Inventor Mobility and Productivity in Italian Regions

Riccardo Cappelli [a], Dirk Czarnitzki [b], Thorsten Doherr [c], Fabio Montobbio [d]

a) (Corresponding author) Phone: +390116704406; Email: riccardo.cappelli@unito.it
Department of Economics and Statistics Cognetti De Martiis, University of Turin, Lungo Dora Siena 100 A, 10153 Turin, Italy.
b) Phone: +3216326906; Email: Dirk.Czarnitzki@kuleuven.be
Department of Managerial Economics, Strategy and Innovation, KU Leuven, Naamsestraat 69, 3000, Leuven, Belgium; and Centre for European Economic Research (ZEW), L 7 1, 68161, Mannheim, Germany.
c) Phone: +4906211235291; Email: doherr@zew.de
Centre for European Economic Research (ZEW), L 7 1, 68161, Mannheim, Germany; and University of Luxembourg, Luxembourg.
d) Phone: +390272342696; Email: fabio.montobbio@unicatt.it
Institute of Economic Policy, Università Cattolica del Sacro Cuore, Largo Gemelli 1, 20123 Milan, Italy; Collegio Carlo Alberto, Via Real Collegio 30, 10024, Moncalieri, Turin, Italy; ICRIOS, Bocconi University, Via Roentgen 1, 20136, Milan, Italy; and

Abstract

This paper describes the inter-regional and international mobility of inventors in Italy and estimates its impact on total factor productivity (TFP) at the regional level for the period 1996-2011. A new database of mobile inventors is constructed and, using a set of geography based instruments to address endogeneity, the paper shows that inventor inflows and outflows affect regional TFP growth. Moreover, the positive effects of the inventors' mobility (inflow) between different applicants take more time to materialize (relative to movements within the same company). Finally, the negative effects of inventor outflows are mainly driven by mobility between applicants.

Keywords: inventor disambiguation**;** inventor mobility; migration; economic growth, regional TFP, Italy

JEL: O47; J61; R23; O30

Inventor Mobility and Productivity in Italian Regions

Abstract

This paper describes the inter-regional and international mobility of inventors in Italy and estimates its impact on total factor productivity (TFP) at the regional level for the period 1996-2011. A new database of mobile inventors is constructed and, using a set of geography based instruments to address endogeneity, the paper shows that inventor inflows and outflows affect regional TFP growth. Moreover, the positive effects of the inventors' mobility (inflow) between different applicants take more time to materialize (relative to movements within the same company). Finally, the negative effects of inventor outflows are mainly driven by mobility between applicants.

Keywords: inventor disambiguation**;** inventor mobility; migration; economic growth, regional TFP, Italy

JEL: O47; J61; R23; O30

Introduction

Understanding the economic impact of diasporas, skilled labor mobility, and migration, on regional economic development is a critical issue for both sending and receiving regions. For the receiving regions, skilled migrants contribute to economic growth and generate significant externalities, promoting creativity, innovation and entrepreneurship (Nathan, 2014). For the sending regions, there is potential brain drain and a loss of human capital (Docquier and Rapoport 2012). In short, "communities on the move" influence the welfare and the economic conditions in both the region of origin and destination.

Several papers have studied how skilled labor mobility and immigration have affected innovative capacity (e.g. Trajtenberg, 2005; Hoisl, 2007; Hunt and Gauthier-Loiselle, 2010) and productivity growth (e.g. Peri, 2012). In addition, skilled labor mobility and diaspora networks are increasingly considered a key vehicle of knowledge spillovers (Almeida and Kogut 1999; Rosenkopf and Almeida, 2003; Breschi and Lissoni, 2009; Trippl, 2011).

Despite the growing interest in the diffusion of knowledge through labour mobility of high skilled workers, few studies analyse the relationship between skilled labour mobility (both outflows and inflows) within countries in Europe and regional economic performance. In fact diaspora effects can arise especially within a country as inter-regional migrants face lower barriers to migration than international migrants (Faggian et al, 2017). To the best of our knowledge, no study directly explores whether inter-regional outflows and inflows of skilled workers within a country affects differential regional economic productivity, as measured by total factor productivity (TFP).

This paper tries to fill this gap by analysing the effects of inventor mobility on TFP growth of Italian regions for the period 1996-2011. Italy is a relevant case because it has experienced very high rates of inter-regional and international mobility of the work force in recent years. In particular, movements of graduated and skilled labour have attracted the attention of policy makers raising important issues of brain drain and brain gain at the regional level (Becker et al. 2004; Fratesi and Percoco; 2014; Marinelli, 2013). Central and southern regions systematically lose their human capital. A recent report shows that five years after graduation (in 2010), 13% of graduates in Italian central regions and 26% of the graduates in the South moved mainly to the North (Alma Laurea, 2016).

This paper faces three main empirical challenges: the construction of a skilled labour mobility index, the estimation of the TFP for Italian regions and the identification of the effect of mobility on TFP. First, this paper builds a skilled labour mobility index using a novel database on Italian inventors of patents filed with the European Patent Office (EPO) to measure the regional rate of inflow and outflow (and, consequently, the net flow) of inventors between regions. Second, to measure TFP at the regional level, it adopts a growth accounting approach, estimating regions' capital stock (Maffezzoli, 2006). Finally, the last empirical challenge is to identify the effects of labor mobility on TFP growth in the presence of endogeneity, simultaneity and omitted variable biases. This paper identifies some regional characteristics, that are likely to be related to skilled mobility and much less to other determinants of productivity at the regional level. These are the geographical distance between the origin and the destination, a common border effect and regional fixed effects (see e.g.:Frenkel and Romer, 1999; Peri, 2012; Miguelez and Moreno, 2015). For each different calendar year, these geographic variables are good predictors of the inventors' inflow and outflow and, at the same time, a priori (being essentially based on geography) are not correlated with shocks on TFP. The empirical estimates are performed using both OLS and IV 2SLS fixed-effects techniques, including also proxies for other relevant determinants of regional productivity growth. The factors controlled for are R&D expenditure per capita, the ratio of patents to R&D expenditure and the population density.

This paper suggests that inventor mobility has a positive a significant impact on TFP. In particular it shows that the inflows of inventor has a positive impact and the outflow of inventor has a negative impact supporting the idea of a substantial economic adverse effect of the loss of human capital. In addition, this paper analyzes heterogeneous effects for "within applicant", i.e. an inventor moves between subsidiaries of the same employer across regions, and "between applicants" movements, i.e. between different firms. It shows that the effects of "between applicants" inventor inflows on TFP take more time to materialize than the "within applicant" inventor inflows. Moreover, the negative effects on TFP by inventor outflows are driven by "between applicants" movements.

The paper is organized as follows: Section 2 discusses why labour mobility and migration affects TFP. Section 3 describes the methodology. Section 4 explains the data. Section 5 focuses ons the empirical results. Section 6 provides some concluding remarks.

Background

Why do skilled labor mobility and, in particular, mobility of inventors affect TFP at the regional level? First, skilled labor mobility produces a better match of jobs and task specialization. Secondly, it stimulates innovation in regions, filling labor shortages in specific sectors, generating absorptive capacity and fostering significant knowledge diffusion. Regional TFP can be generated by entrepreneurial and innovation efforts involving migrants between different regions. For the same set of reasons, the authors of this paper expect that an outflow of skilled labor force induces a decrease in regional TFP.

The first set of explanations builds upon the job matching theory (Jovanovic, 1979); the idea is that inventors stay in jobs in which their productivity appears to be relatively high and that they leave those jobs in which their productivity is perceived to be low. So mobility is likely to increase the match quality between inventors and employers with a consequent increase in inventors' productivity (Hensen et al., 2009; Marinelli, 2013).

A better match between complementary skills induces also task specialization that, in turn, involves all the employees. Mobile inventors may contribute to make the colleagues more inventive. Complementary skills could involve management and entrepreneurship. At the same time, the inventor further benefits from the knowledge of his/her new colleagues[1]. It is worthwhile noting that this view implies that productivity increases *after* that inventor mobility takes place.

A number of papers has analysed migration in US cities. For example, Peri (2012) finds that immigration has a positive effect on state-level TFP. He also finds that a substantial portion of this effect depends upon increased task specialisation of native workers. Relatedly, cross-regional labour mobility generates a more culturally diverse workforce[2]. Individuals coming from different regions have different, complementary skills with respect to workers in the receiving region, and this can lead to the production of new ideas.

These mechanisms can take place also within a country. In Italy, for example, there is substantial labour mobility, and Fratesi and Percoco (2014) - using data on internal mobility for the period 1980-2001 - find a positive relationship between the net inter-regional inflows of skilled people (measured using the educational level) and the GDP per-capita growth of regions. The authors highlight the negative effects of the loss of human capital for the Southern regions of Italy.

A second set of explanations underlines that skilled labour mobility increases productivity because it stimulates innovation activities. Inventors play a crucial role and the literature tends to support the idea international labour movements in science and engineering have a positive effect on innovation (in most of the cases measured by patents). Using a 1940-2000 state panel in US, Hunt and Gauthier-Loiselle (2010) find that an increase in the share of tertiary educated migrants increases the number of patent applications per capita. Kerr (2010) shows that there is localized patent growth in US cities after breakthrough inventions. The spatial reallocation of patenting activities across US cities is faster if the technology has a more mobile workforce.

Concerning Europe, Bosetti et al. (2015), using a panel of twenty European countries, find that skilled migrants contribute positively to the number of patents and citations of scientific publications. Fassio et al. (2015) show that highly-skilled migration has a positive effect on innovation in Germany, France and UK. Finally tracking inventor mobility for 274 EU NUTS2 regions over 8 years, Miguélez and Moreno (2015) show that inventor mobility and co-patenting have a positive effect on regional patents per capita. They interpret their results in terms of absorptive capacity of the regions that benefit from the knowledge and information brought in by mobile inventors and cooperation networks.

Mobile inventors could increase innovation and productivity at the regional level if there are barriers to mobility and self-selection leads them to be more educated, more entrepreneurial or of higher unobserved inventive ability. However, there is substantial evidence that labour mobility of high skilled workers is a key mechanism of knowledge diffusion that overcomes geographic barriers and other constraints. When an employee changes jobs, he/she transfers from the old to the new firm detailed information on the technologies used in the previous employment and also the knowledge, skills and experience embedded in the mobile worker. Previous research shows a positive relationship between mobility and productivity of inventors (Trajtenberg, 2005; Hoisl, 2007).

Inventor mobility brings about learning by hiring (Almeida and Kogut 1999; Rosenkopf and Almeida, 2003; Breschi and Lissoni, 2009). Rosenkopf and Almeida (2003) show that inventor mobility increase the likelihood of knowledge flows between two firms (measured by patent citations) irrespective of the geographic location of the two firms, i.e. in the same or different regions. At more aggregate level, Almeida and Kogut (1999) show the close link between knowledge flows and labour mobility. They show that the mobility of engineers is an important factor explaining the localized diffusion of knowledge (measured by patent citations) within US regions. This stems from the fact that an important part of an invention is represented by the tacit knowledge embedded in engineers.

Beyond these direct effects mediated by market mechanism, inventor mobility positively affects firms and regions performance through knowledge externalities (Breschi and Lissoni, 2009). The mobility of workers creates links between firms through social ties, which involve the worker that moves and the workers in his or her previous firm. These ties favour the diffusion of knowledge among firms and regions (Breschi and Lissoni, 2009; Miguélez and Moreno, 2015; Cappelli and Montobbio, 2016). For example, Agrawal et al. (2006) show that an inventor who moves from one

US region to another is more likely to cite inventors in the previous region, compared to those who have never lived in that region. The social networks between inventors reduce the frictions in knowledge flows exerted by geographical factors such as physical distance.

It is important to note that a substantial part of the empirical work surveyed above deals with the impact of international migration on innovation. In fact, as pointed out by a recent review (Faggian et al, 2017), studies on the determinats of inter-regional mobility of high skilled workers outweight those examining its economic consequences, especially for the origin regions. It is possible that migration within a country is less affected by self selection than international migration. At the same time the impact on receiving regions can be substantial because of an easier matching of complementary skills and a stronger spillover effect generated by cultural and institutional proximity. Cross-regional network effects can play an important role because the community of origin could act as a placing agency, reducing the cost of finding a job in the region of destination.

In parallell, for the region of origin, the outflow of inventors may limit both the absorptive capabilities and the attractiveness for inventors coming from other regions, thereby reinforcing the direct negative consequence associated with the loss of human capital. Consequently, inter-regional inventor mobility may results in within-country divergence of regions.

Methodology

Empirical specification

To conduct the empirical analysis, this paper represents a region with a production function to calculate the TFP which, in turn, depends upon knowledge and technological variables[3]. Knowledge generated in a region is mainly measured using technological input measures like R&D expenditure and number of high skilled people and/or technological output measures like the number of patents and patent forward citations. On the other side, knowledge flows between regions are measured using indirect measures like the stock of foreign R&D or considering explicitly a channel of knowledge flows like inventor mobility and citations between inventors or scientists. In this work, the TFP growth in Italian regions is modelled using the following equation:

$$[1]\ \ln\left(\mathrm{TFP}_{i,t} / \mathrm{TFP}_{i,\,t-1}\right) = \alpha + \beta\ \ln \mathrm{TFP}_{i,t-1} + \varphi\ \mathrm{R\&Dpc}_{i,t-1} + \upsilon\ \mathrm{PATrd}_{i,t-1} + \Omega\ \mathrm{Mobility}_{i,t-1} + \eta\ \mathrm{Density}_{i,t-1} + \varepsilon_{it}$$

where $ln\left(TFP_{i,t} / TFP_{i,\,t-1}\right)$ is the TFP growth rate between year *t-1* and *t* of a given Italian region *i* and ($TFP_{i,t-1}$) represents the regional level of TFP. Following a standard growth accounting approach (Solow, 1957), these variables are constructed using a Cobb-Douglas production function with two input factors, i.e. labour and capital, and constant return to scale (for further details see Appenix A in the supplemental online material). The innovative efforts of regions are measured using R&D expenditure per capita ($R\&Dpc_{i,t-1}$). The ratio of patents and R&D expenditure ($PATrd_{i,t-1}$) is also included to check for an additional effect exerted by successful R&D. This paper uses the inventor mobility indexes ($Mobility_{it-1}$) to measure the inter-regional and international mobility of inventors in Italy. An inventor is considered as being mobile when he or she moves between regions. Three types of inventor mobility indexes are constructed: inflow of inventors from other regions ($Inflow\_rate_{i,t-1}$); outflow of inventors to other regions ($Outflow\_rate_{i,t-1}$); net inflows of inventors ($Netflow\_rate_{i,t-1}$), i.e. the difference between inflow and outflow of inventors. The inventor mobility indexes are expressed as the ratio of number of mobile inventors at year *t* and the regional stock of inventors at theyear *t-1*. In addition, population density ($Density_{i,t-1}$), measured as the number of thousand inhabitants per square kilometre, is included to control for agglomeration effects (Glaeser, 2010). Finally, ε_{it} represents the error term.

Identification strategy

In order to estimate equation [1] this paper needs to address some econometric issues. The literature on labour mobility (see e.g.: Ortega and Peri, 2014) clearly demonstrates that economic factors like the expected earnings in the destination area are important in explaining the observed migration patterns. Then, it could be the case that highly efficient regions attract inventors more than less efficient regions. Moreover, TFP shocks might affect the relative degree of attractiveness of regions. This means that the relationship between TFP and inventor mobility might be bi-directional and introduces an endogeneity problem which might result in biased estimates of the coefficients of the inventor mobility indexes.

In order to solve the endogeneity problem, this paper adopts several strategies. Firstly (see equation [1]), the mobility indexes and all the other independent variables are lagged by one year. Moreover, a fixed-effects OLS estimator is adopted to take into account of all the unobservable factors related to region's attractiveness to inventors. However, these expedients still do not completely solve the potential omitted variable problem and endogeneity bias. To further address these issues, this paper adopts a 2SLS fixed-effects technique.

Following Frankel and Romer (1999), the instruments used in 2SLS estimates are constructed using the predicted values from gravity model estimates where bilateral inventor flows are explained by regions' geographic characteristics. Migration costs related to physical distance and regional national borders clearly affect inventor mobility between regions (Ortega and Peri, 2014; Migueléz and Moreno, 2015), while, on the other side, it can be safely assumed that these geographic factors are not correlated with regional TFP. For further details on the construction of the three instrumental variables, i.e. one for each of the three categories of inventor flows, see Appendix B in the supplemental online material.

Data

This paper constructs a set of variables for all the twenty Italian administrative regions for the period 1995-2011. [4] A first group of variables is constructed using data from the Italian National Institute of Statistics (ISTAT), i.e. total R&D expenditure (used to construct the variable

$R\&Dpc_{i,t-1}$), population and area in square km (used to construct the variable $Density_{i,t-1}$). In addition, this paper relies on other data sources: PATSTAT data to construct the number of patents (used to build the variable $PATrd_{i,t-1}$); EUROSTAT data on the coordinates of regional centroids (used to build the variable $dist_{ij}$). [5]As mentioned above, to develop the empirical analysis this paper has to address the challenge of measuring both geographical inventor mobility and regional TFP.

TFP of Italian regions

This paper measures the TFP of Italian regions, both in level and growth rate, as Solow's residual to GDP once the contribution of two input factors, i.e. labor and capital, are taken into account. This paper relies on ISTAT data to construct the variables on regional TFP: GDP at constant prices[6]; number of full time equivalents[7] as measure of labor input; ratio between compensation of employees and GDP as measure of GDP elasticity to labor. Data on capital stock are not available at regional level, but ISTAT provides data on regional fixed investment for the period 1995- 2011. These short time series data on regional investments allow one to obtain, through a perpetual inventory method, only a partial approximation of the capital stock of regions. Moreover, the quality of the approximation worsens for the first part of the period as the length of the series on fixed investments is getting shorter. In order to reduce this shortcoming of the simple perpetual inventory method, this paper uses the procedure developed by Maffezzoli (2006). The basic idea is to integrate the regional fixed investments data with the time series data on national capital stock (available from ISTAT) in order to construct a measure of regional capital stock using as much as information as possible (forfurther details, see Appendix A in the online supplemental material).

Mobility of Italian inventors

To construct the various patent-based measures, this paper relies on PATSTAT data on EPO patent applications. However, for the identification of the geographical mobility of inventors, the original

PATSTAT data suffer some important limitations because of the "who is who" and the "John Smith" problems (Trajtenberg et al., 2006). The former refers to the fact the name of an inventor with two or more patents may be spelled differently on different patents. The latter refers to the same name sometimes referring to different inventors. To overcome these limitations, this paper builds a separate dataset using a procedure referred to as "name game" analysis (Trajtenberg et al., 2006). To tackle the common name issue, an inventor career in patent data is represented by documents that not only share the name of the inventor but also additional characteristics like the assignee, addresses, co-inventors, citations and so on. Whether a set of mutual characteristic between two documents is sufficient for a valid connection depends on a heuristic plausibility check. The algorithm is based on an hierarchical order of the characteristics starting with the inventor address. All patents sharing a similar address for a given inventor name are considered to be from the same person. These patent clusters by themself are not able to identify mobility, but in association with the next entry in the hierarchy, the assignees, we are able to create an intransitive network between these non-mobile clusters. Traversal of this network leads to an onion like structure of layered clusters, already containing mobility. This process is repeated for the remaining characteristics like citations or co-inventors ending with the international patent classification. To avoid huge clusters of documents perceived as unrelated, a circumstance explained by the intransitivity of the connections, a system of plausibility checks, based on aggregated meta information within a cluster and exogenous assessment of the resulting mobility, has to be passed. If a cluster is not able to fulfill the requirements of a plausibility check, the contained network is repeatedly traversed with increasingly restrictive rule sets until it is separated into plausible sub-clusters. The method is fully described in Doherr (2017).

To verify the overall quality of this procedure in regard of precision and recall rate, we conducted the benchmark Lissoni et. al. (2010) proposed for the algorithm challenge of the APE-INV (Academic Patenting in Europe - Inventors) initiative of the European Science Foundation. It is based on individually verified EPO patent links of 424 French and 121 Swiss researchers, enriched with false positives as noise. Because we disambiguated not only the Italian inventors but the entirety of EPO patents, we were able to follow the guidelines of the APE-INV for these benchmark datasets. Our method achieved a recall rate of 90.98% with a precision of almost 100% (99.9903%).[8] Given these numbers, we are confident to identify inventor mobility in Italy without having concerns related to the disambiguation procedure.

This dataset of Italian inventors allows to identify movement of inventors between regions by observing the patents they filed over time. This paper considers inventors with at least two EPO patent applications and look at the inventors' addresses of these patents. If an inventor, in a given period, has an EPO patent with a certain address and the same inventor, in a later period, appears on an EPO patent with an address in a different region , this paper assumes that this inventor moved from one region to another during the two periods. [9] Since the exact date of an inventor movement cannot be tracked from the patent documents, the inventor flows are computed assuming that the mobile inventors move in the priority year of the patent of the destination region.

The Empirical Analysis

Descriptive evidence

This section displays the main characteristics of the data on TFP and inventor mobility in the twenty Italian regions. Table 1 shows descriptive statistics of the TFP annual growth rates (in percentage values) for the period 1996-2011. The TFP growth rates range from -5.95% (observed for Umbria) to 4.86% (observed for Calabria). The region with the highest average value of TFP growth rates is Basilicata (0.53%); the region with the lowest average value of TFP growth is Molise (-0.46%).

- Table 1 about here -

Table 2 displays the stock of inventors in the year 1995 and the total number of inventor flows for the Italian regions in the period 1995-2010. Column 1 shows the geographical distribution of Italian inventors in 1995, i.e. the stock of Italian inventors with at least one EPO patent application. The total number of Italian inventors in 1995 is 7143. The highest number of inventors (2570) is in

Lombardia and the lowest (5) is in Valle d'Aosta. The other columns show the interregional inventor inflows, outflows and net inflows. For each of these three categories of inventor mobility, the total number of flows (Total) are also splitted in inventor flows between Italian regions (National) and inventor flows between Italian regions and non-Italian regions (International). All of these values are constructed aggregating the annual data on inventor mobility observed during the period 1995-2010. The region with the highest value for total inventor inflows (720) is Lombardia; the region with the lowest value for total inventor inflows is Molise (2). The region with the highest value for total inventor outflow is Lombardia with 695 cases; Molise and Valle d'Aosta are the regions with the lowest value of total inventor outflows with 6 cases each. The region with the highest value for total net inflows is Emilia Romagna with 28 cases (248 cases of inflow and 220 of outflow); the region with the lowest value of total net inflows is Piemonte with a value of -69 (232 cases of inflow and 301 of outflow). Regarding the distinction between national and international flows, it emerges that (on average) two thirds of inventor flows, both inflows and outflows, are represented by inventor mobility within Italy. Considering the international mobility, Tables 3 shows the top 10 countries of origin (destination) of inventor inflows (outflows). USA ranks first in both categories of inventor flows and, with the exception of China ranking tenth as destination country of inventor outflows, the other top 10 countries are European countries.

- Table 2 about here -

-Table 3 about here -

Results

The descriptive statistics and matrix of correlations of the variables used to estimate the determinants of TFP growth rates of Italian regions (equation [1]) are provided in Appendix C in the supplemental online material.

Table 4 shows the results of equation [1] obtained using both OLS FE (models with suffixes **a**) and 2SLS FE estimates (models with suffixes **b**). These analyses are performed using, alternatively, one of the three categories of inventor flows.

In general, the results of OLS and 2SLS FE are very similar. For each of the three 2SLS FE models, a weak identification test is performed by computing the Kleibergen-Paap Wald rk F statistic. The Kleibergen-Paap test coincides with the Angrist and Pischke (2009) test since only [one] endogenous variable is used in 2SLS analysis. The values of these tests are well above the traditional rule of thumb of 10 and the highest critical value (16.38) reported by Stock and Yogo (2005), and, thus, support the relevance of the selected instruments.

As expected, the inflow of inventors (*Inflow_rate*) has a positive effect on regional TFP growth (0.025 in Model 1b). As outlined above, several reasons can explain this result. First, immigrating inventors positively affect innovation capacity of the destination regions increasing the stock of inventors. Second, incoming inventors help destination regions to gain access to different and complementary knowledge. Third, even though the direct beneficiaries of inventor mobility are the local hiring firms, knowledge externalities allow other local actors to benefit from the knowledge embodied in the incoming inventors. In general, this result is in line with the existing literature that support the effectiveness of inventor mobility as channel of knowledge diffusion (Migueléz and Moreno, 2015), but extends this literature by providing a first empirical evidence of the direct contribution of inventor inflows in explaining the changes in the regional TFP.

Table 4 also shows a negative coefficient for outflow of inventors (*Outflow_rate*) (-0.037 in Model 2b). This result suggests that the negative effects associated with inventor outflows, i.e. the reduction in the inventor stock and the weakening in the network ties of inventors within region (brain drain effect), prevails over the positive effects represented by the facilitated access to the knowledge generated outside the region (brain gain effect).

The results also show a positive coefficient for the net inflow of inventors (*Netflow_rate*) (0.048 in Model 1c), which overall confirms the benefits on regional TFP growth associated with the immigration of inventors.

Finally, the effects of the controls are worth mentioning. Lagged TFP level (*lnTFP*) is significant and has a negative sign. Thus, during the period examined there was a process of catching-up, which means that regions with lower levels of TFP per capita, ceteris paribus, show higher TFP growth rates. The results do not show any significant effect of R&D activities (*R&Dpc*) on TFP growth rates. However, there is a positive and significant effect of the variable for patent intensity (*PATrd*). This means that successful R&D activity, i.e. which results in a patent application to the EPO, contributes positively to regional growth. Lastly, population density (*Density*) is positive but not significant.

-Table 4 about here -

"Within applicant" and "between applicants" inventor flows

Are these results driven by "within applicant" or "between applicants" inventor flows? It could make a difference when an inventor just moves between different subsidiaries of the same employer across regions compared to changing both the region and the employer. To address this question the full sample of inventor movements is split into two groups. If the applicant of the origin region's patent is also the assignee of the destination region's patent the inventor movement is considered as "within applicant", i.e. within firm; otherwise it is considered as "between applicants". In case of patents with multiple applicants, inventor movements are considered as "between applicant" when the origin region and destination region patents do not show any applicant in common. The observed percentage of inventors that move "between applicants" at least once is 43.5% (973/2238). A set of 2SLS FE estimates (not shown here but available from the authors upon request) are performed substituting the aggregate mobility indexes with the two mobility indexes computed distinguishing between mobility "within applicant" and mobility "between applicants". It appears that the effect of inventor inflows resulting from "whitin applicant" movements is positive and significant at 10% level, while the effect of inventor inflows resulting from "between applicants" is not significant. Inventor outflows are negative for both types of mobility, but (more interestingly) only the effect of "between applicants" movements is significant. Inventor netflows are positive and significant for both types of movements (at 10% level for "between applicant" movements).

It could be the case that the effects of "between applicants" inventor movements will occur only after few years (e.g. because of different organizational routines between the origin and destination firms). Thus, new 2SLS FE estimates are performed including also the 2 year lagged values of the inventor flows variables. The estimates results (see Table 5) show that, as expected, the coefficient of the 2 year lagged values of inventor inflows resulting from "between applicants" is significantly positive (see Model 4b).

Several robustness checks are performed controlling for (i) the shocks caused by the recent financial crisis, (ii) potential biases in the measurement of the TFP, (iii) the potential effect exterted by human capital (iv) heterogeneous effects related to inventors' inventive performance, (v) uncertainty about the exact date of inventor movements. Overall these robustness checks validate the main results of this paper (for further details, see Appendix D in the online supplemental material).

- Table 5 about here -

Conclusions

This paper combines a new database on geographical mobility of patent inventors with estimates of the regional TFP in Italy in the 1996–2011 period. Using an aggregate production function at the regional level and a set of geography based instruments to address endogeneity, this paper tests the relationship between inventors' geographical mobility and TFP.

Inventors' mobility across regions in Italy is significantly associated with TFP growth. In particular, a 1% increase in the inflow of inventors in a region increases TFP by 2.5%. In parallel a 1% increase in the outflow of inventors in a region decreases TFP by 3.7%. These correlations are robust to including several control variables (such as R&D expenditure per capita, the ratio of patents to R&D expenditure and the population density). The coefficients from the 2SLS estimates imply that the net flow of inventors is correlated with significant productivity gains for the receiving regions and a significant productivity loss for the sending regions.

In addition, the results of this paper show heterogeneous effects for "within applicant", i.e. an inventor moves between subsididiaries of the same employer across regions, and "between applicants" movements, i.e. between different firms. The effects of "between applicants" inventor inflows on TFP take more time to materialize than the "within applicant" inventor inflows. Moreover, the negative effects on TFP by inventor outflows are driven by "between applicants" movements.

The results of this paper highlight that diaspora effects arise also within a country. Comparing to international migrations, inter-regional migrations might be of greater magnitude because of the lower barriers faced and, thus, resulting in a major geographical reallocation of the human capital. This spatial heterogeneous redeployment of human capital might generate different regional growth patterns within the same country depending on the ability of regions to maintain and attract high-skilled people. Cumulative processes might be engendered where the richer regions benefit from the inflow of high skilled people attracted by the higher wages of these regions, which in turn entails a greater demand of high-skilled people and a greater investment in knowledge-intensive activities. Conversely, the poorer regions might enter a vicious circle characterised by outflow of high-skilled people, decrease in demand of high-skilled people and in knowledge-intensive activities (Faggian and McCann, 2009). Our results contribute to shed light on the economic effects of the movement of the regional communities in Italy that transfer their own tacit knowledge and social capital to receiving regions and contribute to task specialization, innovation and, possibly, entrepreneurship. It also raises an important warning flag on the negative effects of the loss human capital in sending regions. In doing so the results of this paper add evidence of the effect of brain drain in Italy where a substantial and increasing portion of geographically mobile workers are tertiary educated (Becker et al. 2004; Fratesi and Percoco, 2014; Marinelli, 2013). The potential detrimental effect of the outflow of skilled workers should be a reason of concern not only for southern region like Puglia and Campania but also northern regions like Liguria and Piemonte.

NOTES

[1] Related to this there is the idea that productivity increases because of improved working conditions. For example, Clark et al. (1998) using the German data of the Socio-Economic Panel show that workers are more likely to leave when they are not satisfied with their jobs.

[2] For an analyisi of the impact of cultural diversity on productivity see, for example, Alesina et al. (2014).

[3] The empirical model of this paper is also in line with the technology gap approach (see e.g.: Fagerberg, 1988), which considers regional economic (or productivity) growth as driven primarily by innovation and takes the distinction between the development of new knowledge in a region and the diffusion of knowledge between regions.

[4] This paper uses a balanced panel dataset. TFP growth data refers to period 1996-2011, while data on the lagged independent variables refer to the period 1995-2010.

[5] Geographical distance, measured in km, is calculated as great circle distance between regions' centroids.

[6] All the variable measured in Euros are expressed at constant prices (reference year: 2005).

[7] Data on total hours worked are not available at regional level.

[8] Precision = true positive / (true positive + false positive); Recall = true positive / (true positive + false negative).

[9] The constructed database contains 52696 inventors, of which 33459 are one-patent inventors. Since this paper considers inventors with a minimum of two patents, the computed inventor mobility measures do not capture potential movements of one-patent inventors. In general, the authors of this paper recognize that the computed inventor mobility indexes underestimate the real inventor flows, and overall the inter-regional flows of high skilled people.

Table Captions

Table 1. Descriptive statistics of TFP annual growth rate (percentage change) - period 1996-2011

Table 2. Stock of inventors in 1995 and total number of inventor flows during the period 1995-2010

Table 3. Inventor inflows and outflows (period 1995-2010): top 10 countries per country of origin and destination

Table 4. Determinants of TFP growth rates - OLS FE and 2SLS FE estimates

Table 5. Determinants of TFP growth rates and distinction between "within applicant" (Models 4a, 5a and 6a) and "between applicants" (Models 4b, 5b and 6b) inventor mobility - 2SLS FE estimates

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(5), 571-591.

Alesina, A., Harnoss, J., and Rapoport, H. (2013). Birthplace Diversity and Economic Prosperity. *National Bureau of Economic Research*, NBER Working Papers No. 18699.

Alma Laurea (2016). XVIII Rapporto Alma Laurea sul Profilo e la Condizione occupazionale dei laureati.https://www.almalaurea.it/sites/almalaurea.it/files/comunicati/2016/cs_almalaurea_indagini_27_4_def.pdf

Almeida, P., and Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, 45(7), 905-917.

Angrist, J., and Pischke, S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Becker, O, Ichino, A., and Peri, G. (2004). How large is the “brain drain” from Italy?. *Giornale degli Economisti e Annali di Economia*. 63 (Anno 117), No. 1 , 1-32.

Bosetti, V., Cattaneo, C., and Verdolini, E. (2015). Migration of skilled workers and innovation: A European Perspective. *Journal of International Economics*, 96(2), 311-322.

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(4), 439-468.

Cappelli, R., and Montobbio, F. (2016). European integration and knowledge flows across European regions. *Regional Studies*, 50(4), 709-727.

Clark, A., Georgellis, Y., and Sanfey, P. (1998). Job satisfaction, wage changes and quits: Evidence from Germany. *Research in Labor Economics*, 17, 95-121

Docquier, F., and Rapoport, H. (2012). Globalization, brain drain, and development. *Journal of Economic Literature*, 50(3), 681–730.

Doherr, T. (2017). Inventor Mobility Index: A Method to Disambiguate Inventor Careers. *ZEW*, Discussion Paper No. 17-018, Mannheim.

Fagerberg, J. (1988). Why growth rates differ. In: Dosi, G., Freeman, C., Nelson, R., Silverberg, G., Soete, L. (Eds.), *Technical change and economic theory*, Pinter Publishers, London.

Faggian, A., and McCann, P. (2009). Human capital and regional development. In R. Capello and P. Nijkamp (ed), *Handbook of regional growth and development theories* . Cheltenham: Edward Elgar.

Faggian, A., Rajbhandari, I., and Dotzel K.R. (2017). The interregional migration of human capital and its regional consequences: a review. *Regional Studies*, 51(1), 128-143.

Fassio, C., Montobbio, F., and Venturini, A., (2015). How do native and migrant workers contribute to innovation? A study on France, Germany and the UK. *IZA, Institute for the Study of Labor,* Discussion paper no. 9062, http://ftp.iza.org/dp7922.pdf

Frankel, J.A., and Romer, D. (1999). Does trade cause growth?. *American Economic Review*, 89 (3), 379–399.

Fratesi, U., and Percoco, M. (2014). Selective Migration, Regional Growth and Convergence: Evidence from Italy. *Regional Studies*, 48(10), 1650-1668, DOI:10.1080/00343404.2013.843162

Glaeser, E. (2010). *Agglomeration Economics*, University of Chicago Press, Chicago.

Hensen, M.M., de Vries, M.R., and Cörvers, F. (2009). The role of geographic mobility in reducing education–job mismatches in the Netherlands. *Papers in Regional Science*, 88(3), 667–682. doi:10.1111/j.1435-5957.2008.00189.x

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

Hunt, J., and Gauthier-Loiselle, M. (2010). How Much Does Immigration Boost Innovation?. *American Economic Journal: Macroeconomics*, 2(2), 31-56.

Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of Political Economy*, 87 (5), 972–990.

Kerr, W.R. (2010). Breakthrough inventions and migrating clusters of innovation. *Journal of Urban Economics*, 67(1), 46–60.

Lissoni, F., Maurino, A.,Pezzoni, M., Tarasconi, G. (2010). APE-INV's "Name Game" Algorithm Challenge: A Guideline for Benchmark Data Analysis & Reporting. *European Science Foundation*, http://www.esf-ape-inv.eu/download/Benchmark_document.pdf.

Maffezzoli, M. (2006). Convergence across Italian regions and the role of technological catch up. *Topics in Macroeconomics*, 6, Article 15.

Marinelli, E. (2013). Sub-national graduate mobility and knowledge flows: An exploratory analysis of onward- and return-migrants in Italy. *Regional Studies*, 47(10), 1618–1633. doi:10.1080/00343404.2012.709608

Miguélez, E. and Moreno, R. (2015). Knowledge flows and the absorptive capacity of regions. *Research Policy*, vol. 44(4), 833-848.

Nathan, M. (2014). The wider economic impacts of high-skilled migrants: a survey of the literature for receiving countries. *IZA Journal of Migration*, 3:4.

Ortega, F., and Peri, G. (2014). Openness and income: The roles of trade and migration. *Journal of International Economics*, 92(2), 231-251.

Peri, G., and Sparber, C. (2009). *Task Specialization, Immigration, and Wages. American Economic Journal: Applied Economics*, 1(3), 135–169.

Peri, G. (2012). The effect of immigration on productivity: evidence from U.S. States. *Review of Economics and Statistics*, 94(1), 48–358.

Rosenkopf, L., and Almeida, P. (2003). Overcoming local search through alliances and mobility. *Management Science*, 49(6), 751-766.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 39, 312–20.

Stock J.H. and Yogo M. (2005). Testing for weak instruments in linear IV regression, In: Stock J.H. and Andrews D.W.K. (Eds.), *Identification and Inference for Econometric Models:Essays in Honor of Thomas J. Rothenberg*. Cambridge University Press.

Trajtenberg, M. (2005). Recombinant ideas: the mobility of inventors and the productivity of research. In: Proceedings of the CEPR- Conference, Munich, May 26–28.

Trajtenberg, M., Shiff, G., and Melamed, R. (2006). The "Names Game": harnessing inventors' patent data for economic research. *National Bureau of Economic Research,* NBER working paper 12479.

Trippl, M. (2013). Scientific mobility and knowledge transfer at the interregional and intraregional level. *Regional Studies*, 47 (10), 1653-1667. DOI:10.1080/00343404.2010.549119.

Table 1. Descriptive statistics of TFP annual growth rate (percentage change) - period 1996-2011

Region	Mean	Std. Dev.	Min	Max	No obs
Abruzzo	0.10	1.55	-2.71	2.75	16
Basilicata	0.53	1.95	-3.39	3.20	16
Calabria	0.28	2.29	-3.99	4.86	16
Campania	0.53	1.24	-2.49	2.39	16
Emilia-Romagna	0.23	2.00	-5.03	3.22	16
Friuli-Venezia Giulia	0.04	2.00	-4.60	3.04	16
Lazio	-0.07	1.27	-2.77	1.92	16
Liguria	0.09	1.72	-4.48	2.94	16
Lombardia	-0.11	1.83	-4.71	4.29	16
Marche	0.07	1.71	-3.78	1.74	16
Molise	-0.46	1.69	-4.85	2.24	16
Piemonte	0.02	1.97	-5.65	3.82	16
Puglia	0.06	1.48	-2.78	2.54	16
Sardegna	-0.15	1.20	-2.95	1.73	16
Sicilia	-0.02	1.31	-3.74	1.99	16
Toscana	0.10	1.39	-3.44	2.08	16
Trentino-Alto Adige	-0.19	1.55	-3.52	2.19	16
Umbria	-0.27	1.80	-5.95	1.94	16
Valle d'Aosta	0.03	2.23	-4.06	3.11	16
Veneto	-0.04	1.79	-4.13	2.44	16
Italy	0.04	1.43	-3.97	2.20	16

Note: the TFP growth rates for Italy are calculated as weighted average of the TFP growth rates of the 20 Italian regions.

Table 2. Stock of inventors in 1995 and total number of inventor flows during the period 1995-2010

Region	STOCK OF INVENTORS IN 1995	INFLOW			OUTFLOW			NETFLOW		
		TOTAL	NATIONAL	INTERNATIONAL	TOTAL	NATIONAL	INTERNATIONAL	TOTAL	NATIONAL	INTERNATIONAL
Abruzzo	81	53	39	14	60	39	21	-7	0	-7
Basilicata	8	5	3	2	8	6	2	-3	-3	0
Calabria	14	10	9	1	9	9	0	1	0	1
Campania	79	46	42	4	68	58	10	-22	-16	-6
Emilia-Romagna	929	248	168	80	220	137	83	28	31	-3
Friuli-Venezia Giulia	226	67	51	16	51	42	9	16	9	7
Lazio	458	210	139	71	176	116	60	34	23	11
Liguria	213	61	44	17	70	59	11	-9	-15	6
Lombardia	2570	720	420	300	695	402	293	25	18	7
Marche	141	46	34	12	40	39	1	6	-5	11
Molise	6	2	1	1	6	5	1	-4	-4	0
Piemonte	992	232	169	63	301	212	89	-69	-43	-26
Puglia	43	29	23	6	47	39	8	-18	-16	-2
Sardegna	34	19	16	3	24	20	4	-5	-4	-1
Sicilia	95	35	26	9	39	28	11	-4	-2	-2
Toscana	376	173	120	53	155	111	44	18	9	9
Trentino-Alto Adige	81	40	24	16	38	19	19	2	5	-3
Umbria	49	29	21	8	31	26	5	-2	-5	3
Valle d'Aosta	5	14	14	0	6	6	0	8	8	0
Veneto	743	199	141	58	200	131	69	-1	10	-11
Italy	7143	2238	1504	734	2244	1504	740	-6	0	-6

Table 3. Inventor inflows and outflows (period 1995-2010): top 10 countries per country of origin and destination

Inventor inflows				Inventor outflows			
Country of origin	Number	Percentage on total		Country of destination	Number	Percentage on total	
		Including Italy	Excluding Italy			Including Italy	Excluding Italy
USA	198	8.85	26.98	USA	209	9.31	28.24
Germany	118	5.27	16.08	Germany	111	4.95	15.00
France	95	4.24	12.94	Switzerland	85	3.79	11.49
United Kingdom	81	3.62	11.04	France	80	3.57	10.81
Switzerland	61	2.73	8.31	United Kingdom	57	2.54	7.70
Sweden	30	1.34	4.09	Netherlands	30	1.34	4.05
Netherlands	29	1.30	3.95	Belgium	26	1.16	3.51
Belgium	22	0.98	3.00	Sweden	25	1.11	3.38
Spain	19	0.85	2.59	Spain	21	0.94	2.84
Austria	7	0.31	0.95	China	10	0.45	1.35

Note: The percentage values under the column Including Italy (Excluding Italy) are calculated including (excluding) inventor mobility within Italy.

Table 4. Determinants of TFP growth rates - OLS FE and 2SLS FE estimates

VARIABLE	INFLOW		OUTFLOW		NETFLOW	
	Model 1a OLS FE	Model 1b 2SLS FE	Model 2a OLS FE	Model 2b 2SLS FE	Model 3a OLS FE	Model 3b 2SLS FE
$\log(TFP_{i,t-1})$	-0.275***	-0.274***	-0.316***	-0.316***	-0.298***	-0.298***
	(0.075)	(0.070)	(0.077)	(0.072)	(0.069)	(0.065)
$R\&Dpc_{i,t-1}$	-0.001	-0.002	0.004	0.004	-0.004	-0.004
	(0.018)	(0.017)	(0.018)	(0.017)	(0.018)	(0.017)
$PATrd_{i,t-1}$	0.014**	0.014**	0.017**	0.017**	0.014*	0.014**
	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)
$Density_{i,t-1}$	0.179	0.179	0.170	0.170	0.168	0.168
	(0.210)	(0.198)	(0.220)	(0.207)	(0.202)	(0.190)
$Inflow\_rate_{i,t-1}$	0.024*	0.025**				
	(0.013)	(0.012)				
$Outflow\_rate_{i,t-1}$			-0.038**	-0.037***		
			(0.014)	(0.014)		
$Netflow\_rate_{i,t-1}$					0.048***	0.048***
					(0.010)	(0.009)
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Instrument		Pred. Inflow rate		Pred. Outflow rate		Pred. Net inflow rate
Observations	320	320	320	320	320	320
Number of regions	20	20	20	20	20	20
R-squared	0.68	0.68	0.68	0.68	0.69	0.69
Log Likelihood	1038.42	1038.42	1039.95	1039.95	1043.13	1043.13

Notes: for the sake of clarity the results of the first stage estimates in 2SLS models are not reported here; the constant term (i.e. the average value of the fixed effects) is not reported here; cluster standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Determinants of TFP growth rates and distinction between "within applicant" (Models 4a, 5a and 6a) and "between applicants" (Models 4b, 5b and 6b) inventor mobility - 2SLS FE estimates

VARIABLE	INFLOW		OUTFLOW		NETFLOW	
	Model 4a	Model 4b	Model 5a	Model 5b	Model 6a	Model 6b
log(TFP $_{i,t-1}$)	-0.275***	-0.282***	-0.291***	-0.317***	-0.288***	-0.304***
	(0.071)	(0.072)	(0.072)	(0.067)	(0.071)	(0.063)
R&Dpc$_{i,t-1}$	-0.001	-0.002	0.004	0.003	0.001	-0.006
	(0.015)	(0.019)	(0.016)	(0.017)	(0.016)	(0.021)
PATrd$_{i,t-1}$	0.015***	0.015**	0.015**	0.015***	0.016***	0.014**
	(0.005)	(0.007)	(0.007)	(0.006)	(0.006)	(0.007)
Density$_{i,t-1}$	0.175	0.169	0.171	0.149	0.183	0.176
	(0.200)	(0.191)	(0.207)	(0.185)	(0.205)	(0.188)
Inflow_rate within $_{i,t-1}$	0.041*					
	(0.023)					
Inflow_rate within $_{i,t-2}$	0.032					
	(0.022)					
Inflow_rate between$_{i,t-1}$		0.015				
		(0.043)				
Inflow_rate between$_{i,t-2}$		0.046***				
		(0.010)				
Outflow_rate within $_{i,t-1}$			-0.022			
			(0.021)			
Outflow_rate within $_{i,t-2}$			0.017			
			(0.011)			
Outflow_rate between$_{i,t-1}$				-0.076***		
				(0.024)		
Outflow_rate between$_{i,t-2}$				0.031		
				(0.063)		
Netflow_rate within$_{i,t-1}$					0.062***	
					(0.011)	
Netflow_rate within$_{i,t-2}$					0.002	
					(0.012)	
Netflow_rate between $_{i,t-1}$						0.047*
						(0.025)
Netflow_rate between $_{i,t-2}$						0.039**
						(0.020)
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Instrument	Pred. Inflow rate	Pred. Inflow rate	Pred. Outflow rate	Pred. Outflow rate	Pred. Net inflow rate	Pred. Net inflow rate
Observations	320	320	320	320	320	320

Number of regions	20	20	20	20	20	20
R-squared	0.68	0.68	0.68	0.68	0.69	0.69
Log Likelihood	1038.99	1038.37	1038.39	1036.42	1040.58	1039.95

Notes: the suffix “within” (“between”) refers to “within applicant” (“between applicants”) movements; for the sake of clarity the results of the first stage estimates in 2SLS models are not reported here; the constant terms (i.e. the average values of fixed effects) are not reported here; cluster standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Appendix

Appendix A. TFP of Italian regions

To calculate the TFP of Italian regions for the period 1996-2011, this paper follows a standard growth accounting approach (Jorgenson, 1995; OECD, 2011; Solow, 1957). The starting point is a Cobb-Douglas production function with constant return to scale, defined as follows:

$$GDP_{i.t} = TFP_{i,t} \; Capital_{i,t}^{1-\beta} \; Labor_{i,t}^{\beta}$$

where $GDP_{i.t}$ is the GDP of region *i* at time *t*. Capital (*Capital*) and labor (Labor) are the two input factors considered, and *1-*β and β are, respectively, the GDP elasticity of capital and the GDP elasticity of labor. It follows the TFP growth rates ($ln\left(TFP_{i,t}/ \, TFP_{i,t-1}\right)$) and the annual TFP levels ($TFP_{i,t}$) of regions can be computed through the following equations:

[A1] $ln\left(TFP_{i,t}/ \, TFP_{i,t-1}\right) = \; ln\left(GDP_{i,t}/ \, GDP_{i,t-1}\right) - \left(1 - \bar{\beta}\right) \, ln\left(Capital_{i,t}/ \, Capital_{i,t-1}\right) - \bar{\beta} \; ln\left(Labor_{i,t}/ \, Labor_{i,t-1}\right)$

[A2] $TFP_{i,t} = \; GDP_{i,t}/\left(Capital_{i,t}^{1-\beta} \; Labor_{i,t}^{\beta}\right)$

A basic problem on the calculation of equations [A1] and [A2] arises because of the lack of data on regional capital stock (*Capital*). At regional level, ISTAT provides only data on regional fixed investments for the period 1995-2011. Unfortunately, the length of these data on regional investments does not allow to independently construct an accurate estimate of the gross and net capital stocks. So, first this paper uses these short regional time series on fixed investment to build a partial approximation of the regional capital stock using the perpetual inventory method. Secondly, this paper uses this partial approximation to build regional shares. Third, this paper applies the regional shares to the complete data available at the national level (see Maffezzoli, 2006; Quatraro, 2009).

The first step is to calculate the partial initial stock of capital at the regional level in the following way:

$$PC_{i,1995} = \frac{I_{i,1995}}{g_i + \delta}$$

where I_i is the real investment in 1995, g_i is the average growth rate of fixed investments in Italian regions and δ is the depreciation rate. The authors of this paper use the average depreciation rate of national capital stock (δ) because data on depreciation rates at regional level are not available.So, this paper assumes a constant depreciation rate across all regions equal to 0.0048.

Then, the authors of this paper calculate the regional partial capital stock for the period 1996-2011:

$$PC_{i,t} = PC_{i,t-1} * (1 - \delta_t) + I_{i,t}$$

Thus, for each region *i*, this papercalculates the share of the partial capital stock in 2011:

$$S_{i,2011} = \frac{PC_{i,2011}}{\sum_i PC_{i,2011}}$$

The authors of this paper use these shares to allocate the net national capital stock NC (provided by ISTAT) in 2011 to the 20 regions such that:

$$C_{i,2011} = S_{i,2011} * NC_{2011}$$

Finally, the authors of this paper follow the procedure outlined in the literature (Maffezzoli, 2006; Quatraro, 2009) to extend the series before 2011:

$$C_{i,t-1} = \frac{C_{i,t} - I_{i,t}}{1 - \delta_t}$$

Appendix B. Instrumental variables in 2SLS fixed-effects estimates

The instrumental variables are constructed adopting the gravity model approach. Gravity model estimates are not performed for interregional net inflows because of inappropriateness of the gravity-type variables. Variables like geographical distance and territorial borders represent migration costs which affect bilateral inventor inflows and inventor outflows in the same way. Thus, these variables are not useful in explaining the bilateral net effects resulting from inflows and outflows.

As first step in building our instruments, the author of this paper perform, both for inventor inflows and outflows, $N=16$ poisson pseudo maximum likelihood (PPML) (Santos Silva and Tenreyro, 2006) separate annual cross-sectional estimates, i.e. one for each year t covered by our data ($N = \sum_t s_t$, where $s_t = 1$ if year t is used in the estimation and $s_t = 0$ otherwise), using the following gravity model equation:

$$\text{[B1]} \quad C_{ij} = \exp\left[\alpha + \gamma \ln\left(\text{dist}_{ij}\right) + \Omega\, \text{Italy}_{ij} + \varphi\, \text{Border}_{ij} + \varrho_i + \eta_j\right] \varepsilon_{ij}$$

where C_{ij} is the variable capturing inventor flows between region i and j (in our case measured by the inventor inflow or outflow rates). As geographic variables we use the logarithm of geographical distance between the two areas ($dist_{ij}$), a dummy set equal to 1 if the two regions are neighbours ($Border_{ij}$) and a dummy set equal to 1 if the two regions belong to Italy ($Italy_{ij}$). Regions specific effect, both for region i and region j (denoted ϱ_i and η_j), are included to take into account of region-specific unobservable effects and to correct for cross-sectional bias (Anderson and van Wincoop, 2003; Baldwin and Taglioni, 2006).

The gravity model estimates (see equation [B1]) are performed aggregating the inventor inflows (outflows) at the level of country of origin (destination) when the origin (destination) region is not an Italian region. As a consequence, C_{ij} represents the observed inventor inflow (outflow) rates from (to) areas j to (from) areas i where destination (origin) areas i are represented by the 20 Italian regions, and origin (destination) areas j are represented by the 20 Italian regions and 52 countries for which at least one inventor inflow (outflow) is observed over the 16 years covered by the paper's analysis. The independent variables used in gravity models are also constructed taking into account the two types of geographical levels considered in our sample, i.e. regional or country level. Thus, for instance, the geographical coordinates of the centroids of Lombardia and Germany are used to measure physical distance between the two geo-areas.

Each of the 16 annual cross-sectional regressions uses1420 observations, i.e. 20 areas *i* * 71 areas *j*. Areas *i* are the 20 Italian regions; areas *j* are: 19 regions of Italy (19 instead of 20 because we do not consider intra-regional flows) and the other 52 countries. The results of these estimates are not shown here but are available from the authors upon request.

As second and final step, for each region *i* we aggregate the predicted bilateral inventor flow rates involving the region as follows:

$$[B2]\ \widehat{C}_i = \sum_{i \neq j} \exp\left[\widehat{\alpha} + \widehat{\gamma}\ \ln\left(dist_{ij}\right) + \widehat{\Omega}\ Italy_{ij} + \widehat{\rho}\ Border_{ij} + \widehat{\varrho}_i + \widehat{\eta}_j\right]$$

After having performed the descripted procedure, we obtain *N=16* predicted values (i.e. *N* values of $\widehat{C}_i$) from estimates that use inventor inflow rates as dependent variable and other *N* values from estimates that use inventor outflows as dependent variable. A variable containing the former values is used to instrument the observed inventor inflow rates, while a variable containing the latter values are used to instrument the observed inventor outflow rates. A third variable constructed as difference between the two instrumental variables is used to instrument the observed inventor net inflow rates.

Appendix C. Descriptive statistics and matrix of correlations

Table C1. Descriptive statistics

Variables	Description	Mean	Std. Dev.	Min	Max
$\log(TFP_{i,t} / TFP_{i,t-1})$	Regional TFP growth rate	0.0004	0.017	-0.060	0.049
$\log(TFP_{i,t-1})$	Regional TFP level (log)	1.895	0.085	1.663	2.080
$Inflow\_rate_{i,t-1}$	Interregional inventor inflow rate	0.020	0.037	0	0.500
$Inflow\_rate\ within_{i,t-1}$	"Whitin applicant" interregional inventor inflow rate	0.011	0.023	0	0.250
$Inflow\_rate\ between_{i,t-1}$	"Between applicants" interregional inventor inflow rate	0.009	0.021	0	0.250
$Outflow\_rate_{i,t-1}$	Interregional inventor outflow rate	0.022	0.038	0	0.333
$Outflow\_rate\ within_{i,t-1}$	"Whitin applicant" interregional inventor outflow rate	0.013	0.029	0	0.250
$Outflow\_rate\ between_{i,t-1}$	"Between applicants" interregional inventor outflow rate	0.009	0.026	0	0.333
$Netflow\_rate_{i,t-1}$	Interregional net inflow rate	-0.002	0.041	-0.333	0.250
$Netflow\_rate\ within_{i,t-1}$	"Whitin applicant" interregional inventor net inflow rate	-0.003	0.026	-0.250	0.083
$Netflow\_rate\ between_{i,t-1}$	"Between applicants" interregional inventor net inflow rate	0.0001	0.032	-0.333	0.250
$R\&Dpc_{i,t-1}$	Regional R&D expenditure per capita (Thousands of Euros)	0.217	0.133	0.018	0.553
$PATrd_{i,t-1}$	Region's number of patents on R&D expenditure (Millions of Euros)	0.236	0.105	0	1.385
$Density_{i,t-1}$	Population density of region (people per square KM)	0.176	0.105	0.035	0.424

Table C2. Matrix of correlations

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	$\log(TFP_{i,t} / TFP_{i,t-1})$	1.000													
2	$\log(TFP_{i,t-1})$	-0.108	1.000												
3	$Inflow\_rate_{i,t-1}$	0.033	-0.095	1.000											
4	Inflow_rate $within_{i,t-1}$	0.036	-0.044	0.858	1.000										
5	Inflow_rate $between_{i,t-1}$	0.020	-0.119	0.838	0.438	1.000									
6	$Outflow\_rate_{i,t-1}$	-0.043	-0.177	0.412	0.373	0.324	1.000								
7	Outflow_rate $within_{i,t-1}$	-0.046	-0.098	0.517	0.536	0.335	0.726	1.000							
8	Outflow_rate $between_{i,t-1}$	-0.011	-0.147	0.019	-0.058	0.095	0.639	-0.065	1.000						
9	$Netflow\_rate_{i,t-1}$	0.070	0.078	0.531	0.436	0.465	-0.554	-0.203	-0.577	1.000					
10	Netflow_rate $within_{i,t-1}$	0.084	0.073	0.166	0.270	0.004	-0.499	-0.668	0.023	0.616	1.000				
11	Netflow_rate $between_{i,t-1}$	0.022	0.040	0.541	0.338	0.587	-0.305	0.275	-0.751	0.778	-0.016	1.000			
12	$R\&Dpc_{i,t-1}$	-0.081	0.670	-0.068	-0.031	-0.086	-0.165	-0.122	-0.103	0.091	0.112	0.027	1.000		
13	$PATrd_{i,t-1}$	0.009	-0.051	0.063	0.015	0.095	-0.127	-0.046	-0.133	0.175	0.065	0.171	0.055	1.000	
14	$Density_{i,t-1}$	0.034	0.346	-0.127	-0.058	-0.160	-0.142	-0.110	-0.082	0.016	0.074	-0.040	0.434	-0.100	1.000

Appendix D. Robustness checks

Various checks are conducted to validate the robustness of the main results of this paper. The results of these robustness checks, not reported here, are available upon request from the authors.

The data used in the analysis include the recent financial crisis period. The shocks caused by the crisis might affect both the inventor flows and the TFP of Italian regions. To exclude the possibility that the relationships between inventor flows and TFP are driven by these shocks, new estimates are performed excluding the period 2009-2011. The obtained results are very similar to those in the main text.

To measure the regions' TFP growth rates the authors of this paper rely on an estimated measure of the regions' capital stocks. To control for potential biases due to possible errors in the measurement of the capital stocks and, thus, of the TFP growth rates, new estimates are performed using as dependent variable the regions' labor productivity growth rates (labor productivity is measured as the ratio between the GDP and the number of full time equivalent workers). The traditional high correlation between TFP and labor productivity is in support of this strategy. The estimates results are very similar to those discussed in the main text.

This paper considers also the potential effect of the human capital. ISTAT provides data on the number of graduates in Science & Technology (S&T) for the period 1998-2011. Using these data, a variable measuring human capital (High skilled) is computed as the ratio between the number of S&T graduates and the total population. The correlation matrix shows that, as expected, High skilled is highly correlated with other control variables already included in the model, i.e. R&D per capita (0.64) and TFP level (0.48). Thus, the authors of this paper argue that R&D per capita and TFP level capture most of the regional differences in human capital. Moreover, the results of 2SLSE FE estimates that include High skilled as additional control variable are similar to those reported in the main text of the paper, and also the coefficient values of High skilled are not statistically significant.

In order to check for heterogeneous effects related to inventors' inventive performance, additional estimates are performed based on a recalculated mobility index using the past inventive productivity of each inventor as a weight. In particular, this paper adopts two alternative measures of inventor productivity, i.e. the inventor's depreciated patent stock and the average number of forward citations associated with the inventor's patents. The former are calculated using the perpetual inventory method with a depreciation rate of 15%. The latter calculated using temporal windows of 3 years to control for the well-known truncation bias problem (Hall et al., 2005). The results are similar to those discussed in the main text in terms of significance levels, but the coefficient values of all the mobility indexes are lower. These results can be explained by a demographic selection. Mobile inventors are, in general, younger than non-mobile inventors and, consequently, the past productivity of the mobile inventors is lower than the past productivity of non-mobile inventors.

Another issue is whether the uncertainty about the exact date of inventor movement affects the estimation results. The data show that the mean value of the temporal lags between the origin region's patent and the destination region's patent is 2.4 years and that in 53% of the cases of inventor movements the observed temporal lag is of 1 year or less. Moreover, the percentage of inventor movements increases to 77% if we consider a temporal lag of 3 years or less. These figures show that in most cases the exact move date is not too uncertain. In addition, as a robustness check, we perform new estimates excluding cases of inventor movements for which the temporal lag between the origin and destination patents is bigger than a temporal threshold value. Using both a temporal value of 5 years and 3 years the obtained results are very similar to those discussed in the main text, except that inventor inflows are significant at 10% level instead of 5% level when we use a temporal threshold value of 3 years. Temporary movements, i.e. cases where the inventor moves from region *i* to region *j* and then back to region *i* are observed for 207 inventors (8.93% of the 2318 mobile inventors). Our inventor flow indexes are constructed considering only the last movement. So, for example, if an inventor moves from Marche to Lombardy both in 1995 and in 2002, it is assumed that this inventor moves from Marche to Lombardy only in 2002. However, since the bias of the effect of temporary movements is only partially mitigated by the correction adopted to control for double movements, 2SLS FE estimates are performed excluding all cases of temporary movements. The estimates results are similar to those reported in the main text, except for the inventor inflows that are significant only for the 2 year lagged values. Moreover, inventor outflows have a stronger effect (-0.086 vs. -0.037).

A question related to the uncertainty about the exact move date is that the mobility variable might capture different lags of the R&D or patent variables. As a further robustness check, two separate set of 2SLS FE estimates are performed including additional control variables: the first one including the lagged values from t-4 to t-2 of both R&D per capita and patent per R&D expenditure; and the other one; the second one including the agged values from t-3 to t-1 of the annual rates of growth of both R&D per capita and patent per R&D expenditure. The results are similar to those reported in the main text of the paper.

References

Anderson, J.E., and van Wincoop, E. (2003). Gravity with gravitas: a solution to the border puzzle. American Economic Review, 93(1), 170-192.

Baldwin, R.E., and Taglioni, D. (2006). Gravity for dummies and dummies for gravity equations. National Bureau of Economic Research. NBER Working Paper 12516.

Hall, B.H., Jaffe, A., and Trajtenberg, M. (2005). Market Value And Patent Citations. Rand Journal of Economics,36, 16-38.

Jorgenson D. W. (1995) *Productivity Volume 1: Post-war US Economic Growth*, Cambridge, MA, MIT Press.

Maffezzoli M. (2006) *Convergence across Italian regions and the role of technological catch up*, Topics in Macroeconomics, 6, Article 15.

OECD (2001) *Measuring Productivity. Measurement of Aggregate and Industry-level Productivity Growth*, Paris, OECD.

Quatraro F. (2009) *Innovation, Structural Change and Productivity Growth: Evidence from Italian Regions, 1980-2003, Cambridge Journal of Economics*, 33(5), 1001-1022.

Santos Silva, J.M.C, and Tenreyro, S. (2006). The log of gravity. The Review of Economics and Statistics, 88(4),641-658.

Solow R. M. (1957) Technical *change and the aggregate production function*, The Review of Economics and Statistics, vol. 39, 312–20.