PAPER (do not quote without permission)

Title: The inventor dynamics of knowledge whispers: A multilayer analysis

Authors and e-mails of them:
Matteo Tubiana[a,b], Ernest Miguélez[c,d] , Rosina Moreno[d]

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

[b] *INAPP, Roma*

[c] GREThA UMR CNRS 5113—Université de Bordeaux.

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

rmoreno@ub.edu

Subject area: La geografía de la innovación y el desbordamiento del conocimiento

Abstract:
The present paper studies the role of knowledge accessibility on individuals' productivity, by exploiting a large dataset of EPO inventors over time and across European cities. The paper simultaneously investigates the role of different layers with which inventors interact and from where they learn and source ideas: the city/metropolitan areas where they resides, the firm/organization for which they work, and their network of collaborators. We take therefore a multilevel approach to the phenomenon. In doing so, we are bridging 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. City-level knowledge diffusion is stronger than firm-level knowledge and network-level one. However, when the type of knowledge accessed and produced is taken into account, network effects prevail, suggesting that complex knowledge diffuses only within strong social connections, while simple one diffuses equally across actors within cities.

Keywords: *knowledge diffusion, firms, networks, regions, patents*
JEL codes: O18, O31, O33, R11

1. Introduction

Knowledge diffusion plays a central role in many disciplines in social sciences. Most of them reach the conclusion that the production of new knowledge arises from the combination and recombination of existing ideas ((Weitzman, 1998; Fleming and Sorenson, 2001). The increasing “burden of knowledge" leads individuals to specialize and become more dependent on other people's skills to innovate (Jones, 2009). Meanwhile, teamwork is seen as a way to realize scale economies in knowledge production and share risks (Powell and Grodal, 2005; Wutchy, Jones and Uzzi, 2007; Singh and Fleming, 2010; Jaravel, Petkova and Bell, 2018). Even children who are more exposed to innovation (e.g., growing up in areas with more inventors) are more likely to become inventors themselves (Bell et al., 2017). Hence, scholars have clearly established that inventors’ productivity depends both on their own knowledge and human capital, as well as on interacting with other inventors and learning from them (Lucas, 2009; Lucas and Moll, 2014; Akcigit et al., 2018).

However, the way in which innovators access knowledge to be recombined has been way less investigated. Economic geography, regional science and the economics of innovation have widely investigated knowledge diffusion and recombination at the regional/local level (Boschma, 2005; Antonelli, Patrucco and Quatraro, 2011). The concept of knowledge externalities has become commonplace in the field to explain gains from agglomeration economies and industry concentration. The role they 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). 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 concepts such as localized knowledge spillovers (Jaffe, Trajtenberg and Henderson, 1993) or Regional Innovation Systems (Asheim and Coenen, 2005, 2006). Agglomerations generate opportunities for repeated face-to-face contacts and exchanges of knowledge and ideas (Amin and Cohendet, 2005). 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 simply because they are co-located in space, in a bond of institutions, supply chains and repetitive transactions and interactions (Audretsch et al., 2012; Audretsch, Lehemann and Hinger, 2015). In this context, cities are seen as platforms to reduce the costs of interacting and allow accessing spatially sticky knowledge (Storper and Venables, 2004; Carlino and Kerr, 2014). As a matter of fact, cities have been traditionally considered focal points of creative activity and further innovation (Jacobs, 1969, p. 196; Feldman and Audretsch, 1999; Duranton and Puga, 2001).

More recently, researchers have questioned the idea that knowledge is "in the air" (Marshall, 1890) in cities and clusters. By contrast, knowledge diffusion is the result of planned and well-structured partnership between individuals, firms and other organizations (Fitjar and Rodríguez-Pose, 2017), thus highlighting the importance of social networks' membership in order to access relevant pieces of knowledge at the local level (Uzzi, 1997, p. 199; Singh, 2005; Breschi and Lissoni, 2009) as well as their role to tap into non-local sources of ideas (Bathelt, Malmberg and Maskell, 2004). What matters most in this strand of studies is the transmission channel. Even though co-location and networks might be sometimes observationally equivalent, the latter are seen as independent from the role of the former, and sometimes it is found even more important in regulating knowledge flows between agents (Breschi and Lissoni, 2001, 2009).

In parallel to these developments, the knowledge-based view of the firm has explored its role in organizing and distributing knowledge and seeking for competitive advantage internally (Kogut and Zander, 1992). This literature defends that the mere existence of the firm facilitates the transferability of knowledge between individuals within firms' boundaries (Allen, 1977; Teece, 1977; Grant, 1996). This is especially the case of knowledge of tacit nature, whose transfer between people is slow, costly and uncertain (Kogut and Zander, 1992). As knowledge is generally not appropriable by means of market transaction, the firm serves as the best platform to organize and share knowledge among the different individuals. To put it differently, the existence of

the firm is a response to the need of coordinating efforts of individual specialists who possess many different types of knowledge, and need to exchange it among them to produce new ideas (Grant, 1996).

In this framework, the contributions of this paper are manifold. First, we assess the importance of accessing knowledge from three different layers: the city, the firm, and the individual's network. While we expect the three levels to be positive and significant, we are agnostic on which layer is going to dominate. As shortly reviewed above, extant research generally focuses on only one level of analysis while neglecting others, with the empirical consequences of ignoring unobserved heterogeneity coming from other layers. Moreover, focusing on one level and ignoring the others also come with the caveat of neglecting the possibility that the three levels of analysis may interact. In the present paper, we do not only assess the impact of each layer while controlling for the other two to avoid confounding effects, but we also gauge the possible complementary or substitutive relation between layers. Next, in our analysis we explicitly differentiate between quantity and quality – by means of forward citations to patents, a typical indicator for patent quality and value (Jaffe and de Rassenfosse, 2017). Finally, we exploit the characteristics of the knowledge accessed by differentiating between different degrees of complexity (Sorenson, Rivkin and Fleming, 2006; Balland and Rigby, 2017). We expect the advantages of accessing complex knowledge to grow the stronger are the connections between the sender and receiver of the message – that is, when they are socially connected, as opposed to when they only live in the same city. Accessing knowledge requires that the receiving partner makes efforts to understand it and acquire it, even correct potential errors in the message transmitted. These difficulties increase as the knowledge (the message) becomes more complex, and can only be overcome with short chains of transmission (within the individuals' networks or within her organization's boundaries). On the contrary, the less connected are the sender and the receiver in the social dimension, the larger the potential of simple knowledge to be accessed, which do not require the full assistance of the sender to be understood.

In addressing these research questions, several interesting novelties are introduced. 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 – except for the few papers mentioned above, such as Akcigit et al. (2018) and Jaravel, Petkova and Bell (2018). In particular, as a novelty in the literature, we explore directly the effect of the pool of local knowledge on individuals' performance. 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.

Our understanding of the factors driving and influencing inventors' production 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.

Second, we focus our territorial part of the analysis at the level of European cities or metropolitan areas[1], contrary to previous studies on the geography of innovation, which has focused on larger NUTS2 areas – some of them barely urbanized. There is substantial theory and evidence that innovation is primarily an urban phenomenon (Bairoch, 1988; Carlino and Kerr, 2014). By contrast, most empirical studies on the geography of innovation have used administrative boundaries such as NUTS2 regions

[1] Even though these are labelled metropolitan areas or metropolitan regions, for the sake of simplicity we will call them cities in the remaining of the paper, following a large tradition in economics on the relationship between cities, or urban areas, and innovation, starting from Jane Jacobs (Jacobs, 1961). It is well documented that innovation is, by large, an urban phenomenon (Bairoch, 1988).

in Europe or US States. This chosen spatial scale of analysis should reflect more closely the dynamics of knowledge interactions and innovation, and complements the existing empirical evidence at the level of regions.

Our empirical analysis uses an underexploited database of disambiguated EPO inventors residing in Europe, from 1978 to 2010 (Pezzoni, Lissoni and Tarasconi, 2014). Using patent data and the information on inventors, their city of residence, their collaborators, and the firms for which they work listed in patent documents, we build an unbalanced panel at the individual, inventor level regressing inventors' performance on several measures of knowledge stock at the level of the city, the firm, and the individual network of collaborators. Our setting allows us to incorporate a large list of fixed effects (city, time, sector, and individual), which allows us to rule out the influence of confounding factors.

We find evidence on the importance of city-level knowledge stocks on inventors' patent production. Doubling the city stocks increases individuals' productivity by 4-5%. These estimates are in line with (though somewhat larger than) elasticities on the effects of agglomeration economies on wages (Rice et al., 2006). We also find positive effects for the firm-level knowledge stocks, but less preponderant in terms of magnitude. Interestingly, when patent quality is taken into account, the picture changes upside down: what matters the most for quality-adjusted patent production is the network-level stock, while the city-level stock is not significant, suggesting that closer relations are critical to share knowledge which will allow individuals to produce high-quality ideas. Further, when exploring whether this is explained by the complexity of the message transmitted, we find some evidence supporting it (although relatively small): the effect of city-level knowledge stocks on patent production and quality-adjusted patent decrease as the level of complexity of the stocks grows. Meanwhile, the effect of network-level stocks on patent and quality-adjusted patent production increases with knowledge complexity – especially for the latter, suggesting that close and strong ties between the sender and the receiver of the message helps in transmitting complex ideas and in transforming them into high value innovations.

In the next Section, we review the literature presented above. Section 3 presents our methodological approach, while Section 4 describes the data building process and the final dataset. Finally, Section 5 outlines the preliminary results and conclusions 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 appreciate 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

determines 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 et al., 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 poor 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, huge 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 main 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. Methods

The effort of opening the 'black box' of knowledge externalities has been a major step forward in our understanding of how knowledge disseminates, stating clearly 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 (see section 4), 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, that is, the city where he lives. This layer is encompassing the other two – though not always, but comprehending also 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. Rather, it is an index of a knowledge potential embedded in the repositories (the network, the firm or the city).

We plan to test to what extent the stock of knowledge at the city 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 (Breschi and Lissoni, 2009). Instead, institutional and formal channels (labour-related mechanisms) could not be enough to account for all the creative knowledge potential of a city. We will investigate the direct impact of both the network and the institutional and territorial environments, with a 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).

3.1. 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 city. 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 common 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 (MA) (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), this approach has been preferred over the FEs approach in a number of settings (Raudenbush, 2009; Bell and Jones, 2015; Bell, Fairbrother and Jones, 2016).

However, 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 unobserved group-invariant characteristics is avoided. In a longitudinal setting, the serial auto-correlation of lowest level variables can be controlled for with clustered or properly modelled autoregressive standard errors. Indeed, 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. For all these reasons, we opt for estimating our models by means of FEs.

3.2. The model

In order to test our hypotheses, the following four-way FEs regression is going to be estimated:

$$AppXinv_{\,i,f,m,t} = CITYstock_{\,c,t-1} + FIRMstock_{\,f,t-1} + NETstock_{\,i,t-1} + Controls_{\,i,f,c,t-1} + \delta_{\,i} + \delta_{\,f} + \delta_{\,c} + \delta_{\,t} + \varepsilon_{\,i,f,m,t}$$

,

where *i* is the inventor, *f* the firm, *c* the city, *t* stands for time and the δs are a set of FEs. The three main explanatory variables account for the knowledge potential of the multilevel structure the inventor is embedded in, respectively the city, the firm and the network of past collaborators. In our specification, 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).

4. Data

As anticipated above, we extensively exploit patent data to build our main 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, the inventor's engagement with one or more firms and the city he/she belongs to.

We match two different patent databases in order to retrieve all the necessary information about our three layers of interest: the ICRIOS Patent Database (2014) (Coffano and Tarasconi, 2014) and the OECD HAN Database (2016). Both databases have the crucial feature of being the output of a process of name-disambiguation: inventors' names in ICRIOS, through inventors' IDs assigned by Pezzoni, Lissoni and Tarasconi (2014), and applicants' names in HAN. Out of this matched dataset, we build our main variables of interest: the number of patent applications per inventor-year (dependent variable), and a series of knowledge (patent) stocks for i) the inventor's network of collaborators in a five year window, ii) the firm-year (proxied by the applicant name listed in the patent document) and iii) the city-year tuples.

Each patent is assigned to a repository (the network, the firm, and the city) 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 the knowledge creation effort and result 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, see Hall, Jaffe and Trajtenberg, 2005).

4.1. Explanatory and Dependent Variables

City. We identify our city boundaries using EUROSTAT's definition of "Metropolitan Regions", which correspond to "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 Metropolitan Region definition, which emerged as relevant according to patent production rates. Indeed, we retained FUAs whose yearly patent production is equal to that of Metropolitan Regions belonging to the upper last quartile of Metropolitan Regions' patent distribution (e.g. Cambridge Area). In order to locate inventors in these cities and compute cities' knowledge stocks, we match our databases to the OECD REGPAT database (Maraut et al., 2008) which provides regionalized information (NUTS3 level for Europe) for all EPO inventors.

Firm. We use the applicant name – usually the owner of the patent – listed in patent documents as a proxy for the firm (or other organizations, such as universities or research centres) for which the inventor works. Homogenised firm names come from the OECD HAN Database 2016, which exploits the ORCID database for applicants' name harmonisation. Applicants who are individuals (physical persons) are removed from the analysis. Some firms are multi-establishment entities, and some of these establishments could be scattered in different cities. In order to account for the localized nature of knowledge production, i.e. its embeddedness into a territory, we compute the knowledge stocks together for all establishments within a given city (so we treat same-city establishments as a unique firm, and multi-city establishments of the same firm as different units). As in the majority of patent documents only the headquarter address is reported, we mark the presence of a firm into a territory thanks to the geo-localization of the inventors' addresses employed by that applicant.

Collaboration Network. The definition of the collaboration network needs to be set. We choose to consider those inventors who collaborated with the focal inventor within a 5-year window up to one year before the focal year. Other definitions may be plausible; we reasonably assume that a past collaboration remains an active source of knowledge for a 5-year period at most (Breschi and Lenzi, 2016). In order to quantify

the network stock, we sum up the individuals' depreciated applications stock of the collaborators.

Inventor's patent production. The dependent variable of our baseline model is a bare count of the yearly patent applications signed by an inventor. We only observe non-zero counts, hence we exploit the variance in the size of each inventor's production. Moreover, in order to control for individual time-invariant unobserved characteristics, we retain only inventors who invented at least twice.

Inventors' quality-adjusted patent production. As an alternative dependent variable, we also computed the count of high-quality yearly patent applications signed by an inventor. High-quality patents are defined as the top-50% patents sorted by their forward citations received – within a time window of 5 years after the priority year of the cited patent – controlling for the technological area and cohort (Waltman et al., 2011; Wohlrabe and Bornmann, 2017). Citations data are retrieved from ICRIOS Patent Database (2014). DOC_DB family data is used to compute the forward citations, hence including direct citations coming from other EPO documents as well as citations coming indirectly from other non-EPO patents (but collapsed by families to avoid double-counting).

As we want to explore further the non-trivial relationship between the knowledge embedded in each layer and inventors' creative production, we qualify each layer's knowledge stock with a measure of modular complexity, as introduced in Fleming and Sorenson (2001, 2004). Modular complexity describes a characteristic of each patent understood as bundle of recombined knowledge pieces (Weitzman, 1998). 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 listed in the patent document. 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 properly.

We operationalize such procedure as follow. In a first step, 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 a second step, we calculate the modularity index for each patent application document: the count of technological classes of the focal patent divided by the sum of their Eases of Recombination – an inverse weighted average. In the third and final step, we compute the stock of patent applications at each layer – as we previously did – split by the level of modular complexity. More precisely, we look at the modular complexity distribution of all patent applications, and we assign to "low" complexity the applications belonging to the lower 50% of the distribution, to "medium" those belonging to the upper 50-to-90% of the distribution, and "high" the remaining ones. Finally, we compute three separate stocks for each level of modular complexity, for each layer. Before running through this three step algorithm, we follow the recommendations by Alstott et al. (2017) and apply a normalization procedure, i.e. a Null Model, to the incidence matrix describing occurrences of technological classes within patent applications. In so doing, we clear the probability that two technological classes recombines from a random component originating from technological class and patent populations sizes. Further documentation is available upon request.

All variables enter the regression models after an Inverse Hyperbolic Sine (IHS) transformation, which is a log-like transformation well-defined at zero (differently from the natural logarithm). Moreover, city and firm knowledge stocks are lagged one year to lessen simultaneity issues, as well as the network stock that is defined up to one year before the focal year.

4.2. 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 keep at a minimum the correlation with the stock variables, instead of the bare count we use the average number of patents in the upper 50% of the distribution corrected by year and technology (Waltman et al., 2011; Waltman and Schreiber, 2013; Wohlrabe and Bornmann, 2017). Hence, these productivity controls really control for efficiency and high standards of the knowledge layers.

To smooth temporal disturbances, we compute the individual inventor's productivity on a time window between *t* and *t-4*. Consequently, averages are computed at each layer. The network's productivity is computed on a 5-year window from *t-1* backwards, whereas the firm level variable is 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). City-level average productivity enters the regressions as the respective value for each city 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). As stated by Crescenzi, Gagliardi and Iammarino (2015), "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. Once again, we overcome data availability constraints exploiting the information in patent data. We set up an algorithm checking if firms are inventing in more than one nation, and if employed inventors declare to live in countries different from that of the firm's headquarter.

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 our city-level analysis by means of the EUROSTAT definition of Metropolitan Regions) on GVA, population and employment, which we use to compute gross value-added per capita (GVA pc), population density and a Herfindahl-Hirschman index of employment specialization. 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 section.

4.3. Final dataset

The inventor-firm-city matching process generated multiple ambiguous assignations, e.g. more than one city or firm 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 firm ID and for the city ID. The rationale behind these algorithms is continuity, i.e. we want to detect when mobility patterns of inventors across firms and cities 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

- 272.404 multiple inventors (inventors that applied for patents more than once)
- 66.288 applicants (mainly firms, though not only)
- 320 MA
- over a total of 30 years covered, from 1980 to 2010.

4.4. Descriptive evidence

Figure 1 shows the cities considered in the present study, coloured according to their level of knowledge stock computed in 1980 and 2010 (the two extremes of our period of analysis). As can be seen, the cities with largest knowledge stocks are situated in the core of Europe. Some cities (particularly in the South of Europe) seem to converge in terms of knowledge stocks (they scale up to belong to the group with largest stocks). However, overall, the cities with the largest stocks are constant over time. Table 1 lists the top-10 firms considered in our study, sorted by their knowledge stocks in 2010. As can be seen, the top firms are usually large multinational companies, in some cases showing up several times in the top list as a consequence of their multi-city presence. Table 2 provides the descriptive statistics of our sample, while table 3 presents the correlation matrix. Some of the variables are highly correlated – particularly the network knowledge stock and the network productivity variable. However, simple correlations might not be adequate in a panel data framework to gauge multicollinearity problems, and therefore we run Variance Inflation Factor (VIF) tests after regressions. Fortunately, these point to the absence of collinearity problems.

The appendix further explores our dataset. Figure A.1 in the Appendix 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 A.2 to A.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 A.3 shows that there are many inventors without any past collaboration.

5. Results

Results of the 4-way FE regressions are shown in Table 4. Models 1 to 5 progressively plug in explanatory variables. From column 1 we learn that the effect of city-level knowledge stock is positive and significant. Doubling the size of the stock augments productivity by 4.9% – results from HIS transformed variables are interpreted as elasticities. These results are not far from the ones found in economic geography when estimating agglomeration effects (Rice et al., 2006), despite not being exactly the same phenomenon. The following columns introduce explanatory variables in a cascading way. Column 2 introduces firm-level stocks, which are positive and significant, but to a lesser extent (way smaller elasticity). Network-level stock is introduced in column 3. Simultaneous consideration of the three relevant knowledge flows levels let us better gauge the respective coefficients than previous research. In terms of magnitude, city knowledge stocks are still much stronger than network-level; city-level stocks' regression coefficient shrinks but still keeps its lead. Columns 4 and 5 introduce the remaining relevant controls, and some interesting findings emerge. First, the role of city-level stocks remain positive and significant, and with similar size as compared to estimates without controls. Inventors take also advantage of the firm-level knowledge pool, although the elasticity is relatively small compared to city-level stocks. Finally, once productivity controls are accounted for, the network-level stock diminishes its coefficient and becomes not significant. The negative and significant signs of controls regarding the firm and the city productivities are a signal that peers innovativeness tends to inhibit inventors capacity to produce new knowledge, instead of providing a stimulating work environment. In column 6 we assess the validity of the assumption of independence across levels by introducing interaction effects between city-, firm- and network-level stocks. In particular, we pursue the question of whether the different levels taken into account are complementary or substitutive among them. We find support for the hypothesis that network and city stocks are substitute (negative coefficient: the marginal effect of each activity decreases in the presence of the other activity). That is, inventors can rely on their networks when city-level knowledge stocks are weak (networks are not necessarily nested in cities). On the contrary, firm and network levels reinforce one another at the margin (positive coefficient), indicating positive feedbacks between high firm-level knowledge stocks and the actual connections of workers. This seems to suggest that firm-level knowledge stocks have stronger effects when its workers connect each other by means of actual collaboration links. Similarly, it suggests that the inventor's ability to exploit efficiently the external knowledge coming from his network depends also 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 city boundaries.

[Insert table 4 here]

Table 5 splits the knowledge stocks in each layer according to their degree of complexity (low, medium and high). Again, interesting results emerge. For low and medium levels of complexity of the stocks, results are similar to previous tables: larger effects for city-level stocks, then firm-level stocks, which are also significant, and finally the negligible influence of network-level knowledge stocks. However, some differences are worth reporting. The network-level coefficient of low complexity is actually significant, but negative. That is, when inventors access knowledge of low complexity through their network, this is not only redundant (they already access it through the city and firm levels) but even have congestion effects: acquiring this redundant knowledge makes them losing precious time for being productive elsewhere. Interestingly, as soon as we augment the degree of complexity of the stocks, the network-level knowledge increases its coefficient and becomes positive, but still not significant. On the contrary, for high levels of complexity, knowledge resists diffusion when inventors access knowledge at the city level, while the coefficient remains strongly significant at the firm level, suggesting the importance of close interactions within organizational boundaries to transfer complex knowledge, where possibly onsite demonstrations and direct monitoring is the rule. Still, network-level knowledge stocks are not significantly different from zero.

Column 4 introduces all the three complexity levels (low, medium and high) for the three focal explanatory variables. Results change a bit, especially for the city-levels stocks. However, we are reluctant to make our preferred outcome due to the huge collinearity between levels of complexity (especially within the city-level).

[Insert table 5 here]

Next, table 6 switches our dependent variable by the quality-adjusted one. As anticipated in the introductory section, results changes upside down. From column 1 we learn that city-level knowledge stocks do not have any effect on the number of high-quality patents the inventors can produce. This result suggests that if inventors aim to produce breakthrough ideas, they cannot rely on the knowledge stock of the area where they live, but only on more direct links and ties. Both the firm- and the network-level stocks are positive and significant now – the latter with the largest elasticity.

One explanation for such results is that inventors need to access and combine the necessary pools of knowledge allowing them to produce breakthrough patents. These are usually more tacit, and therefore closer interactions and trust between the sender and the receiver of the messages are crucial. These pools of knowledge tend to be of highly complex too. To investigate this, we further split our knowledge stocks by their degree of complexity (low, medium and high). We indeed find that the more complex the knowledge stocks are, the higher their effects on high-quality patenting. This is the case for the network-level stocks (possibly where interactions and links between inventors are stronger), while no differences across complexity levels are found for the case of firm-level knowledge stocks.

[Insert table 6 here]

In the Appendix, Tables A.1 and A.2 report some robustness checks. First, we substituted the establishment level variables with the firm level ones (table A.1). Results are very similar to those of Table 4. In Table A.2 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. In unreported results, we present the same regressions without *Network Avg. Productivity*. Indeed, as the correlation matrix in Table 3 shows, the network average productivity and the network-level 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.

6. Conclusions

Understanding the mechanisms of knowledge diffusion is a relevant topic in economics and other social sciences since knowledge creation is at the very base of the innovative dynamics and behind the economic growth of firms, cities and countries. 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 as well as economic geography, 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, city knowledge diffusion stands up as a significant and sizeable force enhancing knowledge production. The firm-level 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 with respect to the territorial environment. Finally, when the quality of the ideas produced is taken into account, and the type of knowledge accessed to produce them, the picture change considerably, and the network-level knowledge pools emerge as the most preponderant source of ideas.

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.

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S. and Sterzi, V. (2018) *Dancing with the Stars: Innovation Through Interactions*. Working Paper, National Bureau of Economic Research.

Allen, T.J. (1977) Managing the flow of technology: Technology transfer and the dissemination of technological information within the R & D organization(Book). *Research supported by the National Science Foundation. Cambridge, Mass., MIT Press, 1977. 329 p*.

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

Alstott, J., Triulzi, G., Yan, B. and Luo, J. (2017) Mapping technology space by normalizing patent networks. *Scientometrics*, **110**, 443–479.

Amin, A. and Cohendet, P. (2005) Geographies of Knowledge Formation in Firms. *Industry and Innovation*, **12**, 465–486.

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. 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.

Antonelli, C., Patrucco, P.P. and Quatraro, F. (2011) Productivity growth and pecuniary knowledge externalities: An empirical analysis of agglomeration economies in European regions. *Economic Geography*, **87**, 23–50.

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.

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

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.

Carlino, G. and Kerr, W.R. (2014) *Agglomeration and Innovation*. Working Paper, National Bureau of Economic Research.

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.

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.

Duranton, G. and Puga, D. (2001) Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products. *American Economic Review*, **91**, 1454–1477.

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

Feldman, M.P. and Audretsch, D.B. (1999) Innovation in cities:: Science-based diversity, specialization and localized competition. *European Economic Review*, **43**, 409–429.

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.

Fitjar, R.D. and Rodríguez-Pose, A. (2017) Nothing is in the Air. *Growth and Change*, **48**, 22–39.

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.

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

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.

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.

Jaffe, A.B. and de Rassenfosse, G. (2017) Patent citation data in social science research: Overview and best practices. *Journal of the Association for Information Science and Technology*, **68**, 1360–1374.

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.

Jaravel, X., Petkova, N. and Bell, A. (2018) Team-Specific Capital and Innovation. *American Economic Review*, **108**, 1034–1073.

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.

Kogut, B. and Zander, U. (1992) Knowledge of the Firm, Combinative Capabilities, and the Replication of Technology. *Organization Science*, **3**, 383–397.

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.

Lucas, R.E. (2009) Ideas and Growth. *Economica*, **76**, 1–19.

Lucas, R.E. and Moll, B. (2014) Knowledge Growth and the Allocation of Time. *Journal of Political Economy*, **122**, 1–51.

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.

Pezzoni, M., Lissoni, F. and Tarasconi, G. (2014) How to kill inventors: testing the Massacrator© algorithm for inventor disambiguation. *Scientometrics*, **101**, 477–504.

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

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.

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.

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

Storper, M. and Venables, A.J. (2004) Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, **4**, 351–370.

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.

Teece, D.J. (1977) Technology transfer by multinational firms: The resource cost of transferring technological know-how. *The economic journal*, **87**, 242–261.

Uzzi, B. (1997) Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness. *Administrative Science Quarterly*, **42**, 35–67.

Waltman, L., van Eck, N.J., van Leeuwen, T.N., Visser, M.S. and van Raan, A.F.J. (2011) Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, **5**, 37–47.

Waltman, L. and Schreiber, M. (2013) On the calculation of percentile-based bibliometric indicators. *Journal of the American Society for Information Science and Technology*, **64**, 372–379.

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*.

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

Wutchy, 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.

Figures and tables

Figure 1. Cities/metro areas considered and their distribution of knowledge stocks

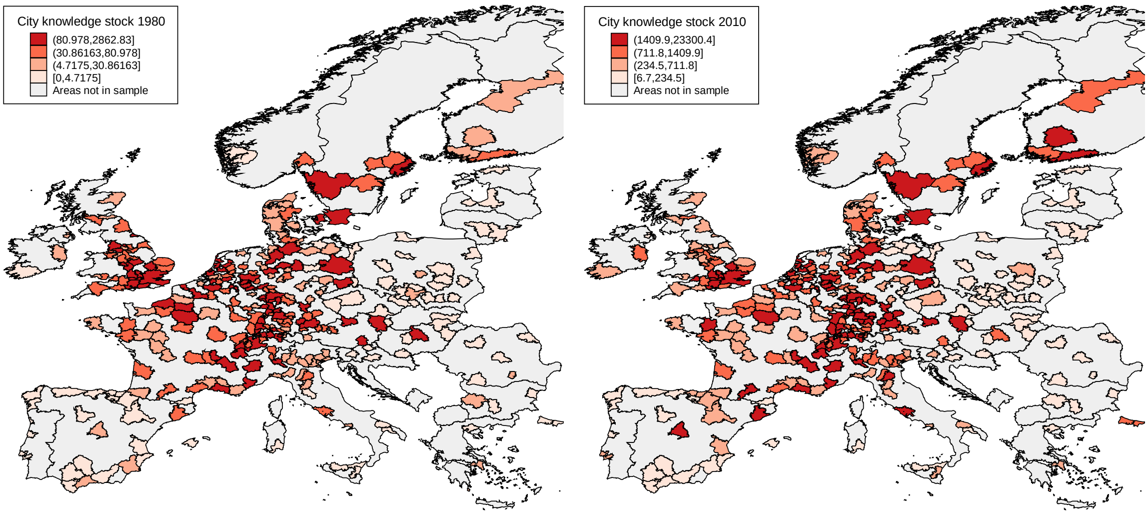

Figure 1.A: City-level knowledge stocks 1980 **Figure 1.B: City-level knowledge stocks 2010**

Table 1. Top-10 firms by their stock of knowledge in 2010 – with their city location

Firm's name	Firm's city location	Stock of knowledge 2010
PHILLIPS	Eindhoven	3224.41
BASF	Mannheim	2338.96
ROBERT BOSCH	Stuttgart	2136.88
BAYER	Köln	2012.80
SIEMENS	München	1982.28
HOECHST	Frankfurt	1176.25
SIEMENS	Nürnberg	1173.57
BASF	Heidelberg	1104.59
BAYER	Düsseldorf	855.85
CIBA GEIGY	Bael	823.38

Table 2. Descriptive statistics.

Statistic	N	Mean	St. Dev.	Min	Max
Patents per inventor-year	818,883	1.203	0.503	0.881	6.704
Quality-adj. patents per inventor-year	818,883	0.59	0.629	0	5
City stock	818,883	8.266	1.326	0	10.751
Firm stock	818,883	3.713	2.563	0	9.521
Network stock	818,883	1.883	1.932	0	9
City stock (low complex.)	818,883	7.284	1.305	0	9.728
City stock (medium complex.)	818,883	7.362	1.374	0	9.861
City stock (high complex.)	818,883	6.575	1.479	0	9.382
Firm stock (low complex.)	818,883	2.425	2.408	0	8.46
Firm stock (medium complex.)	818,883	2.66	2.595	0	8.891
Firm stock (high complex.)	818,883	1.941	2.5	0	8.14
Network stock (low complex.)	818,883	0.846	1.371	0	7
Network stock (medium complex.)	818,883	1.09	1.593	0	9
Network stock (high complex.)	818,883	0.683	1.42	0	9
City avg. productivity	818,882	0.244	0.098	0	0.891
Firm avg. productivity	818,845	0.284	0.266	0	3.886
Network avg. productivity	818,883	0.257	0.393	0	5
Multinational	818,883	0.65	0.477	0	1
Specialization index	694,876	0.242	0.017	0.185	0.475
Population density	694,404	0.507	0.313	0.025	2.12
GVApc	694,866	4.013	0.282	1.356	4.943

Table 3. Correlation Matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1	1											
2	0.59	1										
3	0.08	0.04	1									
4	0.19	0.13	0.36	1								
5	0.32	0.21	0.17	0.35	1							
6	0.14	0.10	0.36	0.30	0.31	1						
7	0.24	0.23	0.21	0.45	0.44	0.44	1					
8	0.28	0.24	0.14	0.31	0.78	0.30	0.50	1				
9	0.14	0.13	0.18	0.51	0.25	0.20	0.29	0.22	1			
10	0.02	0.03	0.15	0.01	0.05	0.15	0.07	0.06	0.02	1		
11	0.04	0.02	0.39	0.13	0.09	0.30	0.14	0.09	0.05	0.35	1	
12	0.08	0.04	0.68	0.23	0.13	0.29	0.19	0.12	0.13	0.21	0.23	1

Note : 1: Patents per inventor-year; 2: Quality-adj. patents per inventor-year; 3: City stock; 4: Firm stock; 5: Network stock; 6: City avg. productivity; 7: Firm avg. productivity; 8: Network avg. productivity; 9: Multinational; 10: Specialization index; 11: Population density; 12: GVApc

Table 4. Baseline results. 4-ways FE OLS regression with clustered standard errors at the city level

	Dependent variable: *applications per Inventor-Year*					
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
City Stock	0.049*** (0.023, 0.076)	0.045*** (0.020, 0.070)	0.041*** (0.017, 0.066)	0.047*** (0.022, 0.071)	0.046*** (0.021, 0.070)	0.046*** (0.023, 0.070)
Firm Stock		0.004* (-0.0001, 0.009)	0.002 (-0.002, 0.006)	0.010*** (0.006, 0.013)	0.009*** (0.005, 0.013)	0.013 (-0.003, 0.029)
Network Stock			0.008*** (0.003, 0.012)	0.002 (-0.002, 0.006)	0.002 (-0.002, 0.007)	0.002 (-0.013, 0.017)
Firm Avg Productivity				-0.184 (-0.457, 0.090)	-0.182 (-0.457, 0.094)	-0.279** (-0.547, -0.011)
City Avg Productivity				-0.140*** (-0.164, -0.116)	-0.139*** (-0.163, -0.115)	-0.114*** (-0.136, -0.092)
Network Avg Productivity				0.040*** (0.026, 0.054)	0.040*** (0.026, 0.054)	0.029*** (0.019, 0.039)
Multinational Firm (T/F)					0.042*** (0.034, 0.051)	0.038*** (0.030, 0.046)
Firm Stock * City Stock						-0.001 (-0.003, 0.001)
Network Stock * Firm Stock						0.004*** (0.002, 0.005)
Network Stock * City Stock						-0.002** (-0.004, -0.00004)
Observations	818,883	818,883	818,883	818,845	818,845	818,845
Inventor, firm, city & time FEs	YES	YES	YES	YES	YES	YES
Adjusted R^2	0.227	0.227	0.227	0.228	0.229	0.301

Table 5. Complexity. 4-ways FE OLS regression with clustered standard errors at the city level

	Dependent variable: *applications per Inventor-Year*			
	(1)	(2)	(3)	(4)
City stock low complex.	0.031***			0.013
	(0.008, 0.053)			(-0.017, 0.043)
Firm stock low complex.	0.013***			0.004**
	(0.010, 0.017)			(0.0005, 0.008)
Network stock low complex.	-0.006***			-0.005***
	(-0.009, -0.003)			(-0.008, -0.002)
City stock medium complex.		0.034***		0.025
		(0.011, 0.056)		(-0.006, 0.056)
Firm stock medium complex.		0.018***		0.013***
		(0.015, 0.021)		(0.010, 0.016)
Network stock medium complex.		0.001		0.002
		(-0.003, 0.005)		(-0.002, 0.005)
City stock high complex.			0.009	-0.002
			(-0.006, 0.025)	(-0.020, 0.016)
Firm stock high complex.			0.017***	0.012***
			(0.014, 0.020)	(0.008, 0.015)
Network stock high complex.			0.002	0.002
			(-0.003, 0.008)	(-0.003, 0.008)
City Avg Productivity	-0.204	-0.246*	-0.197	-0.248*
	(-0.478, 0.069)	(-0.517, 0.025)	(-0.458, 0.064)	(-0.522, 0.025)
Firm Avg Productivity	-0.113***	-0.119***	-0.113***	-0.122***
	(-0.134, -0.092)	(-0.140, -0.097)	(-0.134, -0.092)	(-0.143, -0.101)
Network Avg Productivity	0.054***	0.042***	0.041***	0.042***
	(0.040, 0.069)	(0.030, 0.055)	(0.028, 0.054)	(0.032, 0.053)
Multinational Firm (T/F)	0.037***	0.036***	0.037***	0.035***
	(0.029, 0.045)	(0.028, 0.044)	(0.029, 0.045)	(0.027, 0.043)
Observations	818,845	818,845	818,845	818,845
Inventor, firm, city & time FEs	YES	YES	YES	YES
Adjusted R^2	0.301	0.301	0.301	0.302

Table 6. Quality-adjusted patents and complexity. 4-ways FE OLS regression with clustered standard errors at the city level

	Dependent variable: quality-adjusted applications per Inventor-Year			
	(1)	(2)	(3)	(4)
City Stock	0.006			
	(-0.029, 0.040)			
Firm Stock	0.005**			
	(0.0003, 0.009)			
Network Stock	0.017***			
	(0.014, 0.020)			
City stock low complex.		0.024		
		(-0.009, 0.057)		
Firm stock low complex.		0.007***		
		(0.003, 0.010)		
Network stock low complex.		0.005**		
		(0.001, 0.009)		
City stock medium complex.			0.012	
			(-0.018, 0.042)	
Firm stock medium complex.			0.010***	
			(0.006, 0.013)	
Network stock medium complex.			0.012***	
			(0.008, 0.016)	
City stock high complex.				-0.007
				(-0.027, 0.013)
Firm stock high complex.				0.007***
				(0.003, 0.011)
Network stock high complex.				0.014***
				(0.009, 0.019)
City Avg Productivity	0.205*	0.165	0.142	0.164
	(-0.006, 0.416)	(-0.048, 0.377)	(-0.074, 0.358)	(-0.054, 0.383)
Firm Avg Productivity	-0.349***	-0.344***	-0.346***	-0.342***
	(-0.396, -0.301)	(-0.389, -0.298)	(-0.392, -0.300)	(-0.387, -0.298)
Network Avg Productivity	-0.169***	-0.116***	-0.136***	-0.129***
	(-0.201, -0.137)	(-0.148, -0.084)	(-0.165, -0.107)	(-0.155, -0.103)
Multinational Firm (T/F)	0.028***	0.027***	0.027***	0.028***
	(0.018, 0.038)	(0.018, 0.037)	(0.017, 0.037)	(0.018, 0.038)

Observations	818,845	818,845	818,845	818,845
Inventor, firm, city & time FEs	YES	YES	YES	YES
Adjusted R^2	0.275	0.274	0.275	0.275

Appendix

Figure A.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.*

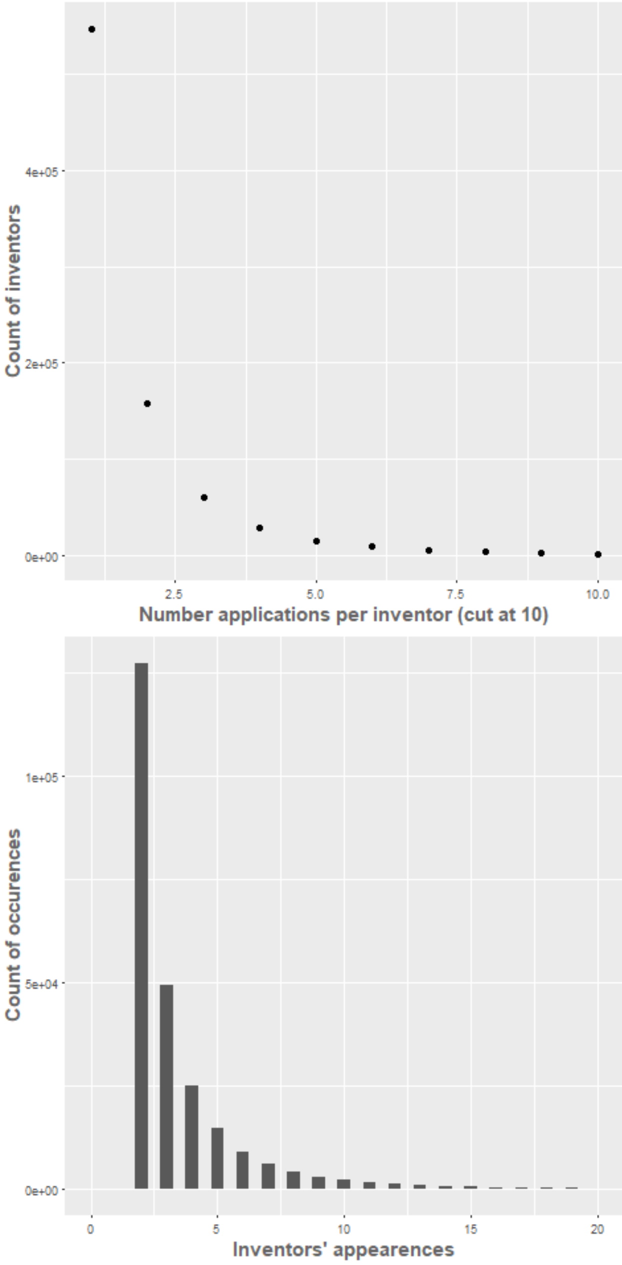

Figure A.2. *Mean Firm and Establishment Knowledge Stock across years and MRs. The overall Pearson Correlation Coefficient is 0.85, indicating that the two measures are indeed strongly related, but still they measure slightly different aspects*

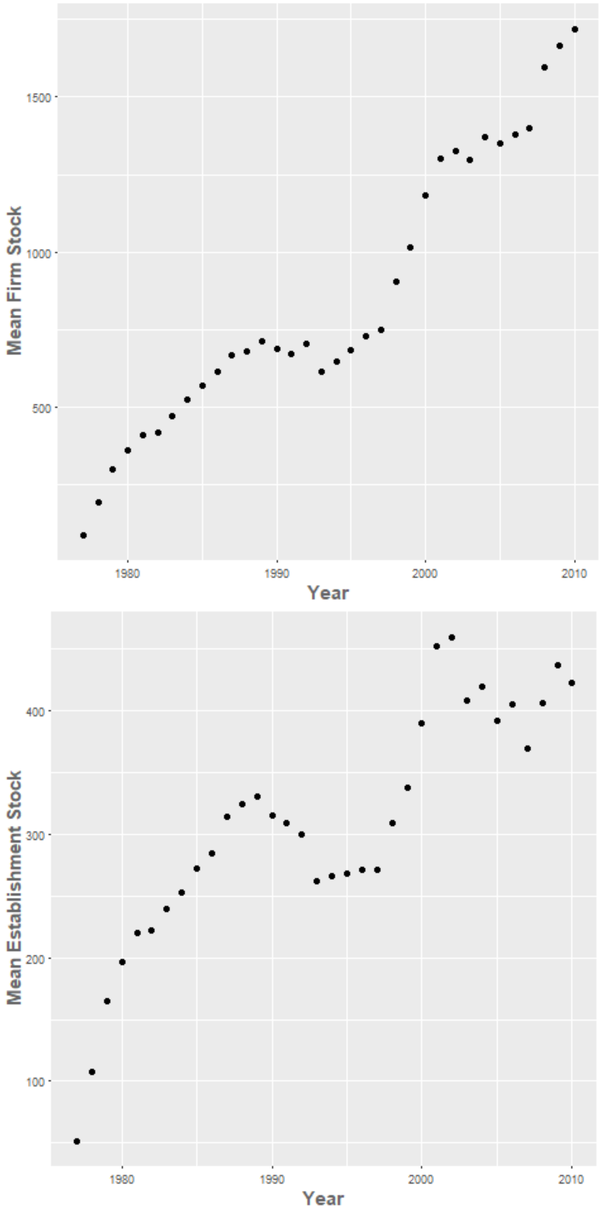

***Figure A.3**. Mean of the Network Stock and Network Size, 1st and 3rd quartile*

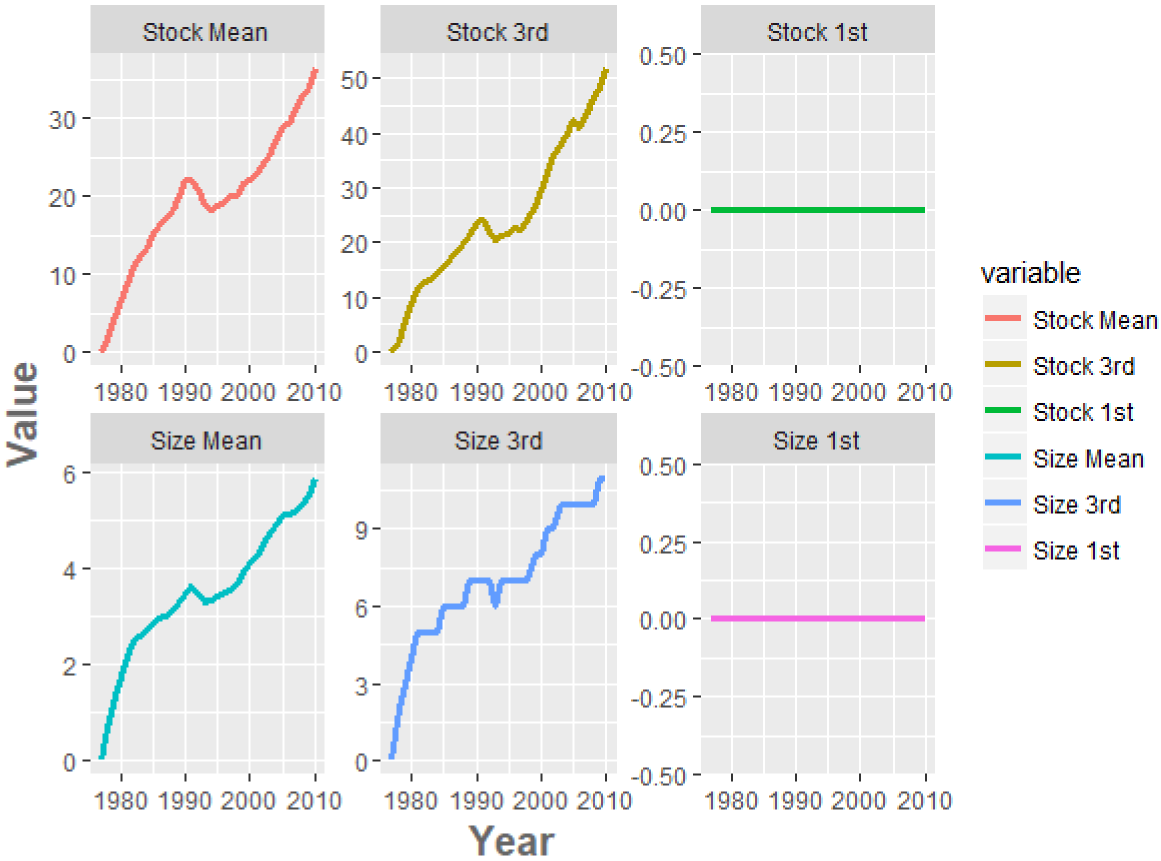

***Figure A.4**. Mean Establishment Stock across Countries.*

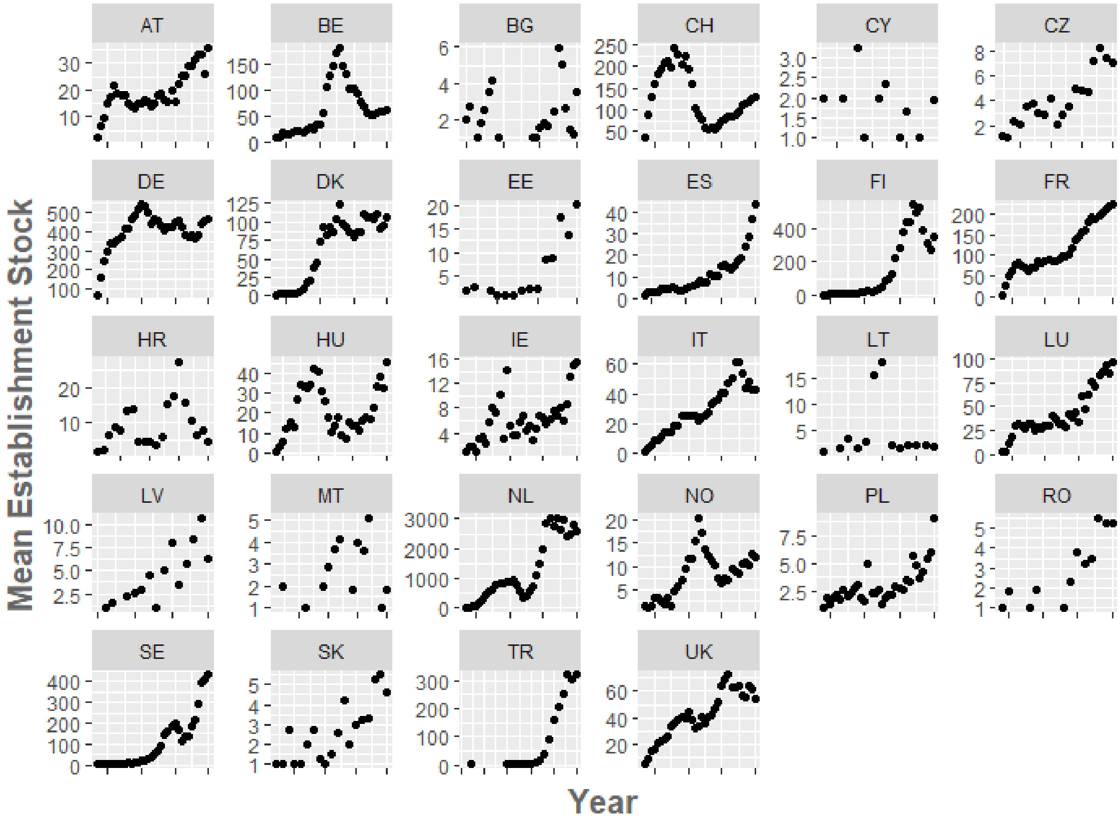

***Figure A.5**. Mean MR Stock across Countries.*

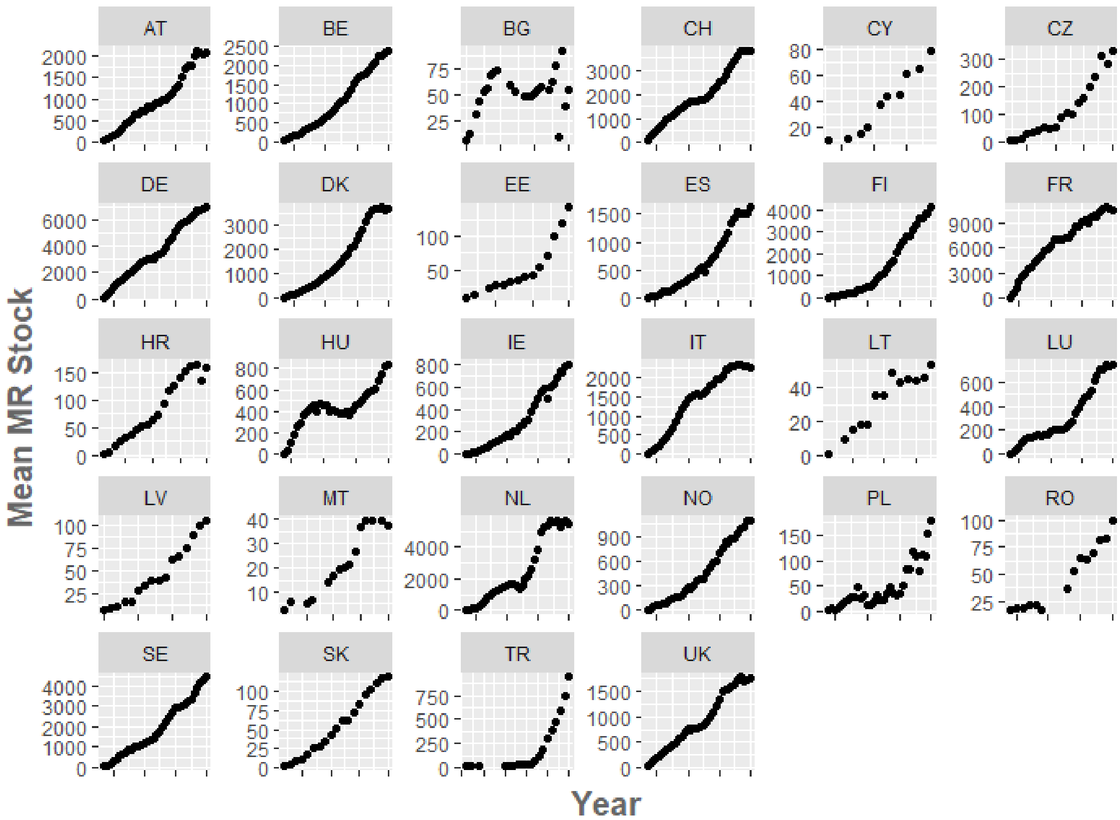

Table A.1. Robustness check (1)

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)
City Stock (lag 1)	0.037***	0.035***	0.041***	0.040***
	(0.012)	(0.012)	(0.012)	(0.012)
Large 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 A.2. 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)
City Stock (lag 1)	0.031**	0.030**
	(0.014)	(0.014)
Firm 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)
Firm Avg Productivity (lag 1, 3-years window)	-0.083***	-0.080***
	(0.004)	(0.005)
City 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 * Firm Stock (lag 1)		0.003***
		(0.001)
Network Stock * City Stock (lag 1)		-0.002
		(0.001)
Firm Stock (lag 1) * City 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: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Paper formatting

The total length of the paper should not exceed 25 pages and maximum size of 2MB. It should be sent in Microsoft Word (.doc) using the attached pattern.

The paper should be divided into numbered sections (bold).

Tables and figures should be numbered consecutively in accordance with their appearance in the text.

Text should be 12-point Times Roman, typed in 1.5 spacing. Before and after paragraph spacing should set to 6 pts.

Page numbers should be in the bottom center (expect the first page).

Bibliography. Please follow these sample references:

World Bank (2002): *Cities on the Move: a World Bank Urban Transport Strategy Review, World Bank, Washington DC.*

Becattini, G. (2002): "Del distrito industrial marshalliano a la ‘teoría del distrito’ contemporánea. Una breve reconstrucción crítica", *Investigaciones Regionales*, nº1, p.9-32.