretsch, D.B. (1995) *Innovation and Industry Evolution*. MIT Press, Cambridge (Mass).

Audretsch, D.B. (2007) Entrepreneurship capital and economic growth. *Oxford Review of Economic Policy*, **23**, 63–78.

Audretsch, D.B., Falck, O., Feldman, M.P. and Heblich, S. (2012) Local Entrepreneurship in Context. *Regional Studies*, **46**, 379–389.

Audretsch, D.B., Keilbach, M.C. and Lehemann, E.E. (2006) *Entrepreneurship and Economic Growth*. Oxford University Press, Oxford.

Audretsch, D., B., Lehemann, E.E. and Hinger, J. (2015) From knowledge to innovation: the role of knowledge spillover entrepreneurship. *Routledge Handbook of the Economics of Knowledge*, Routledge p. Antonelli C. and Link A. N., New York.

Audretsch, D.B. and Lehmann, E.E. (2005) Does the Knowledge Spillover Theory of Entrepreneurship hold for regions? *Research Policy*, **34**, 1191–1202.

Baptista, R. and Mendonça, J. (2010) Proximity to knowledge sources and the location of knowledge-based start-ups. *The Annals of Regional Science*, **45**, 5–29.

Bonaccorsi, A., Colombo, M.G., Guerini, M. and Rossi-Lamastra, C. (2013) University specialization and new firm creation across industries. *Small Business Economics*, **41**, 837–863.

Bourdieu, P. (1986) The forms of Capital. *Handbook of Theory and Research for the Sociology of Education* pp. 241–258. New York: Greenwood Press.

Braunerhjelm, P., Acs, Z.J., Audretsch, D.B. and Carlsson, B. (2010) The missing link: knowledge diffusion and entrepreneurship in endogenous growth. *Small Business Economics*, **34**, 105–125.

Breschi, S. and Lissoni, F. (2001) Knowledge Spillovers and Local Innovation Systems: A Critical Survey. *Industrial and Corporate Change*, **10**, 975–1005.

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

Chiesi, A.M. (2007) Measuring Social Capital and its Effectiveness. The Case of Small Entrepreneurs in Italy. *European Sociological Review*, **23**, 437–453.

Coleman, J.S. (1990) *Foundations of Social Theory*, 3. print. Belknap Press of Harvard Univ. Press, Cambridge, Mass.

Colombelli, A. (2016) The impact of local knowledge bases on the creation of innovative start-ups in Italy. *Small Business Economics*, **47**, 383–396.

Colombelli, A. and Quatraro, F. (2018) New firm formation and regional knowledge production modes: Italian evidence. *Research Policy*, **47**, 139–157.

Colombo, M.G. and Delmastro, M. (2002) How effective are technology incubators? *Research Policy*, **31**, 1103–1122.

Crescenzi, R., Gagliardi, L. and Percoco, M. (2013) Social Capital and the Innovative Performance of Italian Provinces. *Environment and Planning A*, **45**, 908–929.

Dalrymple, M., Hudson, I. and Ford, R. (2003) Finite Mixture, Zero-inflated Poisson and Hurdle models with application to SIDS. *Computational Statistics & Data Analysis*, **41**, 491–504.

Davidsson, P. and Honig, B. (2003) The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing*, **18**, 301–331.

Durlauf, S.N. and Fafchampes, M. (2004) Social Capital. *CSAE WPS/2004-14*.

Ferrary, M. and Granovetter, M. (2009) The role of venture capital firms in Silicon Valley's complex innovation network. *Economy and Society*, **38**, 326–359.

Fukuyama, F. (1996) *Trust: The Social Virtues and the Creation of Prosperity*, 1. Free Press paperback ed. Free Press, New York, NY.

Gehringer, A. (2011) Pecuniary knowledge externalities and innovation: intersectoral linkages and their effects beyond technological spillovers. *Economics of Innovation and New Technology*, **20**, 495–515.

Geroski, P.A. (1995) What do we know about entry? *International Journal of Industrial Organization*, **13**, 421–440.

Granovetter, M.S. (1974) *Getting a Job: A Study of Contacts and Careers*, 2nd ed. University of Chicago Press, Chicago.

Griliches, Z. (1992) The search for R&D spillovers. *Scandinavian Journal of Economics*, 29–47.

Grootaert, C. and van Bastelaer, T. (2001) Understanding and measuring Social Capital: a synthesis of findings and recommendations from the Social Capital Initiative. *Social Capital Initiative Working Paper No. 24*.

Grootaert, C., Van Bastelaer, T. and World Bank (eds). (2002) *Understanding and Measuring Social Capital: A Multidisciplinary Tool for Practitioners*. World Bank, Washington, D.C.

Gui, B. and Sugden, R. (2010) *Economics and Social Interaction: Accounting for Interpersonal Relations*. Cambridge University Press, Cambridge.

Guiso, L., Sapienza, P. and Zingales, L. (2010) Civic Capital as the Missing Link. *SSRN Electronic Journal*.

Gurmu, S. and Trivedi, P. (1996) Excess Zeros in Count Models for Recreational Trips. *Journal of Business & Economic Statistics*, **14**, 469–77.

Hall, B.H., Jaffe, A.B. and Trajtenberg, M. (2005) Market value and patent citations. *The RAND Journal of Economics*, **36**, 16–38.

Hardin, R. (1999) Social Capital. *Competition and Cooperation: Conversations with Nobelists about Economics and Political Science* pp. 170–189. Russell Sage Foundation, New York.

Hausman, J., Hall, B.H. and Griliches, Z. (1984) Econometric Models for Count Data with an Application to the Patents-R & D Relationship. *Econometrica*, **52**, 909.

Ioannides, Y.M. and Loury, L.D. (2004) Job information networks, neighborhood effects, and inequality. *Journal of economic literature*, **42**, 1056–1093.

Jackson, M.O. (2014) Networks in the Understanding of Economic Behaviors. *Journal of Economic Perspectives*, **28**, 3–22.

Kwon, S.-W. and Adler, P.S. (2014) Social Capital: Maturation of a Field of Research. *Academy of Management Review*, **39**, 412–422.

Liao, J. and Welsch, H. (2005) Roles of social capital in venture creation: Key dimensions and research implications. *Journal of small business management*, **43**, 345–362.

Lin, N. (2002) *Social Capital: A Theory of Social Structure and Action*, 1. paperback ed. Cambridge Univ. Press, Cambridge.

Loury, G. (1976) A dynamic theory of racial income differences. *Discussion Paper No. 225, Northwestern University*.

Loury, G.C. (1987) Why Should We Care about Group Inequality? *Social Philosophy and Policy*, **5**, 249.

Lucas, R.E. (1988) On the mechanics of economic development. *Journal of Monetary Economics*, **22**, 3–42.

Nahapiet, J. and Ghoshal, S. (1998) Social Capital, Intellectual Capital, and the Organizational Advantage. *The Academy of Management Review*, **23**, 242.

Nooteboom, B. (2006) Social capital, institutions and trust. *CentER Discussion Paper, 2006-35*.

North, D.C. (1990) *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, Cambridge ; New York.

Owen-Smith, J. and Powell, W.W. (2004) Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science*, **15**, 5–21.

Portes, A. (1998) Social Capital: Its Origins and Applications in Modern Sociology. *Annual Review of Sociology*, **24**, 1–24.

Putnam, R.D. (2000) Bowling Alone: America's Declining Social Capital. *Culture and Politics* (eds L. Crothers), & C. Lockhart), pp. 223–234. Palgrave Macmillan US, New York.

Putnam, R.D., Leonardi, R. and Nanetti, R.Y. (1994) *Making Democracy Work: Civic Traditions in Modern Italy*, 5. print., 1. Princeton paperback print. Princeton Univ. Press, Princeton, NJ.

Qian, H. and Acs, Z.J. (2013) An absorptive capacity theory of knowledge spillover entrepreneurship. *Small Business Economics*, **40**, 185–197.

Qian, H., Acs, Z.J. and Stough, R.R. (2012) Regional systems of entrepreneurship: the nexus of human capital, knowledge and new firm formation. *Journal of Economic Geography*, **13**, 559–587.

Qian, H. and Jung, H. (2017) Solving the knowledge filter puzzle: absorptive capacity, entrepreneurship and regional development. *Small Business Economics*, **48**, 99–114.

Righi, A. (2013) Measuring Social Capital: Official Statistics Initiatives in Italy. *Procedia - Social and Behavioral Sciences*, **72**, 4–22.

Romer, P. (1990) Endogenous technological change. *Journal of Political Economy*, S71-102.

Sabatini, F. (2009) Social capital as social networks: A new framework for measurement and an empirical analysis of its determinants and consequences. *The Journal of Socio-Economics*, **38**, 429–442.

Schenkel, M.T., D'Souza, R.R. and Matthews, C.H. (2012) Entrepreneurial capital: examining linkages in Human and Social Capital of new ventures. *Journal of Developmental Entrepreneurship*, **17**, 1250009.

Scitovsky, T. (1954) Two Concepts of External Economies. *Journal of Political Economy*, **62**, 143–151.

Shane, S. and Venkataraman, S. (2000) The Promise of Enterpreneurship as a Field of Research. *The Academy of Management Review*, **25**, 217.

Solow, R. (2000) Notes on social capital and economic performance (eds P Dasgupta and I Serageldin). *Social Capital. A Multifaceted perspective*.

Stiglitz, J.E. (2002) Information and the Change in the Paradigm in Economics. *The American Economic Review*, **92**, 460–501.

Stuart, T.E. and Sorenson, O. (2005) Social networks and entrepreneurship. *Handbook of entrepreneurship research*, 233–252.

Tosini, D. (2005) Capitale sociale: problemi di costruzione di una teoria. *Quaderno 32 del Diparimento di Sociologia e Ricerca Sociale, Università degli Studi di Trento*.

Trigilia, C. (2001) *Il Capitale Sociale*. Il Mulino.

Tura, T. and Harmaakorpi, V. (2005) Social capital in building regional innovative capability. *Regional Studies*, **39**, 1111–1125.

Venkataraman, S. (1997) The distinctive domain of entrepreneurship research: An editor's perspective. *Advances in Entrepreneurship, Firm Emergence, and Growth*, **3**, 119–138.

Weitzman, M.L. (1998) Recombinant growth. *The Quarterly Journal of Economics*, **113**, 331–360.

Westlund, H. and Bolton, R. (2003) Local social capital and entrepreneurship. *Small business economics*, **21**, 77–113.

Appendix

Figure 4. *Plot of the contribution of each raw variable to the PCA dimensions.*

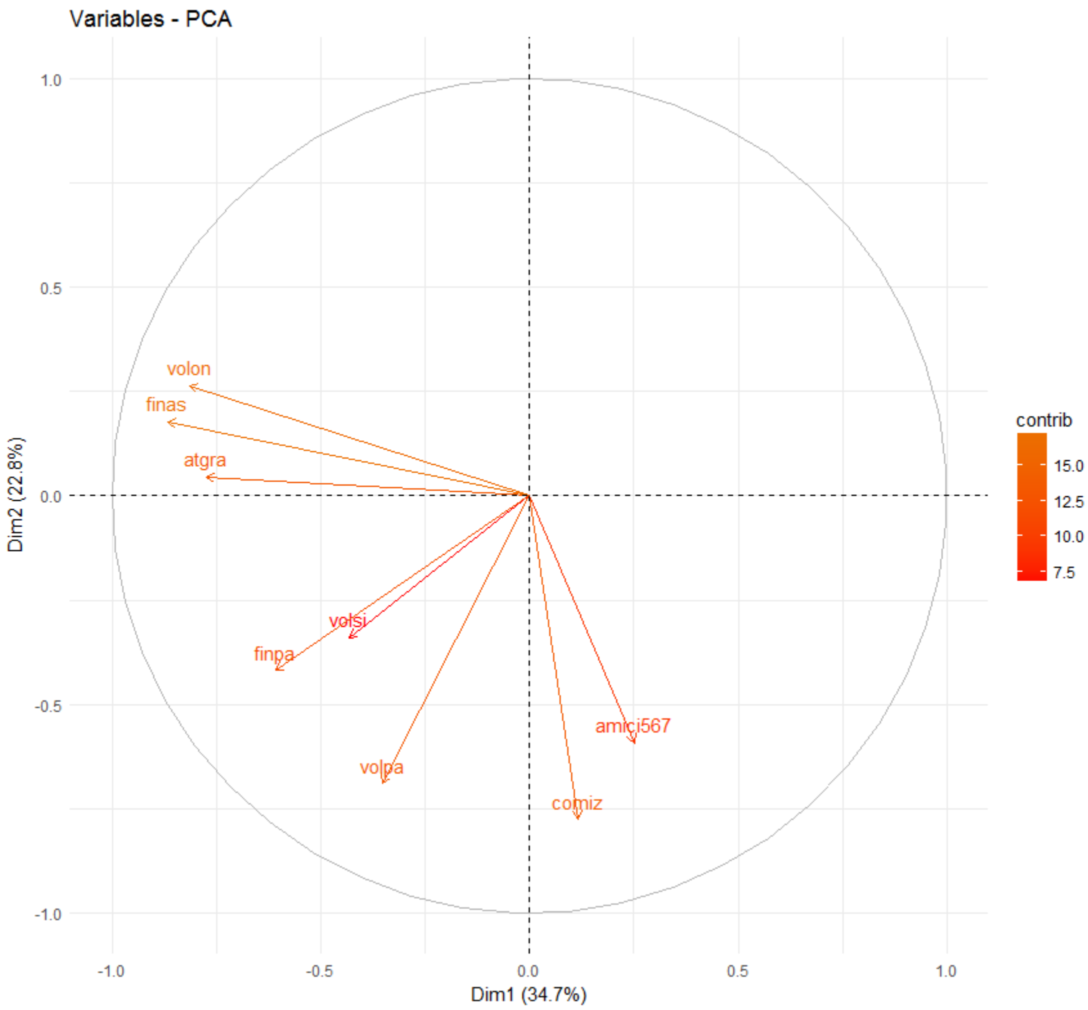


Figure 5*. Quality of the representation of each raw variable respect to each SC dimension (cos2).*

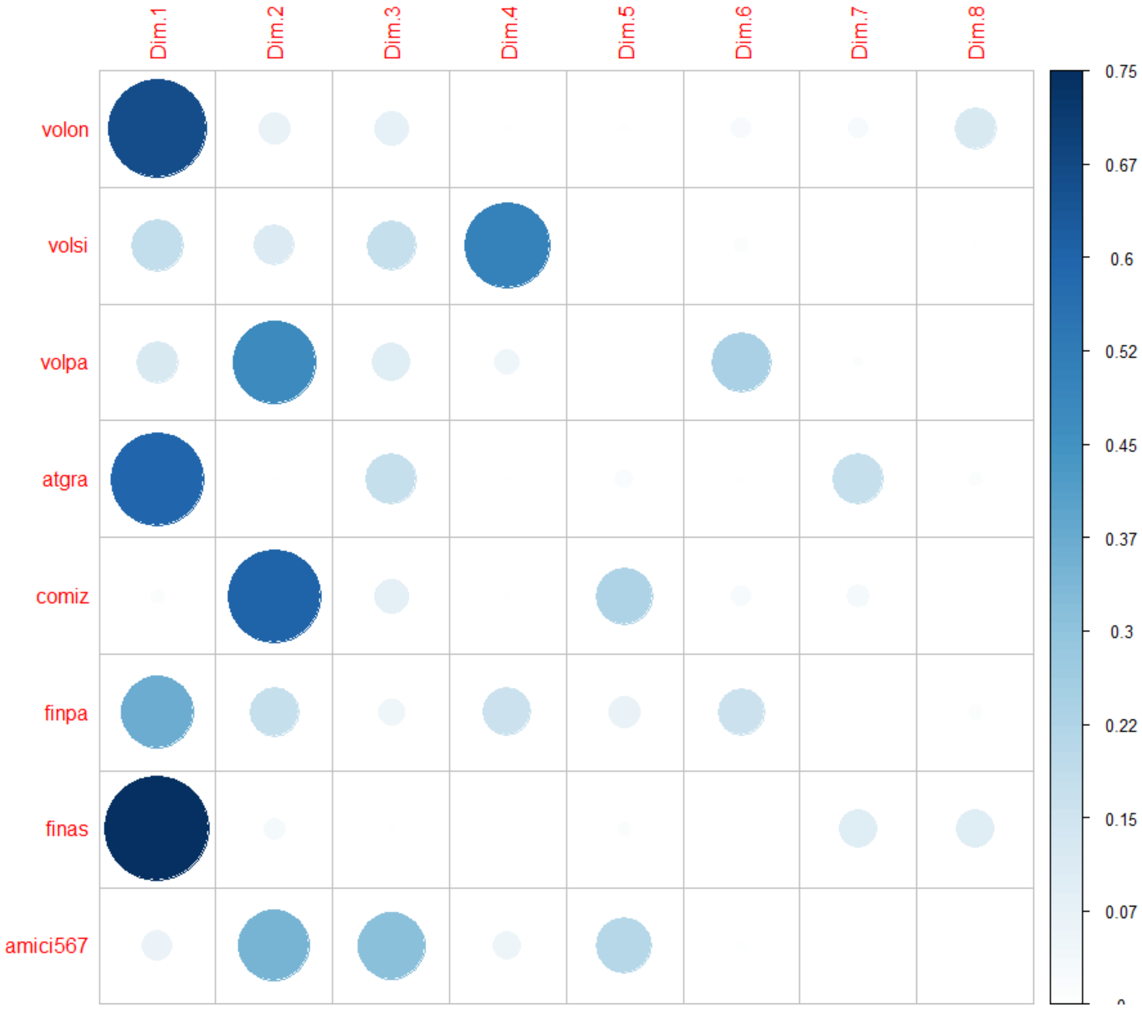


Figure 6. *Scree Plot of PCA.*

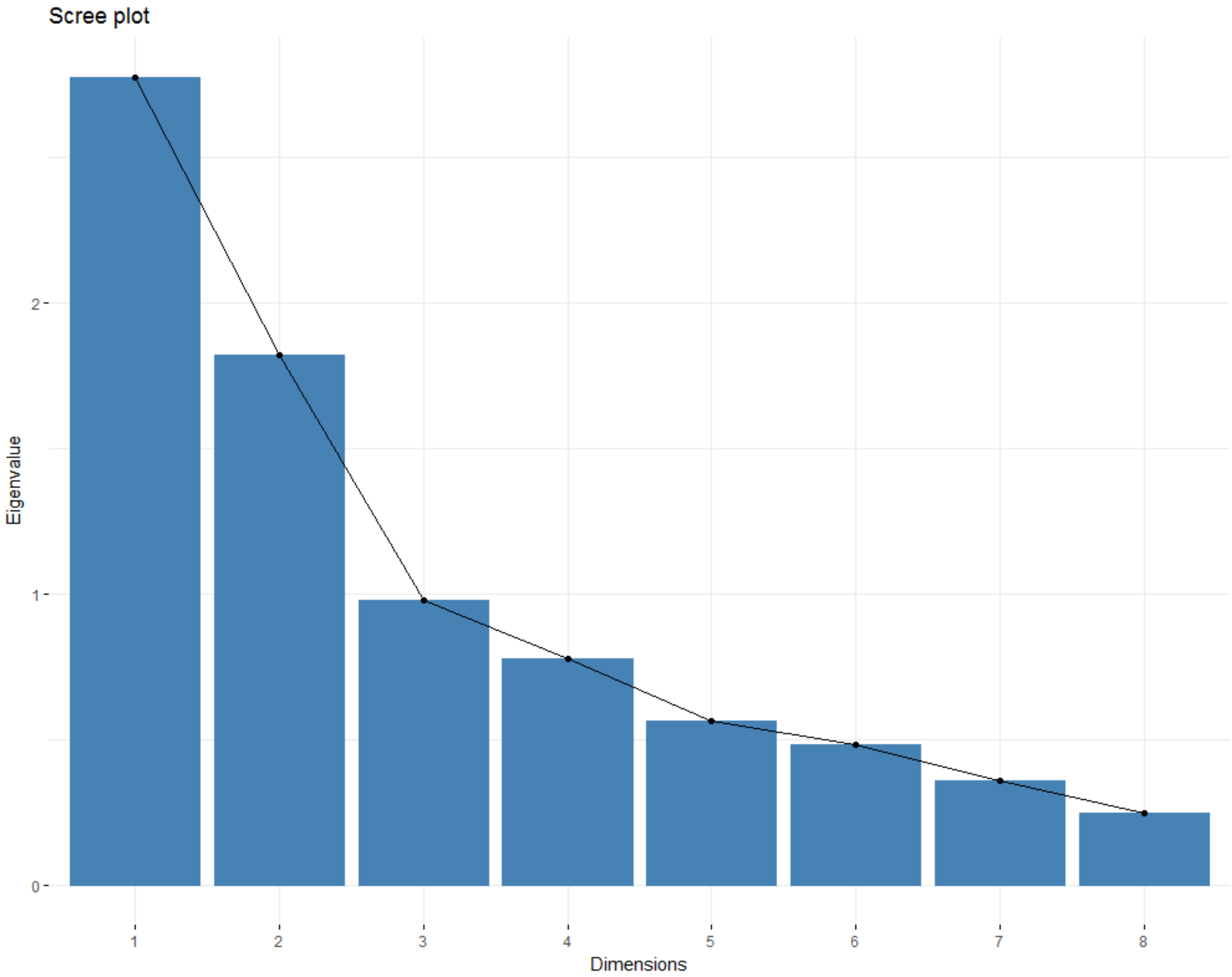


***Figure 7**. Correlation Matrix.*

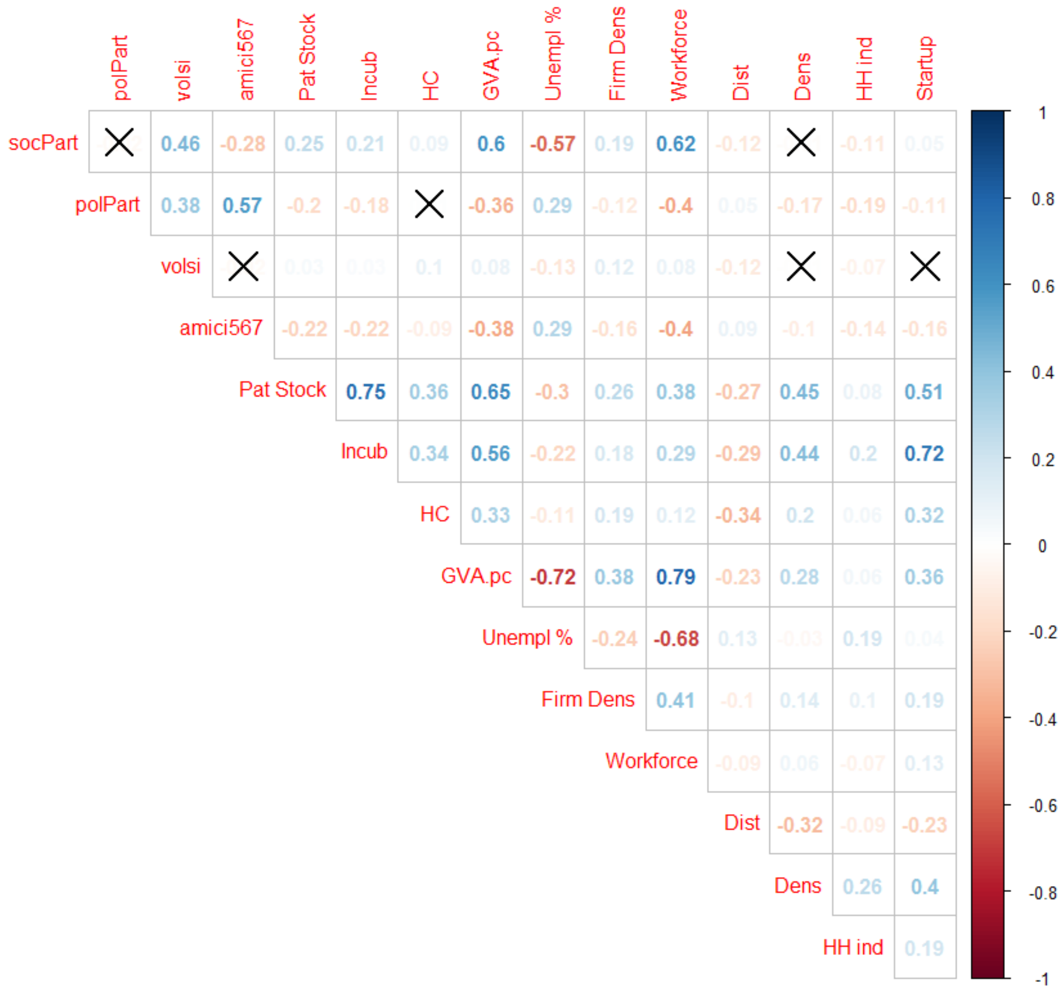


***Figure 8**. QQ-plot of the start-ups count by province from 2009 to 2015.*