The spatial sorting and matching of skills and firms

Mion, Giordano and Naticchioni, Paolo

Université Catholique de Louvain

February 2007

The Spatial Sorting and Matching of Skills and Firms*

Giordano Mion[†]and Paolo Naticchioni[‡]

October 2006

CORE DP 2006/99

Abstract

Using a matched employer-employee database for Italy we look at the spatial distribution of wages across provinces. This rich database allows us to contribute at opening the black box of agglomeration economies exploiting the micro dimension of the interaction among economic agents, both individuals and firms. We provide evidence that firm size and particularly skills are sorted across space, and explain a large portion of the spatial wage variation that could otherwise be attributed to aggregate proxies of agglomeration externalities. Our data further support the assortative matching hypothesis, that is"good" workers match on the labor market with "good" firms, and we further show that assortative matching is not driven by a co-location of workers and firms of similar quality. Finally, we point out that this assortative matching is negatively related to local market size.

Keywords: Spatial Externalities, Panel-Data, Skills, Firms' Heterogeneity, Sorting, Matching.

JEL Codes: J31, J61, R23, R30.

*We would like to thank the research partnership between ISFOL - Area Mercato del Lavoro (Rome) and Dipartimento di Scienze Economiche - University of Rome "La Sapienza" for the access to the INPS Italian database. We are also grateful to Badi Baltagi, Luisito Bertinelli, Pierre-Philippe Combes, Bart Cockx, Gilles Duranton, Laurent Gobillon, Keith Head, Claudio Lucifora, David Margolis, Steven Redding, Eric Strobl, Bruno Van der Linden, and CEPR, IWSES and AIEL Conference participants for helpful comments and suggestions, and to Antonietta Mundo, Renato Santelia for providing us with data on spatial variables.

[†]Giordano Mion: CORE, Université Catholique de Louvain-la-Neuve, and FNRS. Address: CORE-UCL, 34 Voie du Roman Pays, B-1348 LLN, Belgium. Email: mion@core.ucl.ac.be.

[‡]Paolo Naticchioni, Università di Roma "La Sapienza", and Université Catholique de Louvain-la-Neuve. Email: paolo.naticchioni@uniroma1.it

1 Introduction

Imbalances in terms of wages, GDP per capita, growth, and labor markets' outcomes are pervasive features of the economic landscape. Spatial disparities are in fact large in both developed and developing countries attracting a lot of political concern and, in the case of EU, they are so strategically important to be ranked at the top of the political agenda.[1] As for wages, Glaeser and Mare (2001) find that they are 33% higher is US cities compared to outside metropolitan areas. Data evidence on EU as a whole is less systematic. However, a number of country-based studies, like for instance Combes, Duranton and Gobillon (2004), show that wages vary considerably across space.

There is also a large body of empirical literature concerned with measurements of growth and productivity advantages stemming from agglomeration economies such as urbanization and localization externalities, and market potential. However, despite some relevant exceptions,[2] most frameworks deal with aggregate proxies for agglomeration, neglecting the very nature of these externalities which comes from economic agent -individuals and firms- interactions.

The fact that "good" workers and firms are sorted across space is the flip side of agglomeration economies. In this paper, we explicitly deal with two agents' characteristics, namely workers' skills and firms' size. In particular, we estimate a Mincerian equation using a matched panel employer-employee database for Italy and we consider as skills all the time-invariant attributes of a worker that add to his productivity, using individual fixed effects. Exploiting the information of individual data is crucial in order to assess the importance of skills and firm size in explaining standard aggregate measures of agglomeration economies as well as to get some intuition about the underlying mechanisms leading to such sorting. In other words, we aim at contributing to open the black box of agglomeration externalities.

From the view point of theory there at least three explanations for the positive sorting of worker skills across space. First, cities provide valuable consumption amenities for skilled workers, such as theaters, museums, cultural activities, etc. Second, as suggested by Moretti (2004), bigger cities offer higher returns to education (private plus social) fostering the investments in human capital. Third, in a dynamic perspective cities provide a better environment for the accumulation of human capital thanks to more effective face-to-face interactions like in Glaeser and Mare (2001).

As for the sorting of firms, labor market matching models like Kim (1989) and Helsley and Strange (1990) are consistent with bigger and more productive firms being disproportionately located in large

[1] The reduction of income disparities among EU regions involve much of the political debate with Structural and Cohesion Funds, both aiming at the reduction of imbalances, correspond to approximately one third of the EU budget in the period 1994-1999.

[2] See Glaeser and Mare (2001), Henderson (2003), and Combes, Duranton, and Gobillon(2004).

cities. The size of the local labor market improves the quality of matching between firms and workers boosting productivity. A more recent strand of literature, started with the model of Melitz (2003), takes explicitly into account individual firms' heterogeneity in productivity. In these models, firms differ exogenously in their productivity but they can still coexist on the same market because they sell differentiated varieties. The size of the market and the degree of product substitutability determine the minimum efficiency threshold needed to break even and, as the market gets larger, competition drives less productive firms to exit.

Our individual data on wages supports the existence of sorting, for both individuals and firms, and this is robust even after controlling for endogeneity in location choice. In particular, we show that nearly 75% of raw wage variation across Italian provinces is explained by differences in individual skills. Furthermore, we provide evidence that sorting is essentially "static", i.e. due to non migrants. On the contrary, the sorting of firms accounts only for a small fraction of wage variability. These results suggest that our proxies for individual and firm heterogeneity are important, and account for a large portion of agglomeration externalities as measured by standard aggregate indicators.

Another interesting dimension we can exploit in our data is the connection between employer and employee information, testing the existence of a positive correlation between individual skills and firm size (the so called assortative matching). The evidence in the literature concerning assortative matching is mixed. In our data, we do find a positive correlation ("good" workers matching with "good" firms) and we further show that results are not driven by co-location issues. To the best of our knowledge, this is the first analysis of the spatial dimension of assortative matching using an employer-employee database.

Quite interestingly, we also break down this result across provinces, showing that the degree of assortative matching is negatively related to local market size. This result is robust to the use of a more accurate measurement of general human capital as well as to the inclusion of a province specific firm size premium. This original result contrasts with Wheeler (2001), while being in line with Kim (1989). As long as firm-worker specific attributes are taken into account, then one can argue that the larger the size of a market the larger is the possibility of a worker/firm to find a partner with some specific characteristics which are valuable assets for the matching. Therefore, general attributes should became relatively less important the larger is the size of the market.

The rest of the paper is organized as follows. In Section 2 we present the theoretical background of our analysis. In Section 3 we present the data, the spatial variable definitions as well as some descriptive statistics, while in Section 4 we point out what a matched employer employee database can tell us more about spatial imbalances. Section 5 is devoted to the econometric results, in which we also deepen the endogeneity issue. Section 6 further discusses the evidence of sorting and matching coming out from our estimates. Finally, conclusions are reported in Section 7.

2 Theoretical Background

2.1 Agglomeration Economies

The term "agglomeration economies" refers to those externalities stemming from the interaction of agents across space that positively affects local productivity and growth. There is a large body of literature[3], both theoretical and applied, that has identified several mechanisms leading to the emergence of agglomeration economies. However, despite the large number of empirical studies, data limitations and the lack of a unified underlying theoretical framework make it difficult to disentangle the relative importance of different microfoundations that are, to a large extent, observationally equivalent. One of the goal of this paper is to contribute to this debate on the microfoundations of agglomeration economies by looking at the relative importance of individual versus firms' characteristics. Models of agglomeration externalities concern firms and workers, implicitly assuming that one or both types of agents are more productive where the externality is at work (spatial sorting of workers and/or firms). However, most of the underlying theoretical and empirical studies are about firm productivity and location and despite some relevant exceptions, like Glaeser and Mare (2001), Henderson (2003), and Combes, Duranton, and Gobillon (2004), the empirical analysis has focused on aggregate rather than individual data. Our data allow us to match individual workers and firms' characteristics and can help to get a deeper and more accurate understanding of agglomeration economies.

In this paper, we will focus on the most widely investigated agglomeration economies: urbanization externalities, market potential and localization externalities.

Urbanization externalities represent one of the most important local interaction type related to the size of a market. The idea that market size has a positive impact on local productivity goes back to Marshall (1890) and have been formalized by Abdel Rahman, and Fujita (1990) among others. The urban literature has identified various mechanisms leading economic density to foster growth and productivity like knowledge cross-fertilization, increasing returns to scale in a non-tradable intermediate goods sector, matching of differentiated skills, etc.[4] Ciccone and Hall (1996) and Combes (2000) provide evidence in favor of the positive role of local employment density (a proxy for market size) on productivity and employment growth.

The second spatial externality investigated is the one typically arising in models mixing increasing

[3]See for instance the recent reviews of Duranton and Puga (2004) and Rosenthal and Strange (2004).
[4]See Duranton and Puga (2004) for a discussion of the microfoundations of urbanization economies.

returns and transportation costs where the reduced form of agglomeration incentives takes the form of a Market potential function. Loosely speaking, Market potential can be defined as the demand that a firm located in j can have access to in a world where trade is costly. The Market potential concept has been first put forward by Harris (1954) and then formalized by the so called new economic geography literature[5]. The basic mechanism underpinning agglomeration externalities in these models is the trade-off between the need to be close to consumers to reduce transportation costs (dispersion force), and the desire to concentrate production in order to exploit increasing returns to scale (agglomeration force). If trade costs are sufficiently low, then agglomeration dominates and firms massively relocate in areas providing better access to consumers' demand so boosting local factor prices and in particular nominal wages. Mion (2004) and Hanson (2005) provide evidence that local nominal wages are positively related to market potential, the latter being measured either "structurally" or with the simple average of disposable income coming from surrounding locations as originally proposed by Harris (1954).

Finally, localization externalities refer to the advantage stemming from the concentration of a specific industry in a given location (local specialization). The idea that local specialization fosters growth and productivity also goes back to Marshall (1890) that already identifies three channels (labor market pooling, input sharing and knowledge spillovers) through which localization externalities take place. The models of Henderson (1974) and Duranton and Puga (2004) among others provide a micro-foundation for these externalities. On the other hand, the empirical works of Glaeser et al (1992) and Henderson (2003), which use a measure of local sectoral specialization to proxy for such externalities, suggest that localization externalities play an important role in local growth and productivity. Moreover, Rosenthal and Strange (2001) provide some evidence on the relative importance of labor market pooling, input sharing and knowledge spillovers.

2.2 The Spatial Sorting of Workers' Skills and Firms' Size

The fact that "good" workers and firms are sorted across space is the flip side of agglomeration economies. In this paper, we explicitly deal with two agents' characteristics, namely workers' skills and firms' size. One goal of the paper is to quantify the importance of skills and firm size in explaining standard aggregate measures of agglomeration economies as well as to get some intuition about the underlying mechanisms leading to such sorting.

The sorting of workers' skills

[5] See for instance Fujita, Krugman and Venable (1999) or Fujita and Thisse (2002).

The term skills refers here to workers' attributes like ability and education that are in principle time invariant and increase a worker productivity irrespectively of the characteristics of the employer. The idea that skilled workers are sorted across space is, although less explored than other factors, not new in the urban literature and, as previously argued, it might be a crucial element in getting a better understanding of the microfoundations of spatial externalities.

The oldest explanation is the so-called "bright lights" hypothesis. Skilled workers are known to be more mobile. Therefore, as long as extensive consumption amenities and services provided in cities are particularly valuable assets for them, they would prefer cities as the destination of their migrations. Tabuchi and Yoshida (2000) find that the elasticity of "real" wage with respect to city size is negative suggesting that people living in cities do accept a lower real wage to enjoy consumption benefits.

Another possibility, which does not necessarily involves migrations, is that cities provide higher returns to educations so that residents find it profitable to invest more in the accumulation of human capital that would then foster local productivity and growth. Moretti (2004) actually finds some evidence of a positive relation between city density and returns to education (private *plus* social), while Borjas, Bronas and Trejo (1992) find that the skills composition and the size of internal US migrations are well explained by interstate differences in the returns to education which are pretty heterogeneous across space.

On the other hand, Glaeser and Mare (2001) discuss a third dynamic mechanisms. Large cities can be in fact places where the accumulation of human capital can be faster than anywhere else due to the intense and stimulating face to face interactions. The authors further provide evidence in favor of their conjecture by looking at whether individuals moving from a small (large) to a large (small) city experience a substantial wage increase (decrease). In particular they find that only workers moving to a large city get a substantial wage change and that this increase occurs only few years after moving as if they had to accumulate human capital locally before their productivity (and wage) could raise. On the other side workers moving to a small city earn virtually the same wage suggesting that the human capital they had previously accumulated in large cities does not depreciate after the migration, and is a valuable asset also in their new job.

Finally, the link between human capital accumulation and space finds some support even in the more recent new economic geography literature. Redding and Schott (2003) develop a model where if skill-intensive sectors have higher trade costs, more pervasive input-output linkages or stronger increasing returns to scale, then remoteness depresses the skill premium and therefore incentives for human capital accumulation. The authors further provide evidence that countries with a lower market potential display lower levels of educational attainments.

The sorting of firms' size

There are many theoretical frameworks where agglomeration economies translate into higher firms' productivity, with the latter being usually positively related to the size of the firm. Parallel to the case of workers' skills, the link between space, productivity and firm size could be a powerful insight to open the black-box of agglomeration economies.

Bresnahan and Reiss (1991), when looking at the impact of market size on firms' entry decisions and markups, show that markups should be lower the larger the local market. In this framework, firm size should increase with local demand for firms to break even. The positive correlation between firm size and economic density has been documented in Campbell and Hopenhayn (2005).

Labor market matching models like Kim (1989) and Helsley and Strange (1990) are further consistent with bigger and more productive firms being disproportionately located in large cities. The size of the local labor market improves the quality of matching between firms and workers boosting productivity. However, when market size expands, fiercer labor market competition make it feasible only for a less than proportional number of firms to survive, and these firms will employ more workers, i.e; they are bigger.

A more recent strand of literature, started with the model of Melitz (2003), takes explicitly into account individual firms' heterogeneity in productivity. In Melitz (2003), firms differ exogenously in their productivity but they can still coexist on the same market because they sell differentiated varieties. However, due to a fixed cost in production, a firm must be sufficiently productive with respect to competitors to be able to break-even. The size of the market and the degree of product substitutability determine the minimum efficiency needed to survive. As the market gets larger, competition becomes fiercer and less productive firms exit. Therefore, firms are bigger and more productive the larger is the local market. Moreover, Melitz and Ottaviano (2005) show that firms' size and productivity should also be positively correlated to market accessibility which is closely linked to the concept of market potential. Many empirical works, like Pavnick (2002), Syverson (2004), and Bernard et al (2006) have confirmed the main testable implications of these models.

3 Data Description and Spatial Variables

3.1 Data Sources

In this paper we use an administrative database provided by INPS (the Italian Social Security Institute). More specifically, we work on a panel version of this database, elaborated by ISFOL, which matches employer and employee information, a similar database like the one used by Kramarz, Abowd, and Margolis (1999) for France. The sample units are full-time workers in all private sectors but agriculture, covering 14 years from 1985 to 1998.[6]

As far as workers' characteristics are concerned, the database contains many individual information like age, gender, qualification, place of birth, workplace, date of beginning and end of the current worker contract, the social security contributions, if the worker is either part time or full time, the gross yearly wage, and the number of worked weeks and days.

As for firms, we have the following set of information: plant location (province), the average number of employees, the sector, the date of start up and the one of shut down (if any). This means that, contrary to other database, we are able to exactly identify where a job takes place since the headquarter and plant location are two separate pieces of information.

As far as job location is concerned, we use data on the 95 Italian provinces.[7] The choice of provinces represents a good compromise between a detailed classification of the Italian territory and data availability. Provinces are in fact sufficiently big to entirely cover cities area and small enough to provide a rich data variability. Data on yearly sectoral employment at the provincial level are provided by INPS and refers to the period 1986-1998. The corresponding sectoral decomposition is the ATECO 81, which splits the Italian economy in 52 sectors (at 2digit level). Province data on education (year 1991), and households' disposable income (period 1991-2000) are provided by the "Istituto Tagliacarne".

As for historical data, information about provinces population in 1861, 1881, and 1901 comes from a re-elaboration of population census by municipalities operated by ISTAT. Data on local sectoral specialization for the year 1951 comes from the "Ateco51-91" database, which is still provided by ISTAT. Finally, data on surface and crow-fly distances among provinces' centroids comes from Arcview GIS software.

3.2 Spatial Variables definitions

In this section we define our spatial variables. As for urbanization externalities, we measure them like in Ciccone and Hall (1996) and Combes (2000):

$$Dens_{j,t} = \ln \left[\frac{empl_{j,t}}{area_j} \right] \tag{1}$$

[6] The sample scheme has been set up to follow individuals born on the 10^{th} of March, June, September and December, and therefore the proportion of this sample on the Italian employees population is approximately of 1/90. Moreover, apprenticeships and part time workers are excluded from the dataset, since the attention is focused on standard labor market contracts (blue collar, white collar and managers). Further, self-employed are not included in INPS database.

[7] Actually, in 2006 the Italian provinces are 103. The transition between 95 and 103 took place in 1995. In this paper we consider the initial classification of 95 provinces, converting subsequent changes of definitions to the old classification.

where $empl_{j,t}$ is employment in location j at time t, while $area_j$ is a location surface in square km. As standard, we consider the log, and we do the same for all the other location variables in order to interpret parameters as elasticities and to ease comparison with previous studies.

The market potential is instead defined using the following spatially weighted component, as originally introduced by Harris (1954) to measure the "potential" demand for goods and services produced in a location $j = 1, 2, .., J$:

$$MP_{j,t} = \ln \left[\sum_{k \neq j} Y_{k,t} \; d_{jk}^{-1} \right], \tag{2}$$

where $Y_{k,t}$ is an index of purchasing capacity of location k (usually disposable income, as in this paper) at time t, and d_{jk} is the distance between two generic locations j and k.[8] The choice to neglect the disposable income of location j, which is rather standard in the literature,[9] helps to mitigate both endogeneity problems and possible multicollinearity with the density variable.[10] Furthermore, the use of a power minus one function of distance (d_{jk}^{-1}) is justified by the gravity equation literature[11] and the trade nature of this literature. It is true that there exist frameworks dealing with the structural estimation of NEG models, like Head and Mayer (2004) and Mion (2004), which use more sophisticated measures of market potential than (2). However, our goal here is to provide a measure of the magnitude of agglomeration economies and not to estimate the underlying model parameters.[12]

Finally, as a proxy for local sectoral specialization ($Spec_{j,s,t}$) of location j in sector s at time t, we use the following index, as in Combes (2000):

$$Spec_{j,s,t} = \ln \left[\frac{empl_{j,s,t}/empl_{j,t}}{empl_{s,t}/empl_t} \right]. \tag{3}$$

[8] It is interesting to note that Harris (1954) did not provide any model to justify its concept of market potential. This is not surprising since general equilibrium models dealing with increasing returns to scale, space, and product differentiation have been introduced only recently. In particular, Fujita, Krugman and Venables (1999) show that market potential functions can be obtained from many spatial general-equilibrium models, thus providing the theoretical background for the use of such an approach.

[9] See Mion (2004) and Hanson (2005).

[10] High density provinces are usually characterized by a large disposable income. This means that, if we considered in the market potential index the disposable income for location j, then the correlation between density and market potential would be relatively high involving possible multicollinearity problems.

[11] See Disdier and Head (2004).

[12] To this respect Head and Mayer (2004), when comparing Harris' market potential with a more elaborated measure, did not find any strong evidence in favor of the latter in terms of predictive power.

3.3 Database Construction and Descriptive Statistics

In our empirical analysis we focus on the period 1991-1998 for which all individual and spatial data are jointly available. Our unit of analysis is a worker i at year t.[13] We further eliminate those extreme observations below (above) the 1^{st} (99^{th}) percentile of the wage distribution, and consider only workers with at least two observations in the period in order to be able to perform a within transformation on our data.

Moreover, in this paper we focus on male individuals with age between 24 and 39 (when they first enter in the database), *i.e.* 31,457 workers and 200,015 observations. The choice to consider only male workers is quite standard in the wage equation setting, as for instance in Topel (1991). Women wage dynamics is in fact often affected by non-economic factors, meaning that standard economic and spatial covariates are less relevant in explaining their carriers. Furthermore, as we will see in Section 4, workplace changes are crucial for estimations and young male workers represents a relatively homogeneous category with respect to migration. Indeed, the related literature - like Borjas, Bronars, and Trejo (1992) and Dahl (2002) - usually focuses on them.[14] Finally, the need to have a good measure of skills, that we measure as individual fixed effects, leads us to consider only workers with more than four observations ending up with a panel of 24,353 workers and 175,700 observations.[15]

The dependent variable in our regressions is the (log) of gross monthly wage in thousands of Italian liras. Data have been deflated and the base year is 1991. As for individual characteristics, we focus on the standard covariates usually used in a Mincerian equation: age, age^2, and two other dummies for blue collar and white collar with the residual category being managers, as well as time and sectoral dummies.

Moreover, in order to capture firms' heterogeneity we use the log of firm size. The positive and strongly significant relation between wages and firm size has been extensively studied in labor

[13] As for job records of a worker, we consider only one employer-employee match per year. In particular, we assign to each individual i the monthly wage and job characteristics of the longest job record in year t. The choice of the monthly wage - reconstructed using yearly wage and worked weeks - is meant to control for both the actual time worked during a year as well as for differences in actual vs reported working time which can systematically vary across space. More precisely, as wage variable we use the yearly wage paid by the firm to the employee, divided by the number of worked weeks, and then reporting the week wage at the monthly level. We did not use the information of the worked days because Ginzburg et al (1998) claimed that this variable in the south could be underestimated, leading to higher daily wages in this region, which is indeed supposed to be the poorest Italian area.

[14] It is for example well known that male prime-age workers are more mobile. Indeed, in our full sample of 92,579 workers 10.49% of them change location at least once in the observation period, with respect to the 13.54% of male prime-age.

[15] Estimations of spatial externalities in the database with all male prime-age workers with more than two observations (31,457 individuals) are available upon request and are virtually identical. The same conclusions further apply when we consider a wider database not restricted to male prime age (92,579 individuals and 560,040 observations): signs of coefficients are the same even if they are smaller in magnitude, as expected.

economics. The seminal papers are probably the ones of Krueger and Summers (1988) and Brown and Medoff (1989). In particular, this literature points out a persistent positive effect of firm size on wages, identifying several different explanations.[16]

As spatial variables we consider employment density, market potential, and localization externalities as defined (respectively) in (1), (2), and (3). Descriptive statistics of the main variables used in our sample are provided in Table 1.

Spatial imbalances come out quite clearly from our data. The spatial distribution of wages is in fact far from being uniform across Italian provinces suggesting that location matters. This does not have to be taken for granted, since Italy is a country characterized by a very important centralized wage setting where, within each sector, contracts have to respect several national based constraints like a minimum wage. However, it is worth noting that firms are allowed to integrate the national contract with a company specific contract in which, for instance, wages can be increased.

To provide some figures about spatial imbalances, the ratio between the highest average province wage (considering time averages over the period) and the lowest one is 1.52, and this ratio is increasing overtime (1.46 in 1991 and 1.56 in 1998). This result still holds if the different qualifications are taken into account. For instance, the same rate is equal to 1.40 for blue collar workers, 1.53 for white collar and 2.82 for managers. Even considering a less extreme indicator than the max/min like the $90^{th}/10^{th}$ percentile ratio we still derive a relevant wage variation. This ratio is in fact 1.24 for all workers, while being 1.22 for blue collar, 1.17 for white collar and 1.33 for managers.

4 What Can Matched Employer-Employee Panel Data Tell us More About Spatial Imbalances?

Most of the studies that has dealt with the measurement of spatial externalities, like Glaeser et al (1992), Ciccone and Hall (1996), and Mion (2004), use aggregate data on labor, wages and productivity. Using individual level data provides relevant steps forward in the analysis. First of all, it is possible to control for possible composition effects due to individual characteristics like age, gender and qualification.[17]

[16] For instance, some papers claim that only more productive and big size firms can afford to pay efficiency wages in order to attract and keep skilled workers (Krueger and Summers, 1988), while other papers stress the importance of unions power in big size firms and the consequent impact on wages. Further, another explanation concerns the fact that big size firms make use of a better screening device in order to select high skilled workers. For a survey concerning all this literature see Oi and Idson (1999).

[17] Indeed, in our data both the age and the gender of workers are strongly correlated with economic density and the same apply, although to a smaller extent, to market potential. For instance, female workers can be found more easily in big cities with the working population being a little older. Since female workers earn less and at the same time wages are positively correlated with age, the sign of the bias coming from omitting these individual control variables on economic

A second more relevant point in using individual data is that the panel nature of our data allow us to control for a very important source of wage variation: unobserved time-invariant individual characteristics (skills). Education is certainly one important component of individual skills, but the use of individual fixed effects is crucial to capture other important time-invariant features, for instance ability, that would be otherwise neglected. This choice is quite standard in labor economics[18], especially when the core of the analysis is not the estimate of the returns to education. Accounting for such individual effects turns out to be particularly interesting when their distribution across space is uneven. Glaeser and Mare (2001), Moretti (2004), and Redding and Schott (2003) provide the rationale for an interplay between skills and space. In particular, these frameworks suggest that skilled workers should be disproportionately found in regions characterized by high density and market potential. Therefore, skills can explain a not negligible share of the agglomeration externalities as usually measured by raw aggregate proxies.

Another important element we can deal with our data is firms' heterogeneity. On the one side, there is an important literature focusing on the positive relation between firm size and wages. Nevertheless, as long as there is no correlation between firm size and location characteristics, omitting the former would have no impact on the estimates of spatial externalities. To this respect, the recent literature on heterogeneous firms (started with the work of Melitz, 2003) predicts that the size of the local market is positively related to firm size and productivity. Indeed, in our data firm size in positively correlated with density (0.12), and this correlation is strongly significant even after introducing sectoral dummies. All of this suggests that controlling for firm size may be considered as another relevant underlying microeconomic factor behind agglomeration externalities.

Finally, the capability to match employer and employee information allow us to investigate the extent to which skills are correlated with firm size, i.e. the existence of a possible assortative matching. The issue of the presence of an assortative matching has been recently questioned in the literature. However, we are able to go a bit further in the analysis -see section 6.3- by assessing whether the observed correlation is just due to co-location of both big firms and skilled workers in cities.

4.1 Econometric Specification and Identification

Once established the importance of using matched employer-employee data we can proceed in our empirical analysis. In particular we use an augmented Mincerian equation that combines standard features of labor economics with spatial externalities:

density is (in this case) undetermined a priori.

[18] See for instance Krueger and Rouse (1998).

$$w_{i,t} = \mathbf{B}_1^{'}\mathbf{I}_{i,t} + \mathbf{B}_2^{'}\mathbf{F}_{f(i,t),t} + \gamma_0 Dens_{j(i,t),t} + \gamma_1 MP_{j(i,t),t} + \gamma_2 Spec_{j(i,t),\ s(f(i,t),t),\ t} + \delta_t + u_i + \varepsilon_{i,t} \quad (4)$$

where subscript i refers to individuals, t to time, j to location, f to firms and bold variables refer to vectors. The dependent variable is the logarithm of gross monthly wage, u_i is an individual effect (skills), and δ_t is a time effect. The term $\mathbf{I}_{i,t} = \{Age_{i,t},\ Age_{i,t}^2,\ Bc\ dummy_{i,t},\ Wc\ dummy_{i,t}\}'$ is a battery of individual characteristics while $\mathbf{F}_{f(i,t),t}$ contains variables that controls for firm f features. The latter is defined either by $\mathbf{F}_{f(i,t),t} = \{\mathbf{i}_{s(f(i,t),t)},\ \ln(FirmSize_{f(i,t),t})\}'$, where $\mathbf{i}_{s(f(i,t),t)}$ is a set of industry dummies and $\ln(FirmSize_{f(i,t),t})$ is log of firm f size (both time varying), or by a fixed effect $\mathbf{F}_{f(i,t)}$ (one for each firm) as in Abowd, Kramarz, and Margolis (1999). Finally, $Dens_{j(i,t),\ t}$, $MP_{j(i,t),\ t}$ and $Spec_{j(i,t),\ s(f(i,t),t),\ t}$ indicate density, market potential and specialization, as defined in (1), (2) and (3).[19]

Identification of spatial variables is crucial in our analysis. As long as OLS are used, both the between and within variance identify γ_0, γ_1, and γ_2 and there is no major issue. However, with our preferred within estimates, identification essentially comes (although not exclusively) from workers changing sector/location. For instance, in the within dimension, only 14% of the variability of density is actually due to density changing over time in a given location with the remaining 86% coming from workers' migrations.[20] As for firms' characteristics, a similar problem arises for the within estimations of sectoral dummies and the coefficient of $\ln(FirmSize_{f(i,t),t})$, which are essentially identified by workers changing employer. Finally, when using the Abowd, Kramarz, and Margolis (1999) estimator (AKM), spatial variables are identified by their time variability only.

5 Regression results

In order to provide a clear overview of the relation between wages, skills, firms and spatial externalities, we present in Table 2 estimations of (4) based on OLS, within estimates without firm size, within estimates with firm size, as well as the firm and individual fixed effects estimator (AKM) of Abowd, Kramarz, and Margolis (1999), from column (1) to (4) respectively. In all specifications, except AKM

[19] It is worth noting that in our notation the sectoral index s (referring to the 52 Ateco81 sectors), the location index j (referring to the 95 provinces), and the firm index f depend ultimately upon the couple (i, t) because they vary when an individual changes sector and/or province and/or firm at time t.

[20] In order to compute these shares of the within variance we first attribute to each worker the same (initial) location for the entire period and then we compute the (without migration) within variance of density. Finally, we compare this within variance with the non-restricted variance that includes workers' movements.

where sectoral dummies cannot be separately identified from firm effects,[21] a complete set of time and sectoral dummies are included.[22]

As for density and market potential, it is quite straightforward to observe that taking into account individual effects dampens simple OLS results (going from column (1) to (2)). Nevertheless, these variables are always significant and elasticities are in line with economic meaningful values. In fact, according to OLS, doubling density increases wages of 2.21%. Previous findings of Ciccone and Hall (1996) for US and Combes, Duranton and Gobillon (2004) for France found (respectively) something around 5% and 3%. Our rather low value is probably due to the already mentioned fact that Italy is characterized by a high degree of wage compression, for instance due to the sectoral minimum wages set at the national level. Taking into account individual effects further reduce this estimate, as showed in the within estimations of column (2), where the coefficient goes down to 0.74%.[23] These simple estimates suggest a strong positive correlation between individual skills, as measured by u_i, and density. Indeed, our within estimations provides a (significant) correlation of 0.21, which suggests that sorting of skills in space is at work. These findings confirm those of Combes, Duranton and Gobillon (2004) for France.

Concerning market potential, spatial sorting is also at work. OLS estimates suggest that doubling market potential lead to a 10.88% increase in wages. Interestingly, in their aggregate analysis of the impact of market potential on sectoral EU wages, Head and Mayer (2006) find a very similar result. However, taking into account individual skills push down elasticity to 5% in the within estimates, in this way showing that using individual data definitely matters in the estimates of spatial externalities. The (significant) correlation between the u_i and market potential is 0.09 which is significantly lower than the one with density but still suggestive of a positive link between skills and those agglomeration externalities stemming from NEG models. This result is consistent with the theoretical and empirical findings of Redding and Schott (2003).

In columns (3) and (4), we further account for firms' heterogeneity by means of (respectively) firm size and firm fixed effects. This is, to our knowledge, the first empirical framework dealing with the joint individual and firm content of spatial externalities. Considering column (3), the firm size elasticity with respect to wages is 1.94%, which is in line with previous findings for other countries. For instance, Brown and Medoff (1989) derive an elasticity value of around 3% for the US. As for the

[21] Sectoral dummies can be separately identified from firm effects as long as firms change sector. However, there are few of such changes in the data and even Abowd, Kramarz and Margolis (1999) do not deal with them.

[22] It is worth noting that our estimates on the impact of $Age_{i,t}$ and $Age^2_{i,t}$, and the two dummies for blue and white collar are in line with previous findings (Naticchioni and Panigo, 2004).

[23] Although this elasticity might seem really low, Di Addario and Patacchini (2006) find a very close result. Using a similar database on individual wages where they also have information of workers' education the authors find that doubling density leads to a 0.53% increase in wages.

impact on spatial variables, considering firm size slightly decrease the elasticities of density (0.56%) and market potential (4.56%). Indeed, the size of firms is significantly correlated with both and in particular with density (0.12), which is the one that experiences the strongest fall. However, compared to the sorting of individuals, the sorting of firms entails a much weaker impact on spatial externalities.[24]

These findings are confirmed by the more general AKM estimation in which the elasticity of density slightly falls to 0.66% while market potential remains substantially stable compared to estimation in column (2).[25] The idea that firms' heterogeneity may lead to dampen the magnitude of spatial externalities has been put forward by Baldwin and Okubo (2004). Our results suggest that the effect of firms' heterogeneity is small, especially if compared to the one induced by individuals.[26]

Finally, the impact of localization externalities, as proxied by our specialization measure, is always very low in all specifications (between 0.55% and 0.01%) and weakly significant. This is consistent with previous works on Italy and in particular with Cingano (2003). For this reason, we do not devote to them much attention, and we will not discuss them further.

5.1 Endogeneity issues

In this subsection we explore the issue of endogeneity. Although within and AKM estimations provide useful insights on the issue of spatial sorting, the reliability of computed elasticities are in fact conditional upon the validity of the underlying moments' restrictions. In particular, it is assumed that $\mathrm{Cov}(\varepsilon_{i,s}, \mathbf{X}_{i,t})$, where $\mathbf{X}_{i,t}$ represents the vector of all covariates, is equal to zero $\forall\ s,\ t$. However, as pointed out by Combes, Duranton, and Gobillon (2004), some local characteristics are likely to be endogenous to local wages. For instance, provinces experiencing a positive technological shock at time t may attract migrants and thus lead to a positive correlation between density and/or market

[24] Another interesting issue is the sectoral scope of our analysis. One could in fact argue that there may be a considerable sectoral heterogeneity with respect to spatial externalities. In Mion and Naticchioni (2005) we show estimations obtained on the sub-sample of manufacturing workers that still confirm that the sorting of skills (firms) is very strong (weak). Furthermore, coefficients are only slightly different from the ones of Table 2.

[25] In particular, we use the order dependent person first method. As for the conditioning variables Z we use the interactions between (mean) individual characteristics (age, age^2, density, and market potential) and (mean) firms characteristics (firm size, firm size^2, and a 9 industry classification based on Ateco 81 one digit). Separate identification of individual and firms effects require 'connections' within a group of workers and firms (see Abowd, Creecy, and Kramarz, 2002). In our estimations we have 24,353 individuals and 28,719 firms forming 15,186 groups. We use all groups and consequently have 37,886 separately identifiable individual and firm effects. It is important to stress that these fixed effects fully account for both individual and firm time invariant characteristics. The fact of having many groups is in fact only a constrain for the separate identification and comparability of firm and individual effects. Therefore, as comparison is only meaningful within a group, we do not report the correlation between the firm and workers' effects and we base our analysis of the interaction between skills and firms' characteristics on individual effects and firm size.

[26] It is worth noting that in the AKM estimation the identification of parameters γ_1 and γ_2 is due to the time-variability of density and market potential only. Parameters are not so different from column (4) where identification was basically driven by migrations, showing that the self-selection bias problem may not be so important for our estimations, as we will see later on.

potential and the residual term. In particular, exogeneity of the location choice is violated whenever workers make their employment choice on the basis of the actual wages at date t.

We deal with endogeneity by means of IV estimations that exploit the idea of Ciccone and Hall (1996) of using deeply lagged values of the endogenous variables as instruments. In column (5) of table 2 we show our within-IV results, which we believe are the most reliable estimates of spatial externalities we can provide. More specifically, we use as instruments for spatial variables data on specialization in 1951, density of population in 1861, 1881, and 1901, as well as a proxy for market potential, calculated replacing aggregate disposable income of a province by its population in equation (2), for the years 1861, 1881, and 1901.[27] The use of deeply lagged levels of specialization, density and market potential obey to the logic (expressed in Ciccone and Hall, 1996) that, as long as early patterns of agglomeration do not reflect factors that influence productivity today, then they can be used as instruments. To this respect, the presence of a structural break would provide the condition for a natural experiment. Ciccone and Hall (1996) use late US 19th century data that are previous to world war one, right after the civil war, and just at the beginning of railroad network construction. Our instruments of density and market potential for Italy meet these needs as the Italian State was created just after Garibaldi expedition in 1860 and the railroad network did not really develop until late 19th century. Crucially, the Sargan test on over-identifying restrictions do not reject the validity of our instruments, and this is quite a strong result considering that, with almost 200,000 observations, the power of the test should be reasonably high.

As one can see, accounting for endogeneity alters the magnitude of spatial externalities. Compared to within estimation in column (3) of Table 2, density goes from 0.56% to 0.20% and is only significant at 10% level. By contrast market potential is just slightly affected going from 4.53% to 4.64%. This suggests that local economic density is much more affected by endogeneity than market potential. Nevertheless, differences with respect to within estimations are not significant at 1% and caution is needed. As for the relative importance of density and market potential, all our estimations suggest that the latter is more important in explaining spatial wage variation. The elasticity corresponding to market potential is in fact always higher (with a gap that is statistically significant) and, when considering "standardized" elasticities, this result still holds with the one corresponding to density (market potential) being 0.0067 (0.0366).[28] Both absolute and standardized elasticities thus suggest

[27] Note that in this way we assume that disposable income is proportional to the population.

[28] Such standardized (or beta) coefficients are defined as the product of the estimated coefficient and the standard deviation of its corresponding independent variable, divided by the standard deviation of the dependent variable. They actually convert the regression coefficients into units of sample standard deviations giving a measure of how much variability of the dependent variable may be explained by the regressor. See Wooldridge (2003, Section 6.1) for a further description of this transformation.

that, in Italy, pecuniary externalities play a crucial role in the spatial distribution of wages.[29]

6 The sorting of skills and firms across space and their interaction

Estimations in Table 2 show clearly that workers' skills and firm size are crucial components of agglomeration economies. All parameters reflecting standard proxies of agglomeration economies are in fact dampened when firm size and (especially) individual fixed effects are taken into account. In this Section we provide further evidence on the magnitude and nature of firms and individuals' sorting across space.

6.1 Evidence on the sorting of workers' skills

We have already argued that the correlation between fixed effects u_i (our measure of workers' skills) is positive and significant with respect to both density (0.21) and market potential (0.09).[30] As for urbanization externalities (proxied by employment density), we first split provinces in low density (LD) and high density (HD) on the basis of the median of the (time average of) density in our database.[31] In Table 3, we then use the individual fixed effects obtained from the regression in column (5) of Table 2 and compute summary statistics of the distribution of skills of residents as well as of migrants (based on workplace) from and to LD and HD provinces.[32] Values represents mean (across individuals) log wage deviations from the overall mean of u_i (zero) and can thus be interpreted as approximate percentage deviations.

Looking at the first column of Table 3, which shows the mean of the u_i in LD and HD provinces, it is very clear that workers in HD provinces are much more skilled (0.0474) compared to those in

[29] In Mion and Naticchioni (2005) we carried out several robustness checks. First, we replicate for comparability the estimation methodology used by Combes, Duranton and Gobillon (2004), in which there is a further time-location specific error component ($v_{j(i,t),t}$) that can be thought as an idiosyncratic technological shock. In particular, we recover the parameters of spatial externalities from a second step regression in which the dependent variable is a fixed location effect $\beta_{j(i,t),t}$ estimated in a first step within regression. Compared to our strategy, their methodology has the advantage of accounting for the heteroschedasticity that comes from time-location shocks. However, when they first recover their dummies without instrumenting, the endogeneity of spatial variables is still at work and can seriously bias estimates. Indeed, in our replication of their estimation, the Sargan test does not accept the validity of instruments. However, we were forced by the lower number of observations to identify time invariant location effects only ($\beta_{j(i,t)}$). Second, we performed separate estimations for four macro areas, in order to check whether spatial effects are possibly due to different returns on age, qualification, or firm size across space, and also to be sure that the North is not driving our results. Also in this case results change only slightly.

[30] The same correlation is negative and very small (-0.01) with respect to specialization. Therefore, being both the latter correlation not particularly strong and our instruments relatively weaker for our proxy of localization externalities, we focus here on the link between skills, market potential and urbanization externalities. The same reasoning applies to the analysis of firms' sorting.

[31] The HD provinces are Torino, Varese, Milano, Vicenza, Venezia, Trieste, Bologna, Roma, Genova, Como, Bergamo, Treviso, Padova, Modena, Firenze, and Napoli. Note that the median is computed across individual observations.

[32] Some individuals moving more than once may actually "score" on more than one category. The same applies to the analysis resumed in Tables 4, 5, 6, and 7.

LD provinces (-0.0469). The average skills gap thus corresponds to a 9.43% difference in mean wage between HD and LD provinces and standard errors (in parenthesis) reveal that this gap is highly significant. The row spatial variation of wages between LD and HD provinces (simple mean of wages across individuals and time in the two groups) equals to 12.78%. Therefore, almost 75% of row wage variability across city size is explained by the sorting of workers' skills.

Columns (2) and (3) of Table 3 look at the dynamic aspect of skills' distribution: migrations. Before interpreting results, it is convenient to remind the reader that there exists a large body of literature concerned with the link between migrations and skills. In particular, this literature strongly suggests that migrants are not a random sample from the population of origin, with skills being a crucial elements of this self-selection process (see, Borjas, 1987). However, even if there exists a sorting of migrants by their skills, it is not always the case that people migrating are the most skilled. As pointed out by Borjas (1987), selection may be either positive or negative depending on the characteristics of the location of origin and destination. To this respect, Borjas, Bronas and Trejo (1992) find that the skills composition and the size of internal US migrations are well explained by interstate differences in the returns to skill: States that pay low returns to skills will see their best workers leaving (positive selection), while States that pay high returns will experience an outflow of unskilled workers (negative selection).[33]

Coherently with the idea that returns to skills are positively related to market size (Moretti, 2004), the sorting in our data is positive from LD to HD provinces because the average skills of migrants (0.0028) is significantly higher than those of the population of origin (-0.0469) although not as good as those of the population of destination (0.0474). The reverse is true, and so sorting is negative as expected, from HD to LD provinces. Furthermore, migrations within each group (LD to LD and HD to HD) suggest that a sizeable positive sorting exists among HD provinces. However, one can wonder to what extent these results hold when considering long-term migrations. To answer this question we report in Table 4 the same summary statistics as in Table 3 with the difference that migrations are now defined as working in a province different from the one where the worker was born. Furthermore, the first column now shows the distribution of skills based on birthplace instead of workplace. As one can see, the sorting of migrants is qualitatively identical.

However, there is an additional remark to underline: the sorting of migrants has a little impact on the overall sorting of skills across provinces. Although the sorting of migrants widens the skills gap between LD and HD provinces - comparing average skills' gap based on birthplace in column

[33]In the migration literature, returns to skills are relative to the average worker and the distribution is supposed symmetric. This means that, for the same average wage across space, a low skilled worker will receive a lower wage in a high returns to skills location.

(1) of Table 4 with the one based on workplace in column (1) of Table 3 - the difference is relatively small. Those who are born in one of the two groups and do not move are already sorted in space and, although 15% of people change group within the working age, this flow does not have a major impact. Overall, these results suggest that spatial sorting is a pervasive feature affecting not only static means but also the dynamic evolution of skills' distribution. However, much of the sorting comes from the static component (non movers), while both the static and the dynamic evidence are coherent with large cities being more attractive for skilled workers.[34]

Tables 5 and 6 provide similar summary data for the sorting of skills by market potential. As suggested by the lower correlation between the u_i and market potential, sorting is relatively less severe here. The "static" sorting of skills by market potential - column (1) of Table 5 - corresponds to a 5.89% gap (compared to the 9.43% with respect to density) between wages of low market potential (LMP) vs high market potential (HMP) provinces. Nevertheless, this 5.89% has to be compared with the 8.16% row wage variation between LMP and HMP provinces. Even though in absolute terms the skills' distribution by market potential turns into a smaller wage variation, skills still explain almost 75% of row spatial wage variation. As for the sorting of migrants we obtain similar, although less clear, results than in the case of the density breakdown. Sorting is certainly positive from LMP to HMP provinces while there is little evidence of a negative sorting from HMP to LMP units. Nevertheless, we can certainly claim that people migrating towards HMP destinations are always "better" than those migrating to LMP provinces. These results are confirmed when looking at migrations based on birthplace (Table 6), which further indicates that long-term migrations have a smaller impact of the distributions of skills compared to the analysis based on density. Overall, these results suggest that spatial sorting has (again) both a static and a dynamic component, which are both coherent with the idea that higher market potential location are more attractive for skilled workers. Furthermore, the static sorting (non movers) is still predominant.

So far, our findings are in line with the idea that returns to skills increase with both density and market potential pushing (essentially non-movers) individuals to invest more in human capital acquisition like in Moretti (2004) and Redding and Schott (2003). Nevertheless, we can't really rule out dynamic human capital accumulation mechanisms of the type described in Glaeser and Mare (2001). Our data do not allow us to observe directly human capital accumulation across space. However we can, as Glaeser and Mare (2001), look at a possible asymmetric effect on wages after moving from (to) a large city as an indirect evidence of a wage growth effect. If urbanization externalities have

[34] It is worth noting that migrations are very important to provide the required variability to identify our spatial coefficients. Hence, one might argue that job-to-job decisions could entail an impact (a bias) on the identification of the spatial coefficients provided in section 5. In Mion and Naticchioni (2005) we point out that the sign of this bias goes in opposite directions depending on the type of migration and eventually cancels out.

a dynamic content, meaning that human capital (skills) is a stock and its accumulation is faster in cities, then wages do not necessarily fall when people that have accumulated such skills in big cities move back to the periphery. Moreover, the gains of people migrating towards HD provinces should not be entirely captured by their wages immediately after migration because such wages will be increasing only over time. We test the consistency of this story with our data in Table 7. Following Glaeser and Mare (2001), we construct a dummy for non-movers living in HD provinces and different dummies for movers. For workers leaving or moving to a HD provinces we use two sets of interacted year dummies to examine the wage dynamics before and after the migration. The first column refers to OLS estimations and the residual category are non-movers living in LD provinces. The second column refers to within estimations and the residual category is all non-movers. All other variables in equation (4) have been used as additional regressors. As one can see, there is some evidence of the wage growth hypothesis only in OLS estimations, where wages increase over time for people moving to HD provinces while no relevant wage losses are observed for people leaving HD locations. However, the evidence on within estimation is more mixed. Results based on a market potential partition of provinces -available upon request- are qualitatively similar and still provide only a mild support for the wage-growth hypothesis.

6.2 Evidence on the sorting of firm size

Unfortunately, our data on firms are not as good as data on individuals and we can't analyze the dynamic content of the spatial sorting of firm size by means of re-locations. However, we can quantify that static aspect of firms' sorting that turns out to be much weaker than the sorting of individuals by skills.

Table 8 shows the average firm size effect by LD and HD provinces (column 1) and LMP and HMP provinces (column 2). Values are obtained multiplying the firm size by the coefficient associated to this variables from the regression of column (5) in Table 2 and then averaging across space and time. Figures represent the average percentage contribution to local wages implied by the local size distribution of firms. As we already stressed, firm size is positively correlated (across firms) to both density (0.13) and market potential (0.07). However, looking at the average firm size effect by LD and HD provinces - column (1) of Table 8 - reveals that it only accounts for a 0.72% wage gap in favor of HD provinces as compared to the 9.43% row spatial variation of wages across the two groups. This percentage is far below the gap explained by workers' skills. To the extent that firm size is a good proxy of firm quality, our results suggest that microfoundations of agglomeration economies aiming at explaining wage differences across space should emphasize on workers' heterogeneity. These findings are confirmed by the breakdown of the firm size effect by LMP and HMP provinces. Firm size distribution differences across provinces imply a 0.45% wage advantage of HMP provinces to be compared with the 8.16% row wage variation.

6.3 Assortative Matching Across Space

Our data suggest that firms and individuals are sorted across space based on (respectively) their size and skills. However, as we are able to match workers and firms, we can also investigate whether "good" workers are matched to "good" firms (the so called assortative matching). There is a pretty large number of studies concerned with this issue of matching. However, both theoretical and empirical contributions point out that, although quite intuitive at first glance, the assortative matching result is not as robust as expected. From the view point of theory, Abowd et al. (2004) and Postel-Vinay and Robin (2006) discuss the role of matching frictions, (un)directed search, wage formation process, scalar heterogeneity and constant returns to scale in order to obtain such a result.

As for empirical evidence, results largely depends on the proxy used to qualify a "good" firm. Improving on Abowd et al (1999), Abowd et al. (2003) use firm fixed effects to measure firm quality and a special algorithm to obtain the otherwise computationally unfeasible exact OLS solution. They find that the correlation with individual effects (skills) is either slightly positive (for US) or negative (for France). However, Postel-Vinay and Robin (2006) argue that the identification of firm fixed effects crucially depends upon workers' mobility across firms that might be too scarce to get sufficiently accurate estimates of firm effects. Furthermore, using French firm level data on log of value added per worker as a proxy of firm quality, they show that the correlation with worker effects is now positive (0.27). In our data we actually find that the correlation between skills (individual fixed effects) and firm size is positive (0.35) and significant. To the extent that firm size is a meaningful variable to capture the underlying firm productivity (an information that we do not have), our results thus support the presence of an assortative matching. To this respect, Postel-Vinay and Robin (2002) show that firm size is a valuable information in order to recover both a firm hiring effort and productivity.

However, even if we observe such an assortative matching, it might in principle just be due to co-location. Melitz and Ottaviano (2005) suggest that big and more productive plants should locate in thick markets. At the same time, according to Glaeser and Mare (2001) and Moretti (2004), skilled workers should be expected to be found disproportionately in big cities. Therefore the "aggregate" correlation (0.35) we found between skills and firm size might just be a spurious result. The fact that we specifically deal with space in our framework allows us to explore this issue more in details. The conditional (with respect to density, market potential and specialization) correlation between the u_i and firm size is only slightly lower (0.34 compared to 0.35) to the unconditional one. This means - and this is one of the main contribution of our paper - that co-location is not really an issue, suggesting that there is a deeper underlying complementarity between skills and firm size, like the one traced for instance in Yeaple (2005).

Even more important, the correlation is positive and significant in each of the 95 Italian provinces, ranging from 0.08 to 0.68. This latter result confirms that assortative matching comes essentially from local labor market interactions. However, the large variability of the correlation across space opens the issue of how local labor market conditions affect such matching.

In our framework, skills are measured by time-invariant effects that are specific to the individual and not to the individual-firm match. Therefore, they may be properly interpreted as worker attributes (like ability and education) that are in principle time invariant and increase a worker productivity irrespectively of the characteristics of the employer. At the same time firm size is also a firm-specific characteristic which is not directly related to the match. Consequently, the correlation between skills and firm size in a given labor marker measures the local degree of interrelation between general individual and firm attributes. To this respect, Wheeler (2001) proposes a spatial labor matching model where firms and individuals are "vertically" differentiated, and shows that the correlation between workers and firms' general quality attributes should increase with market size. Wheeler (2001) further argues that the observed higher educational attainment and skill premium in US cities indirectly support this hypothesis. This quite intuitive approach concerning the positive relation between the degree of assortative matching and market size is rejected in our data. Figure 1 shows the scatter plot of the (time average of) density and degree of assortative matching in each of the 95 Italian provinces together with the regression line. As one can notice, this relation is actually negatively signed and the correlation between the two variables (-0.39) is actually significant at 1% confidence level.

One might argue that our measure of skills is made up by both general and specific (to the match) components, since it is approximated by individual fixed effects. In order to mark out the impact of specific human capital we restrict our analysis to the sub-sample of 10,965 workers that change firm in the estimation period. In this subsample, our measure of skills is more strictly related to general human capital because it does not change across different matches. When considering this subsample, the correlation between density and local assortative matching is still negative (-0.40) and significant.

Another possible explanation might refer to the heterogeneity in the firm size premium across provinces, stemming from both local labor market conditions and economic structure, which are only partially captured in our set of covariates. In order to control for this possibility we allow the coefficient of firm size to be province-specific. Quite interestingly, we find that this explanation partially holds, since the correlation between firm size and skills reduces from -0.39 to a significant -0.16, suggesting that the firm size premium decreases with market size.

The negative correlation between the degree of assortative matching and market size is thus a robust result. One way to interpret this negative relationship, which represents one of the contribution of this papers, is to consider the "horizontal" dimension of skills and productivity like in Helsey and Strange (1990). As long as firm-worker specific attributes are taken into account, then the larger the size of a market the larger is the possibility of a worker/firm to find a partner with some specific characteristics which are valuable assets for the matching. Therefore, general attributes should became relatively less important the larger is the size of the market. In this line, Kim (1989) proposes a framework that combines both general and specific skills with market size. Workers and firms are horizontally differentiated regarding to (respectively) their job requirements and abilities. The mismatch between a given worker ability and a firm job requirement represents a cost that can be reduced by a worker investment in general human capital. However, a worker may also decide to invest in specific human capital to increase his skill-specific productivity. When the size of the market increases, workers invest more in specific human capital than in general human capital, because they can find more easily a firm with job requirements close to their ability.

7 Conclusions

In this paper we show that using individual data is crucial in order to get into the black box of agglomeration externalities. Our results provide evidence that the spatial sorting of firm size and (especially) skills explains a large portion of the raw spatial wage variation, so dampening the magnitude of aggregate proxies for urbanization and localization externalities, as well as for market potential. Using the matched nature of our database we can also state that, even after controlling for co-location issues, an assortative matching between skilled workers and large firms is at work in the Italian labor market. Moreover, we are able to further characterize this result showing that the assortative matching is negatively related to the size of local labor market, which is consistent with the framework developed by Kim (1989).

Our results suggest that the microfoundation of agglomeration externalities should devote more effort to model the process of human capital accumulation across space. Furthermore, there is a need to collect further evidence about the specific *vs* general nature of human capital investments in cities. Finally, it seems important to encourage a wider and deeper discussion concerning the interactions between skills and firm heterogeneity, both at the empirical and theoretical levels.

References

[1] Abdel Rahman, H. M., and M. Fujita (1990). "Product variety, Marshallian externalities, and city size". Journal of Regional Science Vol. 30(2), 165-183.

[2] Abowd, J., F. Kramarz, and D. Margolis (1999). "High-wage workers and high-wage firms". Econometrica Vol. 67, 251-333.

[3] Abowd, J. M., Kramarz, F., Lengermann, P., and Roux, S. (2003). "Interindustry and firm-size wage differentials in the United States and France", Cornell University working paper.

[4] Abowd, J., F. Kramarz, P. Lengermann, and S. Perez-Duarte (2004). "Are Good Workers Employed by Good Firms? A test of a simple assortative matching model for France and the United States". Mimeo

[5] Abowd, J., R. Creecy, and F. Kramarz (2002). "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data". Cornell University, mimeo. $http://instruct1.cit.cornell.edu/\tilde{}jma7/abowd-creecy-kramarz-computation.pdf$

[6] Baldwin, R.E., and T. Okubo (2004). "Heterogeneous firms, agglomeration and economic geography: Selection and sorting". CEPR discussion paper 4602.

[7] Bernard, A.B., Jensen, J.B , and P.K., Schott (2006). "Trade Costs, Firms and Productivity". The Journal of Monetary Economics, forthcoming.

[8] Borjas G.J., (1987). "Self-Selection and Earnings of Immigrants". The American Economic Review Vol. 77(4), 531-553.

[9] Borjas G.J., Bronars S.G., and S.J. Trejo (1992). "Self-selection and Internal Migrations in the United States". Journal of Urban Economics Vol. 32, 159-185.

[10] Bresnahan T., and P., Reiss (1991). "Entry and Competition in Concentrated Markets". Journal of Political Economy Vol. 99, 977-1009.

[11] Brown, G., and J. Medoff (1989). "The Employer Size Wage Effect". Journal of Political Economy Vol. 97, 1027-1057.

[12] Campbell, J.R. and Hopenhayn, H.A., (2005), "Market Size Matter", Journal of Industrial Economics, Vol. 53(1), pp. 1-25.

[13] Ciccone, A., R. Hall (1996). "Productivity and the Density of Economic Activity". American Economic Review Vol. 86, 54-70.

[14] Cingano, F. (2003). "Returns to Specific Skills in Industrial Districts". Labour Economics Vol. 10, 149-164.

[15] Combes, P.P. (2000). "Economic structure and local growth: France, 1984-1993". Journal of Urban Economics Vol 47, 329-355.

[16] Combes P.P., G., Duranton, and L., Gobillon (2004). "Spatial wage disparities: Sorting matters!". CEPR discussion paper 4240.

[17] Dahl, G.B. (2002). "Mobility and the Return to Education: Testing a Roy Model with Multiple Markets". Econometrica Vol. 70(6), 2367-2420.

[18] Di Addario, S., and E., Patacchini (2005). "Is There an Urban Wage Premium in Italy?". Bank of Italy, Banca D'Italia, Temi di Discussione, no. 570.

[19] Disdier, A. C., and K., Head (2004). "The Puzzling Persistence of the Distance Effect on Bilateral Trade"Ṁimeo, *http://pacific.commerce.ubc.ca/keith/Papers/meta.pdf*.

[20] Duranton G., and D., Puga (2004). "Microfoundations of urban agglomeration economies". Handbook of Regional and Urban Economics, Vol.4., J.V. Henderson and J.-F. Thisse (eds.), Elsevier-North Holland, Amsterdam.

[21] Fujita, M., P., Krugman, and A., J., Venables. "The spatial economy. Cities, regions and international trade". MIT Press, Cambridge 1999.

[22] Fujita, M., and J.F., Thisse. "Economics of Agglomeration". Cambridge University Press, Cambdridge 2002.

[23] Glaeser, E.L., H., Kallal, J., Sheinkman, and A., Shleifer (1992). "Growth in Cities". Journal of Political Economy Vol. 100, 1126-1152.

[24] Glaeser, E.L., and D.C., Mare (2001). "Cities and Skills" Journal of Labor Economics Vol. 19(2) 316-342.

[25] Ginzburg, A., G., Solinas, M., Scaltriti, and R., Zoboli (1998). "Un nuovo autunno caldo nel mezzogiorno? Note in margine al dibattito sui differenziali territoriali" Politica Economica Vol.14(3).

[26] Hanson, G.H. (2005). "Market potential, increasing returns, and geographic concentration". Journal of International Economics, forthcoming.

[27] Harris, C.D. (1954). "The Market as a Factor in the Localization of Industry in the United States". Annals of the Association of American Geographers Vol. 44, 315-348.

[28] Head K., and T., Mayer (2004). "Market Potential and the Location of Japanese Investment in the European Union". Review of Economics and Statistics Vol. 86(4), 959-972.

[29] Head, K., and T., Mayer (2006). "Regional Wage and Employment Responses to Market Potential in the EU". Regional Science and Urban Economics, forthcoming.

[30] Helsley, R.W., and W.C., Strange (1990). "Matching and agglomeration economies in a system of cities". Regional Science and Urban Economics Vol. 20(2), 189Ű212.

[31] Henderson, J.V. (1974). "The sizes and types of cities". American Economic Review Vol. 64(4), 640-656.

[32] Henderson, J.V. (2003). "Marshall's Scale Economies". Journal of Urban Economics Vol. 53, 1-28.

[33] Kim, Sunwoong (1989). "Labor specialization and the extent of the market". Journal of Political Economy Vol. 97(3), 692-705.

[34] Krueger, A.B., and L.H. Summers (1988). "Efficiency Wages and the Inter-Industry Wage Structure". American Economic Review Vol. 56(2), 259-293.

[35] Krueger, A.B., and C., Rouse (1998). "The Effects of Workplace Education on Earnings, Turnover and Job Performance". Journal of Labor Economics, Vol. 16(1), 61-94.

[36] Marshall, Alfred. “Principles of Economics”. 1890. London: Macmillan.

[37] Melitz, M. (2003). “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity”. Econometrica Vol. 71(6), 1695-1725.

[38] Melitz M. and G. Ottaviano (2005) Market size, trade and productivity, NBER Working Paper No.11393.

[39] Mion, G. (2004). “Spatial Externalities and Empirical Analysis: The Case of Italy”. Journal of Urban Economics Vol. 56(1), 97-118.

[40] Mion, G. and P.Naticchioni (2005). “Urbanization Externalities, Market Potential and Spatial Sorting of Skills and Firms”. CEPR Discussion Paper, no. 5172.

[41] Moretti, E. (2004). “Human Capital Externalities in Cities”. Handbook of Regional and Urban Economics, Vol.4., J.V. Henderson and J.F. Thisse (eds.), Elsevier-North Holland, Amsterdam.

[42] Naticchioni, P., and D., Panigo (2004) “Employment protection, job-tenure and short term mobility wage gains: A New Explanation for the Italian Case”. CEPREMAP Working Papers 2(Couverture Orange).

[43] Oi, W.I., and T.L., Idson (1999). “Firm Size and Wages”, in Ashenfelter, O., and Card, D. (eds.), Handbook of Labor Economics, vol.3, Amsterdam: Elsevier Science, North-Holland, 2165-2214.

[44] Pavcnik, N. (2002). “Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants”. Review of Economic Studies Vol. 69, 245-276.

[45] Postel-Vinay, F. and J.M. Robin (2002). “Wage Dispersion with Worker and Employer Heterogeneity”. Econometrica Vol. 70(6), 2295-350.

[46] Postel-Vinay, F. and J.M., Robin (2006). t'tMicroeconometric Search-Matching Models and Matched Employer-Employee Data”, in R. Blundell, W. Newey and T. Persson, editors, Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress, Volume 2, Cambridge: Cambridge University Press. Forthcoming.

[47] Redding, S., and P.K., Schott (2003). “Distance, skill deepening and development: will peripheral countries ever get rich?”. Journal of Development Economics Vol. 72, 515-541.

[48] Rosenthal, S.S., and W.C., Strange (2001). “The Determinants of Agglomeration”. Journal of Urban Economics Vol. 50, 191-229.

[49] Rosenthal, S.S., and W.C., Stange (2004). “Evidence on the Nature and Sources of Agglomeration Economie”. Handbook of Regional and Urban Economics, Vol.4., J.V. Henderson and J.-F. Thisse (eds.), Elsevier-North Holland, Amsterdam.

[50] Syverson, C. (2004). “Market Structure and Productivity: A Concrete Example”. Journal of Political Economy Vol. 112(6), 1181-1222.

[51] Tabuchi, T., and A., Yoshida (2000). “Separating Agglomeration Economies in Consumption and Production” Journal of Urban Economics Vol. 48, 70-84.

[52] Topel, R. (1991). “Specific capital, mobility, and wages: Wages rise with job seniority”. Journal of Political Economy Vol. 99, 145-176.

[53] Wheeler C.H. (2001). "Search, Sorting and Urban Agglomeration". Journal of Labor Economics Vol. 19(4), 879-899.

[54] Wooldridge, Jeffrey M. (2003). "Introductory Econometrics". 2 ed. South-Western.

[55] Yeaple, S. (2005). "A Simple Model of Firm Heterogeneity, International Trade, and Wages" . Journal of International Economics Vol. 65(1), 1-20.

Table 1. Summary Statistics

Variable	Observ.	Mean	Std. Dev.	Min	Max
ln(wage)	175700	6.4824	0.3625	3.1180	8.2036
Age	175700	34.1257	5.0713	24.0000	46.0000
Age^2	175700	1190.2820	350.7040	576.0000	2116.0000
Firmsize	175700	4.6278	2.7433	0.0000	12.2699
Bc Dummy	175700	0.6528	0.4761	0.0000	1.0000
Wc Dummy	175700	0.3319	0.4709	0.0000	1.0000
Specialization	175700	0.0622	1.0412	-8.7156	4.9706
Density	175700	3.9525	1.2167	0.6903	6.2398
Market Potential	175700	8.5566	0.2864	7.7637	9.0872
Specialization in 1951	175700	-0.0679	0.8887	-7.2878	3.6926
Density pop in 1861	175700	4.7640	0.6980	2.9671	6.7048
Density pop in 1881	175700	4.9202	0.6847	3.0494	6.8617
Density pop in 1901	175700	5.0708	0.7029	3.1931	6.9914
Market Potential in 1861	175700	12.5091	0.0667	12.4026	12.7018
Market Potential in 1881	175700	12.6485	0.0638	12.5404	12.8337
Market Potential in 1901	175700	12.7721	0.0633	12.6706	12.9480

All variables (except Sex, Age, Ageˆ2 , Bc Dummy, and Wc Dummy) are, coherently with their definition in the text, expressed in natural logarithm. Wages are in log of thousands liras while Market Potential is in log of billions liras. Both are in real terms (base 1991). Market Potential and Density in 1861-1901 are computed on the basis of inhabitants.

Table 2. Regression results. Dependend valiable ln(wage)

	(1)	(2)	(3)	(4)	(5)
Age	0.0481***	0.0456***	0.0466***	0.0464***	0.0455***
	(0.0015)	(0.0080)	(0.0080)	(0.0075)	(0.0079)
Age^2	-0.0005***	-0.0005***	-0.0005***	-0.0005***	-0.0005***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Firmsize			0.0194***		0.0193***
			(0.0004)		(0.0004)
Bc Dummy	-0.7619***	-0.2132***	-0.2149***	-0.2162***	-0.2114***
	(0.0049)	(0.0041)	(0.0041)	(0.0038)	(0.0041)
Wc Dummy	-0.4891***	-0.1452***	-0.1466***	-0.1468***	-0.1461***
	(0.0049)	(0.0031)	(0.0031)	(0.0029)	(0.0031)
Specialization	0.0055***	0.0015*	0.0008	0.0001	-0.0037
	(0.0006)	(0.0008)	(0.0008)	(0.0008)	(0.0035)
Density	0.0221***	0.0074***	0.0056***	0.0066***	0.0020*
	(0.0005)	(0.0010)	(0.0011)	(0.0010)	(0.0011)
Market Potential	0.1088***	0.0500***	0.0453***	0.0516***	0.0464***
	(0.0021)	(0.0058)	(0.0058)	(0.0054)	(0.0134)
Estimation mehtod	OLS	Within	Within	AKM	IV-Within
Time & Sector Dummies	Yes	Yes	Yes	Time only	Yes
Firm fixed effects	No	No	No	Yes	No
R^2	0.5249	0.3596	0.4412		
Endog. Test (df=4)					1.1722
N. of individuals	24353	24353	24353	24353	24353
N. of observations	175700	175700	175700	175700	175700

Standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels.

Table 3. Skills distribution across provinces and migrants characteristics based on workplace: Density

		Migration Flows	
	Whole population	LD provinces (destination)	HD provinces (destination)
LD provinces (origin)	-0.0469*** (0.0021)	-0.0614*** (0.0076) [1184]	0.0028 (0.0083) [1037]
HD provinces (origin)	0.0474*** (0.0023)	-0.0009 (0.0083) [1042]	0.1089*** (0.0102) [881]

Standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 4. Skills distribution across provinces and migrants characteristics based on birthplace: Density

		Migration Flows	
	Whole population	LD provinces (destination)	HD provinces (destination)
LD provinces (origin)	-0.0329*** (0.0023)	-0.0407*** (0.0054) [2235]	0.0513*** (0.0049) [2839]
HD provinces (origin)	0.0469*** (0.0028)	0.0121 (0.0089) [885]	0.1170*** (0.0079) [1278]

Standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 5. Skills distribution across provinces and migrants characteristics based on workplace: Market Potential

	Whole population	Migration Flows: LMP provinces (destination)	Migration Flows: HMP provinces (destination)
LMP provinces (origin)	-0.0279*** (0.0024)	-0.0497*** (0.0075) [1251]	0.0374*** (0.0103) [817]
HMP provinces (origin)	0.0310*** (0.0023)	0.0395*** (0.0106) [722]	0.0416*** (0.0083) [1187]

Standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 6. Skills distribution across provinces and migrants characteristics based on birthplace: Market Potential

	Whole population	Migration Flows: LMP provinces (destination)	Migration Flows: HMP provinces (destination)
LMP provinces (origin)	-0.0294*** (0.0025)	-0.0143*** (0.0051) [2644]	0.0185*** (0.0059) [2104]
HMP provinces (origin)	0.0393*** (0.0027)	0.0807*** (0.0123) [501]	0.0912*** (0.0063) [1906]

Standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 7. Analysis of the wage growth hypothesis. Dependent variable ln(wage)

	(1)	(2)
Non movers living in HD province	0.0501***	dropped
	(0.0012)	
Moving to a HD province:		
Observed 4 or more years before a move	-0.0142*	-0.0096*
	(0.0081)	(0.0056)
Observed 2-3 years before a move	0.0018	-0.0083*
	(0.0066)	(0.0048)
Observed 1 year before a move	0.0030	-0.0146**
	(0.0078)	(0.0051)
Observed 1 year after a move	0.0305***	0.0044
	(0.0083)	(0.0054)
Observed 2-3 years after a move	0.0379***	0.0021
	(0.0072)	(0.0052)
Observed 4 or more years after a move	0.0594***	0.0073
	(0.0095)	(0.0064)
Leaving from a HD province:		
Observed 4 or more years before a move	0.0264***	0.0107**
	(0.0075)	(0.0053)
Observed 2-3 years before a move	0.0165**	0.0107
	(0.0066)	(0.0048)
Observed 1 year before a move	-0.0006	0.0022
	(0.0077)	(0.005)
Observed 1 year after a move	-0.0096	-0.001
	(0.0087)	(0.0055)
Observed 2-3 years after a move	-0.0125*	-0.0029
	(0.0074)	(0.0053)
Observed 4 or more years after a move	0.0087	0.002
	(0.0104)	(0.0069)
Estimation method	OLS	Within
R^2	0.5325	0.4026
N. of individuals	24353	24353
N. of observations	175700	175700

Standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Regressions contains Age, Ageˆ2, Firmsize, Bc Dummy, Wc Dummy, Specialization, Market Potential as well as year and sectoral dummies.

Table 8. Distribution of the firm size effect by density and market potential

	Firm size premium by density	Firm size premium by market potential
LD provinces or LMP provinces	0.0607*** (0.0003) [15835]	0.0618*** (0.0003) [14640]
HD provinces or HMP provinces	0.0679*** (0.0003) [13818]	0.0663*** (0.0003) [14903]

Standard errors in parentheses with ***, ** and * respectively denoting significance at the 1%, 5% and 10% levels. Number of firms is in square brackets

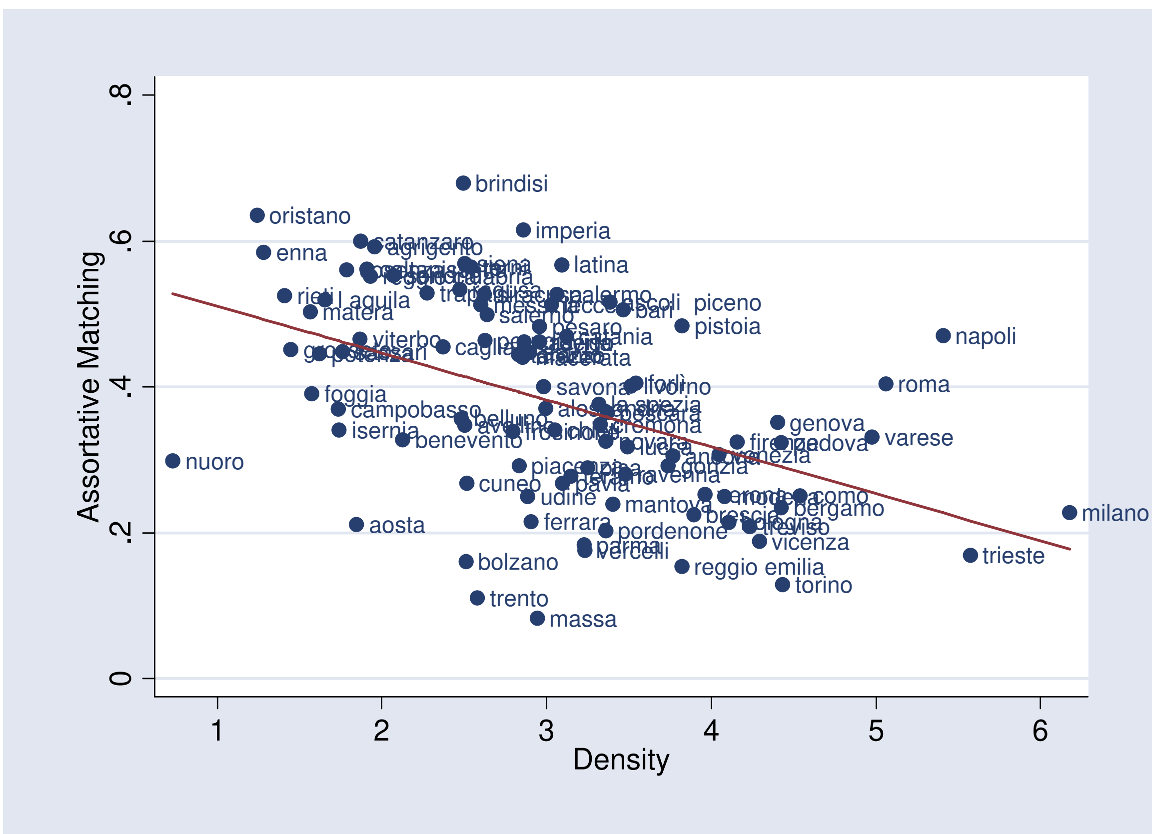

Figure 1: Assortative matching and market density