The Spatial Sorting and Matching of Skills and Firms*

Giordano Mion† and Paolo Naticchioni‡

November 2007

Abstract

In this paper we make use of a matched employer-employee database for Italy to look at the spatial distribution of wages. Using this rich database we aim to open up the black box of agglomeration economies exploiting the micro dimension of interaction among economic agents, both individuals and firms. We provide evidence that firm size and, especially, skills are sorted across space and account for a large portion of the spatial wage variation. Our data also support the assortative matching hypothesis, which we show not to be driven by co-location of good workers and firms. Finally, we point out that assortative matching is negatively related to local market size.

Keywords: Spatial Externalities, Individual Panel-Data, Skills, Firms' Heterogeneity, Sorting, Assortative 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, Pier-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, AIEL, DICE and INRS 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, and FNRS. Address: CORE-UCL, 34 Voie du Roman Pays, B-1348 LLN, Belgium. Email: giordano.mion@uclouvain.be.

‡Paolo Naticchioni, Università di Roma "La Sapienza", and Université Catholique de Louvain. 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. There are, in fact, striking spatial disparities in both the developed and developing countries, arousing considerable political concern and, in the case of the EU, they are so strategically important as to be ranked at the top of the political agenda.[1] As for wages, Glaeser and Mare (2001) find that they are 33% higher in the US cities than in the outer metropolitan areas. The data evidence on the EU as a whole is less systematic. However, a number of country-based studies, as for instance Combes, Duranton and Gobillon (2007), show that wages vary considerably across space.

There is also a large body of empirical literature concerned with measurements of the growth and productivity advantages stemming from agglomeration economies such as urbanization and localization externalities, and market potential. However, despite some significant exceptions,[2] most frameworks deal with aggregate proxies for agglomeration, neglecting the very nature of these externalities that derives from local interaction between economic agents, both individuals and firms.

From the viewpoint of theory, there are at least three explanations for the fact that skilled workers are disproportionately located in cities (sorting of skills). First, cities provide valuable consumption amenities for skilled workers, such as theaters, museums, cultural activities, etc. Second, as suggested by Moretti (2004), the bigger cities offer higher returns to education (private plus social), fostering 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, as 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 cities. The size of the local labor market improves the quality of matching between firms and workers, boosting productivity. Furthermore, a relatively recent strand of the literature, starting with the model of Melitz (2003), explicitly takes into account the relation between the heterogeneity of firms in productivity and the size of the market.

The empirical literature – Glaeser et al (1992), Ciccone and Hall (1996), and Mion (2004), for example – has used aggregate data on labor, wages and productivity to deal with the measurement of spatial externalities.

[1] The reduction of income disparities among EU regions involves much of the political debate with Structural and Cohesion Funds, both aiming at the reduction of imbalances. Further, these funds 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 (2007).

In this paper, we make use of a matched employer-employee panel for Italy in the period 1991-1998. We show that analysis can advance significantly using individual level data. First, it is possible to control for possible composition effects due to individual characteristics like age, gender and occupation. Second, thanks to the panel nature of our data we can control for a very important source of wage variation, namely 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 otherwise be neglected. Furthermore, using an employer-employee panel we are also able to investigate the spatial sorting of firms. Firm heterogeneity is measured by the size of the firm, which has been shown to represent a proper proxy for firm productivity (Postel-Vinay and Robin, 2002).

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 gaining some insight into the underlying mechanisms leading to such sorting. In other words, our aim is to help open the black box of agglomeration externalities.

Our analysis basically proceeds in two parts. In the first part we investigate the spatial sorting of skills and firms using a standard wage equation. Spatial variable coefficients concerning urbanization, market potential and specialization externalities fall sharply when introducing skills and firms heterogeneity, providing a preliminary clue as to the relevance of sorting. These findings are robust to different econometric specifications and endogeneity issues. The first step estimates also provide us with an appropriate measure of skills, i.e. the individual fixed effects.

In the second part of the analysis we further characterize the spatial sorting of individuals and firms, as well as the degree of assortative matching between skilled workers and large firms across locations. In particular, we show that nearly 75% of raw wage variation across Italian provinces is accounted for by differences in individual skills between low and high-density provinces, as well as between low and high market potential provinces. Furthermore, we provide evidence that sorting is essentially static, i.e. due to non-migrants. On the other hand, the sorting of firms accounts only for a small fraction of wage variability.

Another interesting related point 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, i.e. assortative matching. The evidence in the literature concerning assortative matching is mixed. In our data we do in fact find a positive correlation, and we further show that the results are not driven by matters of co-location.

A further interesting finding is obtained when 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 coming in line with Kim (1989). As long as firm-worker specific attributes are taken into account one can argue that the larger the size of a market, the higher is the probability for a worker/firm to find a partner with some specific characteristics that represent valuable assets for the matching. Therefore, the larger the size of the market, the less the general attributes should count.

With respect to the existing empirical literature, our paper introduces two major original contributions. First, it represents the first analysis of both firm and worker sorting using individual wage level data. Combes, Duranton and Gobillon (2007) emphasize the sorting of skills but they cannot account for the sorting of firms, since their database does not contain firm information. Second, we extend the Abowd, Kramarz and Margolis (1999) analysis of assortative matching to a spatial dimension.

The rest of the paper is organized as follows. In Section 2 we present the theoretical background to our analysis. In Section 3 we present the data, the spatial variable definitions and some descriptive statistics. Section 4 is devoted to the econometric results, and we also deal with endogeneity. Section 5 investigates the issue of individual and firm sorting and the one of assortative matching. Finally, our conclusions are set out in Section 6.

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 affect local productivity and growth. There is a large body of literature,[3] both theoretical and applied, that has identified a number of mechanisms leading to the emergence of agglomeration economies. Despite the large number of empirical studies, however, data limitations and the lack of a unified underlying theoretical framework make it difficult to disentangle the relative importance of different microfoundations. One of the goals of this paper is to contribute to this debate on microfoundations by shedding lights on the relative importance of individual versus firm characteristics.

In this paper we focus on the most widely investigated agglomeration economies: urbanization externalities, market potential and localization externalities. As for urbanization externalities, the idea that market size has a positive impact on local productivity goes back to Marshall (1890) and

[3] See, for instance, the recent reviews of Duranton and Puga (2004) and Rosenthal and Strange (2004).

has been formalized by Abdel Rahman, and Fujita (1990) among others. Duranton and Puga (2004) discuss different microfoundations of urbanization economies.

The second spatial externality investigated is market potential, defined as the demand that a firm located in a given region can have access to in a world where trade is costly. It was first put forward by Harris (1954) and then formalized by the so-called new economic geography literature: if trade costs are sufficiently low, then agglomeration of economic activities occurs, boosting local wages.[4]

Finally, localization externalities refer to the advantage stemming from the concentration of a specific industry across space. The idea that local specialization fosters productivity also goes back to Marshall (1890). The models of Henderson (1974) and Duranton and Puga (2004) among others provide micro-foundations for these externalities.

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

The fact that 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 firm size.

The sorting of workers' skills

The term skills refers here to workers' attributes, like ability and education, which are in principle time-invariant and increase worker productivity irrespectively of the characteristics of the employer. The idea that skilled workers are sorted across space has not been so closely examined as other aspects, but is not new in the urban literature.

The oldest explanation is the so-called bright lights hypothesis. Skilled workers are known to be more mobile. Therefore, as long as the extensive consumption amenities and services provided in cities represent particularly valuable assets for them, their migrations should preferentially take them to the cities. Another possibility, which does not necessarily involve migrations, is that cities offer higher returns to education, so that residents find it profitable to invest more in the accumulation of human capital, as in Moretti (2004). On the other hand, Glaeser and Mare (2001) discuss a third dynamic mechanism. Large cities can in fact be places where the accumulation of human capital may be faster than anywhere else due to intense, stimulating face-to-face interactions.

Finally, the link between human capital accumulation and space also finds some support in the new economic geography literature. Redding and Schott (2003) develop a model where remoteness (defined as poor access to foreign markets) depresses the skill premium and therefore offers fewer incentives for human capital accumulation.

[4]See, for instance, Fujita, Krugman and Venables (1999).

The sorting of firm size

There are many theoretical frameworks where agglomeration economies translate into higher firm productivity, the latter usually being positively related to the size of the firm. 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 cities. More specifically, 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 makes it feasible only for a less than proportional number of firms to survive, and these firms will employ more workers, i.e. they grow bigger.

A more recent strand of literature, starting with the Melitz (2003) model, takes explicitly into account individual firm heterogeneity in productivity. In Melitz (2003), firms differ exogenously in their productivity. Due to fixed cost in production, a firm must be sufficiently productive with respect to competitors to be able to break even. As the market gets larger, competition becomes fiercer, the least productive firms exit, and the surviving firms grow larger.

The local interaction between workers and firms characteristics

The spatial literature agrees on the fact that one should observe phenomena of co-location of the better workers and firms in agglomerated areas. However, there is no general consensus as to which worker and firm characteristics should matter most within a given location (for instance concerning issues related to specific *vs* general human capital). On the evidence of our data we can provide some useful pointers.

While, on the one hand, Wheeler (2001) proposes a spatial labor matching model where firms and individuals are vertically differentiated, and shows that the correlation between the general quality attributes of workers and firms (i.e. not specific to the match) should increase with market size, it could be argued that workers are also horizontally differentiated, as in Helsley and Strange (1990): as long as firm and worker specific attributes are taken into account, the larger the market, the higher the probability will be for worker/firm to find a partner with some valuable specific characteristics for the match.

In this line, Kim (1989) proposes a framework that combines both general and specific skills. He finds that, as the size of the market increases, workers invest more in specific human capital.

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 database similar to the one used by Abowd, Kramarz 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.[5]

As far as workers' characteristics are concerned, the database contains individual information like age, gender, occupation, place of birth, workplace, date of beginning and end (if any) of the current contract, part/full time dummy, the gross yearly wage, and the number of worked weeks and days.

As for firms, we have the following set of data: plant location (province), average number of employees, industry, date of start up and shut down (if any). This means that, unlike other databases, with this we are able to identify exactly where a job occurs since the headquarters and plant locations are two separate pieces of information.

As far as job location is concerned, we use data on the 95 Italian provinces. The choice of provinces is a convenient compromise between detailed classification of the Italian territory and data availability. The provinces are in fact large enough to cover entire cities area and small enough to afford rich data variability. Data on yearly sectoral employment at the provincial level are provided by INPS and refer to the period 1986-1998. The corresponding industrial decomposition is ATECO 81, which divides the Italian economy into 52 industries (at 2digit level). Households' disposable income (period 1991-2003) is provided by the Istituto Tagliacarne.

As for the historical variables, data on province populations in 1861, 1881, and 1901 come from the processing of population censuses by municipalities, carried out by ISTAT. Data on local industrial specialization for the year 1951 come from the ATECO-91 database, which continues to be provided by ISTAT. Finally, the data on surface and crow-fly distances between province centroids come from Arcview GIS software.

[5]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 in the population of Italian employees is approximately 1/90. Moreover, apprenticeships and part time workers are excluded from the dataset, since attention focuses on standard labor market contracts (blue collar, white collar and managers). Further, the self-employed are not included in INPS database.

3.2 Spatial Variables Definitions

In this section we define our spatial variables, computed at the aggregate provincial level. As for urbanization externalities, we measure them as in Ciccone and Hall (1996) and Combes (2000):

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

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

Market potential is 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 $j = 1, 2, .., \Phi$:

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

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. Fujita, Krugman and Venables (1999) show that market potential functions like (2) can be derived from different spatial general-equilibrium models. The choice to neglect the disposable income of location j, which is fairly standard in the literature, helps to mitigate both endogeneity problems and possible multicollinearity with the density variable.[6]

Finally, as a proxy for local industrial specialization ($Spec_{j,s,t}$) of location j in sector s at time t, we use the following index, as in Combes (2000), which compares the local to the national degree of specialization for a given industry:

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

3.3 Database Construction

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 in year t.[7] We further eliminate those extreme observations below (above) the 1 st (99 th) percentile of the wage distribution, and consider only

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

[7] We consider only one employer-employee observation per year. In particular, if more than one observation per year is available we assign to each individual i the monthly wage and job characteristics of the longest job record.

workers with at least three observations in the period in order to be able to perform reliable within transformation on our data. This leads us to an unbalanced panel of 92,579 individuals corresponding to 560,040 observations over the span 1991-1998.

However, in the core part of the paper we focus on male individuals aged between 24 and 39 (when they first enter 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 shown in Section 5.1, workplace changes are crucial for the identification of our estimates, and male prime-age workers represent a relatively homogeneous category with respect to migrations. Indeed, the related literature - like Dahl (2002) - usually focuses on them. Finally, the need to have a good measure of skills, which 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.[8]

The dependent variable in our regressions is the (log) of gross monthly wage in thousands of Italian liras.[9] The 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, the residual category being managers, as well as time and sectoral dummies.

Moreover, in order to capture firm 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 economics. The seminal papers are those by Krueger and Summers (1988) and Brown and Medoff (1989). However, to the best of our knowledge, we are the first to analyze the impact of the unequal distribution of firm size across space on wages.

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.

---

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

[9] 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 and 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, then translating the weekly wage at the monthly level. We have not used the variable concerning worked days since Ginzburg et al. (1998) claim that this variable could be underestimated in Southern Italy, leading to higher daily wages in this region which is known to be poorer that Northern Italy.

4 Estimation

The first step in the investigation of the spatial sorting of skills and firms is the estimation of a wage equation. The degree to which the estimates of spatial agglomeration coefficients reduce when controlling for skills and firm heterogeneity provides an initial insight into the relevance of sorting. This Section is devoted to the issue.

These estimations will also provide us with an accurate measure of firm and individual heterogeneity, which will be used in the next Section to analyze both the impact of sorting on the spatial distribution of wages and the degree of assortative matching between skilled workers and large firms across locations.

4.1 Econometric Specification, Identification, and Results

We use an augmented Mincerian equation that combines standard features of labor economics with spatial characteristics, and go on to estimate four different models, beginning with OLS estimates of model (M1), which contains neither individual fixed effects nor firm characteristics:

$$w_{i,t} = \beta_1 Age_{i,t}, + \beta_2 Age^2_{i,t} + \beta_3 Bc_{i,t} + \beta_4 Wc_{i,t} + \gamma_1 Dens_{j,t} + \gamma_2 MP_{j,t} + \gamma_3 Spec_{j,\ s,\ t} + \mathbf{i}_s + \delta_t + \varepsilon_{i,t} \quad \text{(M1)}$$

where subscript i refers to individuals, t to time, j to location, and s to sector. $Bc_{i,t}$ and $Wc_{i,t}$ are dummies referring (respectively) to blue collar and white collar occupations (with the residual category being manager), while $\mathbf{i}_s$ (δ_t) is a sector (time) dummy.[10]

Column 1 of Table 2 gives our OLS estimates. Standard individual covariates have all the expected sign and we discuss them no further, nor do we include industry or time fixed effects in the table. As for spatial covariates, we find that doubling density increases wages by 2.21%. Previous findings by Ciccone and Hall (1996) for the US and Combes, Duranton and Gobillon (2007) - CDG (2007) henceforth - for France indicated (respectively) something in the region of 5% and 3%. Our rather low value is probably due to the fact that Italy is characterized by a centralized wage setting where, within each sector, contracts have to respect several nationally based constraints like the minimum wage. Nevertheless, it is worth noting that firms are allowed to supplement the national contract with a company-specific contract. Furthermore, individual bargaining can take place between employers and employees, allowing for further wage variability. With regard to market potential, the OLS estimates

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

suggest that doubling market potential leads to a 10.88% increase in wages. Interestingly, in their aggregate analysis of the impact of market potential on European Union wages, Head and Mayer (2006) find a very similar result.

To gain some insights as to the extent to which spatial externalities are due to spatial sorting of workers we need to introduce a proxy for skills in the analysis. We thus estimate the model (M2), which contains individual effects u_i that we allow to be correlated with spatial covariates:

$$w_{i,t} = \beta_1 Age_{i,t}, + \beta_2 Age^2_{i,t} + \beta_3 Bc_{i,t} + \beta_4 Wc_{i,t} + \gamma_1 Dens_{j,t} + \gamma_2 MP_{j,t} + \gamma_3 Spec_{j,\ s,\ t} + \mathbf{i}_s + \delta_t + u_i + \varepsilon_{i,t}. \tag{M2}$$

Taking into account individual effects reduces our estimates. As shown in the within estimates of column (2), the density coefficient falls to 0.74%, suggesting that sorting of skills is at work. Indeed, our within estimations provide a (significant) correlation of 0.21 between the u_i and the density. As for market potential, the within estimates push down elasticity to 5%. The (significant) correlation between the u_i and market potential is 0.09, which is substantially lower than the correlation with density but still suggestive of a positive link between skills and those agglomeration externalities stemming from NEG models. This evidence strongly suggests that skilled individuals are over-represented in locations characterized by high density and/or high market potential.

In model (M3) we introduce firm heterogeneity. This is, to our knowledge, the first empirical framework dealing with the joint individual and firm content of spatial externalities, using individual wage data. We account for firm heterogeneity by means of (log) firm size (FS), a key variable of the spatial literature dealing with productivity differences across space:

$$w_{i,t} = \beta_1 Age_{i,t}, + \beta_2 Age^2_{i,t} + \beta_3 Bc_{i,t} + \beta_4 Wc_{i,t} + \gamma_1 Dens_{j,t} + \gamma_2 MP_{j,t} + \gamma_3 Spec_{j,\ s,\ t} + \mathbf{i}_s + \delta_t + u_i + \beta_5 FS_{f,t} + \varepsilon_{i,t}. \tag{M3}$$

where subscript f refers to a specific firm.

From Column (3) of Table 2 it comes out that the firm size elasticity with respect to wages is equal to 1.94%. As for the impact on spatial variables, including 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 falls most sharply. However, compared to the sorting of individuals, the sorting of firms entails a much weaker impact on the estimates of spatial externalities.

One might argue that firm size only partially accounts for firm heterogeneity. One way to deal with this issue is to introduce firm fixed effects along with individual fixed effects. This model was developed by Abowd et al. (1999) - AKM (1999) henceforth - and further refined in Abowd et al. (2002), providing the tools to compute the exact solution of the estimation procedure. In particular, the model allows for estimation of time-invariant unobservable firm and individual fixed effects, potentially correlated with covariates. Using this procedure we estimate model (M4), in which we drop sectoral dummies which are no longer identifiable:[11]

$$w_{i,t} = \beta_1 Age_{i,t}, + \beta_2 Age^2_{i,t} + \beta_3 Bc_{i,t} + \beta_4 Wc_{i,t} + \gamma_1 Dens_{j,t} + \gamma_2 MP_{j,t} + \gamma_3 Spec_{j,\ s,\ t} + \delta_t + u_i + \psi_f + \varepsilon_{i,t} \tag{M4}$$

Column (4) of Table 2 shows that the density coefficient drops to 0.14%, and is no longer significant, while the market potential coefficient is still significant although it falls to 2.17%. One tempting interpretation of these results is that wage differentials across provinces with different densities are entirely related to sorting of skills and firms. On the other hand, sorting does not completely explain the market potential externalities, leaving some room for local interactions among firms and individuals affecting wages.

However, the AKM model works best with very large datasets that include sufficient job mobility. Actually, the identification of individual and firm fixed effects in this procedure is based on job changes, which also represent the main source of identification of our spatial externalities should they entail job-location changes. This means that, given the relatively small size of our sample, much of the variability that identifies spatial variables is killed by firm-effect estimation. What essentially identifies the spatial variable coefficients in model (M4) is thus the time variability of density and market potential together with location changes occurring within the same firm.

Furthermore, it is not entirely clear how these firm-fixed effects are to be interpreted. On the one hand, they could reflect heterogeneity in firm productivity, which represents an important component of spatial externalities theory, while, on the other hand, they might be related to other factors, such as union power within firms (Kramarz, 2007), which are not related to spatial issues. For all these reasons, we consider model (M3) as the preferred specification that accounts for both skill and firm heterogeneity.

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

[11] In particular, we use the STATA procedure FELSDVREG developed by Cornelissen (2006). We keep all groups in the estimates, meaning by group a set of workers and firms connected by a sufficient number of job changes, to identify separately all individual and firm fixed effects. This implies that firm and individual effects are not comparable across groups, making the overall correlation between them meaningless (see Abowd, Creecy, and Kramarz, 2002). In our estimations we have 24,353 individuals, 28,719 firms, and 15,186 groups. With so many groups, most of the variability is exploited to estimate the firm effects, leaving little variation for the identification of spatial variables.

much attention to them.[12]

4.2 Endogeneity and sample selection issues

In this subsection we explore the issue of endogeneity and sample selection for our preferred model (M3). As for endogeneity, the reliability of computed elasticities is in fact conditional upon the validity of the underlying moments' restrictions. In particular, it is assumed that $\text{Cov}(\varepsilon_{i,s}, \mathbf{X}_{i,t})$, where $\mathbf{X}_{i,t}$ represents the vector of all covariates, is equal to zero $\forall\, s, t$. However, as pointed out by CDG (2007), 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 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 in column (5) of Table 2 by means of IV estimates exploiting the idea of Ciccone and Hall (1996) that use deeply lagged values of the endogenous variables as instruments. 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 in equation (2) aggregate disposable income of a province j by its population, for the years 1861, 1881, and 1901.[13]

The use of deeply lagged levels of specialization, density and market potential responds to the logic that, as long as early patterns of agglomeration do not reflect factors that influence productivity today, they can be used as instruments. In 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 prior to World War I, immediately after the Civil War, and at the very 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's expedition in 1860 and the railroad network did not really develop until the late 19th century. Therefore, agglomeration was, by that time, the outcome of a rural society in which both changes in transportation technology and manufacturing growth were just about to begin reshaping the distribution of economic activities. A broader perspective on the quality of such super-lags as instruments for density and market potential is provided in Combes et al. (2007). In particular they show that, for both wage and firm TFP spatial regressions, super-lags

[12] 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, the coefficients differ only slightly from those in Table 2.

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

perform at least as well as other instruments (like detailed soil and climate information) that are a priori more closely tied to the local determinants of the old rural population.

Table A1 in the appendix provides the core information on first stage regressions. For each of the three endogenous variables (specialization, density and market potential) first stage estimates of instruments are all highly significant, with the F-test of the joint significance of instruments performing extremely well. As for the specification tests, the Anderson canonical LR test statistic equals 8,964 with 5 degrees of freedom, the 1% critical value being 15.09. The null hypothesis of under-identification is thus strongly rejected. Furthermore, the Cragg-Donald statistic is very high (1,318), rejecting weak identification. The Stock-Yogo critical threshold for a 5% maximal relative bias is in fact much smaller (13.95). Finally, the Hansen test statistic on over-identifying restrictions equals 1.17 with 4 degrees of freedom, the 1% critical value being 13.28. Therefore, the validity of the instruments is not rejected, and this is quite a strong result considering that, with almost 200,000 observations, the power of the test should be pretty high.

We believe that the IV results of model (M3) are the most reliable estimates of spatial externalities we can provide. As can be seen, accounting for endogeneity slightly alters the magnitude of spatial externalities. Compared to the within estimates in column (3) of Table 2, density goes from 0.56% to 0.20% and is only significant at the 10% level. By contrast, the market potential coefficient does not change much going from 4.53% to 4.64%. This suggests that local economic density is more affected by endogeneity than market potential.

Another possible endogeneity issue concerning model (M3) is that, using within estimations, spatial externalities are mainly identified by job-location changes. However, movers are likely not to be a random sample of the workforce. Actually, job-location changes can be due to workers improving on their previous jobs, workers who have been fired from an ongoing firms, and workers displaced because the firm closed down. These different patterns of movers might generate heterogeneity when investigating mobility issues. A way to cope with this issue is to focus on dismissed workers, as in Dustmann and Meghir (2005). The intuition is simple: if we assume that firm closure is exogenous conditional on observables, workers that have been displaced represent a random sample of the workforce, entailing that their job changes are not related to their past choices. Moreover, we also intend to use this methodology concerning dismissed workers to control for possible sample selection arising from our primary focus on male prime age.

For all these reasons, we consider the wider unrestricted database with males and females aged 18-64. From this sample we select only those workers who underwent job dismissal, and we then take into account their wage dynamics in the first job after dismissal.[14] This procedure leads us to an unbalanced panel of 5,051 individuals corresponding to 16,876 observations. In this framework identification comes essentially from time variation in spatial externalities and from location changes within the same firm. As shown in column (6) of Table 2, the coefficients of spatial externalities are still positive and significant, with a magnitude similar to the within estimation of column (3) for density, while market potential is slightly higher, confirming previous findings.

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

The estimations in Table 2 evidence the existence of sorting of individual skills and firms across space. In this Section we further characterize such sorting.

We start with a simple variance decomposition based on IV estimates of model (M3). Looking at the standard deviations of estimated effects in column (1) of Table 3, it comes out the main source of wage variation in the model is given by skills (u_i). Firm size is also fairly relevant, coming just after the age and time dummies. By contrast, spatial effects (density and market potential) entail a relatively small variation in wages with market potential proving more important. These results confirm the findings of CDG (2007) for the skills and spatial components,[15] although the role of firm heterogeneity is not considered in their analysis.

Turning to correlation between estimated effects, it can be seen from column (2) that wages are indeed positively correlated with both density (0.20) and market potential (0.10). However, the strongest correlations are with skills, occupation and firm size. Columns (3) and (4) show that both skills and firm size are positively related to spatial variables (sorting). These correlations provide the rationale for the reduction of spatial variable coefficients when controlling for individual and firm heterogeneity.

In the following Subsections we further characterize the magnitude and the nature (static *vs* dynamic) of sorting, as well as the interaction between skills and firm size across space. To the best

[14] We consider only the first job because the subsequent jobs might partially suffer from the same endogeneity problems concerning job-location choices. Furthermore, it is worth noting that in our database we have no information about dismissed workers should they become unemployed, or as to whether they enter either public employment or self-employment activities. In particular, after dismissal an individual might exit from the panel, working as self-employed or in the public sector, and then going back into the database of private employees. For these reasons, we drop from our analysis individuals who re-enter the database after more than 6 months as from dismissal, since they probably had other employment spells in that period. Finally, note that since we consider only the first job after dismissal sectoral dummies are not identifiable.

[15] Although the results are not directly comparable because of the two-step nature of the CDG (2007) estimation procedure, in France a somewhat higher share of wage variability is related to space. As already mentioned, the pervasive presence of sectoral minimum wages at the national level and centralized wage setting might account for a higher wage compression in Italy.

of our knowledge, we are the first to provide such results within a wage regression framework using individual data.

5.1 Evidence on sorting of workers' skills

We start with the analysis of sorting related to density. We first split the provinces into low density (LD) and high density (HD), defined with respect to the median of the (time average of) density in our database.[16] We then use the individual fixed effects obtained from IV estimations of model (M3) to compute the average skill level of workers in LD and HD provinces. In particular, we compute two different average skill levels, respectively based on workplace (first column of Table 4) and birthplace (second column).[17] Average skill in LD and HD represents log wage deviations from the overall mean of u_i, which is equal to zero, and can thus be interpreted as approximate percentage wage differences.

As for the average skill level based on workplace, the first column of Table 4 reveals that individuals working in HD provinces are far more skilled (0.0474) than those in LD provinces (-0.0469). The skills gap between average skills in LD and HD corresponds to a 9.43% difference in mean wage between HD and LD provinces, and the standard errors (in brackets) reveal that the gap is highly significant. The whole row spatial variation of wages between LD and HD provinces (grand mean of wages across individuals and time in the two groups) comes to 12.78%. Therefore, almost 75% of the row wage variability across city size is accounted for by the sorting of workers' skills.

In order to analyze the dynamic feature of skill distribution, we present in Table 5 a transition matrix constructed using migrations, defined as the occurrence of a workplace change in the panel. The Table shows the average skills of those workers moving across LD and HD locations together with standard errors (in brackets) and number of movers (in square brackets). Applying Borjas's definition (1987), the dynamic sorting is positive from LD to HD provinces, because the average skill of migrants (0.0028) is significantly higher than that of the population of origin (-0.0469) while being lower than that of the population of destination (0.0474). The reverse is also true: sorting is negative from HD to LD provinces. Furthermore, the migrations within each group (LD to LD and HD to HD) suggest that a sizeable positive sorting exists among HD provinces. These findings are in line with the theoretical background of Borjas (1987), where the interaction between skills and the characteristics of the location of origin and destination determine the positive or negative nature of migrant sorting. Our results further qualify location density as a key element in understanding the sorting of migrants.

[16] 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.

[17] Individuals moving across provinces may actually score on more than one category. The same remark applies to Tables 5 to 9.

However, one may wonder whether our results are robust to extension to long-term migrations. To answer this question we set out in Table 6 the same transition matrix as in Table 5, with the difference that migrations are now defined with respect to the birthplace, i.e., working in a province different from that of the worker's birth. As can be seen, the sorting of migrants is qualitatively identical. However, there is an additional point to be made: the sorting of long-term migrants has little impact on the overall sorting of skills across provinces. Although some of the best blood born in LD move to HD provinces, the average skill gap based on workplace in column (1) of Table 4 is just a little larger than the gap based on birthplace in column (2) of Table 4. Those who are born in either an LD or an HD province and do not move are already sorted in space and, although 15% of the people change location within the working age, this flow has no major impact.

Overall, these results suggest that spatial sorting is a pervasive feature affecting not only static means but also the dynamic evolution of skill distribution. However, much of the sorting comes from the static component (non-movers), while both the static and the dynamic evidence are consistent with large cities being more attractive for skilled workers.

Tables 7, 8 and 9 provide a similar analysis for the sorting of skills by market potential. The sorting of skills by market potential based on workplace - column (1) of Table 7 - corresponds to a 5.89% gap between wages of low market potential (LMP) *vs* high market potential (HMP) provinces. Nevertheless, this 5.89% is to be compared with the 8.16% row wage variation between LMP and HMP provinces. Thus skills still account for almost 75% of row spatial wage variation. As for the sorting of migrants, we obtain similar – although less clear – results to those in the case of density (Table 8). These results are confirmed when looking at migrations based on birthplace (Table 9), which also point out that long-term migrations have a smaller impact on the distributions of skills compared to the findings emerging from the analysis based on density.

5.2 Evidence on the sorting of firm size

Unfortunately, our data on firms are not as detailed as the data on individuals and we cannot analyze the dynamic content (relocations) of the spatial sorting of firms by size. However, we can still quantify the impact on wages of the spatial sorting of firms.

Table 10 shows the average firm size effect, computed using the predicted values of $\beta_5 FS_{f,t}$ in the IV estimation of model (M3), for LD and HD provinces (column 1) and for LMP and HMP provinces (column 2). As previously noted, firm size is positively correlated to both density (0.11) and market potential (0.07). The difference in average firm size effect between LD and HD provinces, in column (1), accounts for a 0.72% wage gap in favor of HD provinces as compared to the 12.78% row spatial variation of wages across the two groups. The share of row spatial wage variation explained by firm sorting (5.6%) is far below the gap accounted for by workers' skills (75%). To the extent that firm size is a good proxy for firm productivity, our results suggest that microfoundations of agglomeration economies aiming at explaining wage differences across space should focus on worker heterogeneity. These findings are confirmed with breakdown of the firm size effect by LMP and HMP provinces. Firm size distribution differences across provinces imply a 0.45% wage advantage for HMP provinces to be compared with the 8.16% row wage variation.

5.3 Assortative Matching Across Space

Our data suggest that firms and individuals are sorted across space. However, as we are able to match workers and firms, we can also investigate assortative matching patterns, i.e., whether good workers are matched to good firms. A considerable number of studies address the issue of assortative matching. The first empirical paper is AKM (1999). Other important empirical contributions are Abowd et al. (2002) and Postel-Vinay and Robin (2006).[18] We improve on the existing literature by analyzing the degree of assortative matching across space.

Using our preferred IV estimations of model (M3), we actually find that the correlation between skills 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 – information that we do not have – our results support the presence of assortative matching.[19] In this respect, Postel-Vinay and Robin (2002) show that firm size is a valuable proxy in order to recover both the firm hiring effort and productivity.[20]

The assortative matching might in principle just be related to co-location of skilled workers and large firms in locations with high density and/or market potential. However, the correlation is positive and significant in each of the 95 Italian provinces, ranging from 0.08 to 0.68. Moreover, the conditional (with respect to density, market potential and specialization) correlation between the u_i and firm size is only slightly lower than the unconditional one (0.34 compared to 0.35). This means that co-location is not really an issue, suggesting that there is a deeper underlying complementarity between skills and firm size.

[18] From the view point of labor economics 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 an assortative matching.

[19] As for the empirical support of assortative matching, in the existing literature the results largely depend on the proxy used to qualify a good firm. Abowd et al. (2002) use firm fixed effects and find that the correlation with individual effects (skills) is either slightly positive (for the 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, which might prove too scarce to yield 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 positive (0.27).

[20] Note that, since we keep all groups, we cannot use individual and firm fixed effects obtained using the AKM model to compute their correlation across space. These effects are in fact comparable only within a group.

A way to characterize this evidence further is to analyze how the degree of assortative matching varies with respect to the spatial variables, and in particular with density. In this respect, it is worth noting that in our analysis skills are measured by time-invariant effects that are specific to the individual and not to the individual-firm match. Therefore, they may rightly be viewed as workers' attributes related to general human capital that, like ability and education, are time-invariant and enhance productivity irrespective 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 matching between general attributes of individuals and firms. 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 can be seen, this relation is negatively sloped and the correlation between the two variables (-0.39) is actually significant at 1% confidence level. Two possible robustness checks can be applied to investigate further this intriguing negative correlation between assortative matching and density.

First, one might argue that individual fixed effects imperfectly capture general human capital attributes, since their estimates are also affected by match-specific factors. In order to mark out the impact of match-specific effects we restrict our analysis to the sub-sample of 10,965 workers that change firm in the estimation period. In this sub-sample, our measure of skills is more strictly related to general human capital attributes that do not change across different matches. When considering this sub-sample, the correlation between density and local assortative matching is still negative (-0.40) and significant.

Second, another possible explanation might refer to 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. Interestingly enough, we find that this explanation partially holds, since the correlation between firm size and skills reduces from -0.39 to a significant and still negative -0.16, suggesting that the introduction of province-specific firm size coefficents alters the computation of individual fixed effects, in turn affecting the degree of assortative matching.

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 an original finding of our paper, is to consider the horizontal dimension of skills and productivity, as in Helsley and Strange (1990). As long as firm-worker specific attributes are taken into account, then the larger the size of the market, the higher is the probability for a worker/firm of finding a partner with some specific characteristics that are valuable assets for the match. Therefore, general attributes should become relatively less important with larger markets. In this line, Kim (1989) proposes a theoretical framework that combines both general and specific skills with market size. Kim (1989) shows that when the size of the market increases, workers invest more in specific human capital than in general human capital, because they can more easily find a firm that appreciates their specific asset.

6 Conclusions

In this paper we show that using individual data is crucial in order to penetrate the black box of agglomeration externalities. Our results provide evidence that the spatial sorting of firms and (especially) skills accounts for a large portion of the raw spatial wage variation. Exploiting the matched employer-employee 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 characterize this result further, showing that the assortative matching is negatively related to the size of local labor market.

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, and to encourage deeper discussion on the interactions between skills and firm heterogeneity, at both 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., 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/~jma7/abowd − creecy − kramarz − computation.pdf*

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

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

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

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

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

[9] Combes P.P., G., Duranton, and L., Gobillon (2007). "Spatial wage disparities: Sorting matters!". Forthcoming Journal of Urban Economics.

[10] Combes P.P., G., Duranton, L., Gobillon , and S. Roux (2007). "Estimating Agglomeration Effects: Does Playing with Different Instruments Give a Consistent Tune? ". Mimeo

[11] Cornelissen, T. (2006), "Using STATA for a memory saving fixed effect estimation of the three way error component model", FDZ working paper, no.3/2006.

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

[13] 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.

[14] Dustmann, C., Meghir, C. (2005), "Wages, Experience and Seniority", Review of Economic Studies, Vol. 72, No. 1, pp. 77-108.

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

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

[17] 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).

[18] Hanson, G.H. (2005). "Market potential, increasing returns, and geographic concentration". Journal of International Economics Vol., 67 1-24..

[19] 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.

[20] Head, K., and T., Mayer (2006). "Regional Wage and Employment Responses to Market Potential in the EU". Regional Science and Urban Economics Vol. 36(5), 573-595.

[21] 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.

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

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

[24] Kim, Sunwoong (1989). "Labor specialization and the extent of the market". Journal of Political Economy Vol. 97(3), 692Ű705.

[25] Kramarz, F. (2007). "Outsourcing, Unions, Wages, and Employment: Evidence from Data Matching Imports, Firms, and Workers, ". CREST working paper.

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

[27] Marshall, Alfred. "Principles of Economics". 1890. London: Macmillan.

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

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

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

[31] 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.

[32] Postel-Vinay, F. and J.M., Robin (2006). "Microeconometric 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.

[33] 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.

[34] 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.

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

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

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, consistently 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 variable ln(wage)

	(1)	(2)	(3)	(4)	(5)	(6)
Age	0.0481***	0.0456***	0.0466***	0.0686***	0.0455***	0.0584***
	(0.0015)	(0.0080)	(0.0080)	(0.00715)	(0.0079)	(0.0028)
Age^2	-0.0005***	-0.0005***	-0.0005***	-.0004***	-0.0005***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.000)
Firmsize			0.0194***		0.0193***	0.0080**
			(0.0004)		(0.0004)	(0.0041)
Bc Dummy	-0.7619***	-0.2132***	-0.2149***	-0.1760***	-0.2114***	-0.1285***
	(0.0049)	(0.0041)	(0.0041)	(0.0040)	(0.0041)	(0.0141)
Wc Dummy	-0.4891***	-0.1452***	-0.1466***	-0.1237***	-0.1461***	-0.0487***
	(0.0049)	(0.0031)	(0.0031)	(0.0028)	(0.0031)	(0.0072)
Specialization	0.0055***	0.0015*	0.0008	0.0008	-0.0037	0.0110***
	(0.0006)	(0.0008)	(0.0008)	(0.0009)	(0.0035)	(.0029)
Density	0.0221***	0.0074***	0.0056***	0.0014	0.0020*	0.0061*
	(0.0005)	(0.0010)	(0.0011)	(0.0014)	(0.0011)	(0.0037)
Market Potential	0.1088***	0.0500***	0.0453***	0.0217***	0.0464***	0.1193***
	(0.0021)	(0.0058)	(0.0058)	(0.0077)	(0.0134)	(0.0240)
Estimated model	(M1)	(M2)	(M3)	(M4)	(M3)	(M3)
Estimation mehtod	OLS	Within	Within	AKM	IV-Within	Dismissed
Time & Sector Dummies	Yes	Yes	Yes	Time only	Yes	Time only
Individual fixed effects	No	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	No	No	No	Yes	No	No
N. of individuals	24353	24353	24353	24353	24353	5051
N. of firms	28719	28719	28719	28719	28719	2778
N. of observations	175700	175700	175700	175700	175700	16876

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

Table 3. Summary statistics for the variance decomposition - IV estimation of model (M3)

		Std dev	Simple correlation with: (w)	(u)	(β_5 FS)
Effect of		(1)	(2)	(3)	(4)
log real wage	(w)	**0.3625**	1.000	0.854	0.473
residuals	(ε)	0.1097	0.303	0.000	0.000
worker effects	(β_1 Age+β_2 Age^2+β_3 Bc+β_4 Wc) + u	0.2918	0.901	0.974	0.375
age	(β_1 Age+β_2 Age^2)	0.0554	0.340	0.081	0.131
occupation	(β_3 Bc+β_4 Wc)	0.0384	0.551	0.482	0.284
worker fixed-effects	(u)	**0.2612**	**0.854**	1.000	**0.349**
time dummies	(δ_t)	0.0890	0.316	-0.005	0.022
sectoral composition	(**i**+γ_3 Spec)	0.0424	0.262	0.129	0.168
sectoral dummies	(**i**)	0.0417	0.260	0.129	0.154
specialization	(γ_3 Spec)	0.0038	0.068	0.015	0.183
firm size	(β_5 FS)	**0.0532**	**0.473**	**0.349**	1.000
spatial effects	(γ_1 Dens+ γ_2 MP)	**0.0141**	0.133	0.119	0.085
density	(γ_1 Dens)	**0.0024**	**0.203**	**0.213**	**0.110**
market potential	(γ_2 MP)	**0.0133**	**0.103**	**0.087**	**0.071**

175,700 observations. All correlations between the effects that are not orthogonal by definition are significant at 1%.

Table 4. Average skills distribution across provinces based on either workplace or birthplace density

	Average skills by workplace	Average skills by birthplace
LD provinces	-0.0469*** (0.0021)	-0.0329*** (0.0023)
HD provinces	0.0474*** (0.0023)	0.0469*** (0.0028)

Standard errors in parenthesis with ***, ** and * respectively denoting significantly different from zero at the 1%, 5% and 10% levels.

Table 5. Analysis of migrants' skills: transition matrix based on workplace density

	LD provinces (destination)	HD provinces (destination)
LD provinces (origin)	-0.0614*** (0.0076) [1184]	0.0028 (0.0083) [1037]
HD provinces (origin)	-0.0009 (0.0083) [1042]	0.1089*** (0.0102) [881]

Standard errors in parenthesis with ***, ** and * respectively denoting significantly different from zero at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 6. Analysis of migrants' skills: transition matrix based on birthplace density

	LD provinces (destination)	HD provinces (destination)
LD provinces (origin)	-0.0407*** (0.0054) [2235]	0.0513*** (0.0049) [2839]
HD provinces (origin)	0.0121 (0.0089) [885]	0.1170*** (0.0079) [1278]

Standard errors in parenthesis with ***, ** and * respectively denoting significantly different from zero at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 7. Average skills distribution across provinces based on either workplace or birthplace market potential

	Average skills by workplace	Average skills by birthplace
LMP provinces	-0.0279*** (0.0024)	-0.0294*** (0.0025)
HMP provinces	0.0310*** (0.0023)	0.0393*** (0.0027)

Standard errors in parenthesis with ***, ** and * respectively denoting significantly different from zero at the 1%, 5% and 10% levels.

Table 8. Analysis of migrants' skills: transition matrix based on workplace market potential

	LMP provinces (destination)	HMP provinces (destination)
LMP provinces (origin)	-0.0497*** (0.0075) [1251]	0.0374*** (0.0103) [817]
HMP provinces (origin)	0.0395*** (0.0106) [722]	0.0416*** (0.0083) [1187]

Standard errors in parenthesis with ***, ** and * respectively denoting significantly different from zero at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 9. Analysis of migrants' skills: transition matrix based on birthplace market potential

	LMP provinces (destination)	HMP provinces (destination)
LMP provinces (origin)	-0.0143*** (0.0051) [2644]	0.0185*** (0.0059) [2104]
HMP provinces (origin)	0.0807*** (0.0123) [501]	0.0912*** (0.0063) [1906]

Standard errors in parenthesis with ***, ** and * respectively denoting significantly different from zero at the 1%, 5% and 10% levels. Number of migrants are in square brackets

Table 10. Distribution of the firm size effect (β_5 FS) by the density and market potential of firms' location

	Average firm size effect by density (1)		Average firm size effect by market potential (2)
LD provinces	0.0607*** (0.0003) [15835]	LMP provinces	0.0618*** (0.0003) [14640]
HD provinces	0.0679*** (0.0003) [13818]	HMP provinces	0.0663*** (0.0003) [14903]

Standard errors in parenthesis with ***, ** and * respectively denoting significantly different from zero at the 1%, 5% and 10% levels. Number of firms is in square brackets

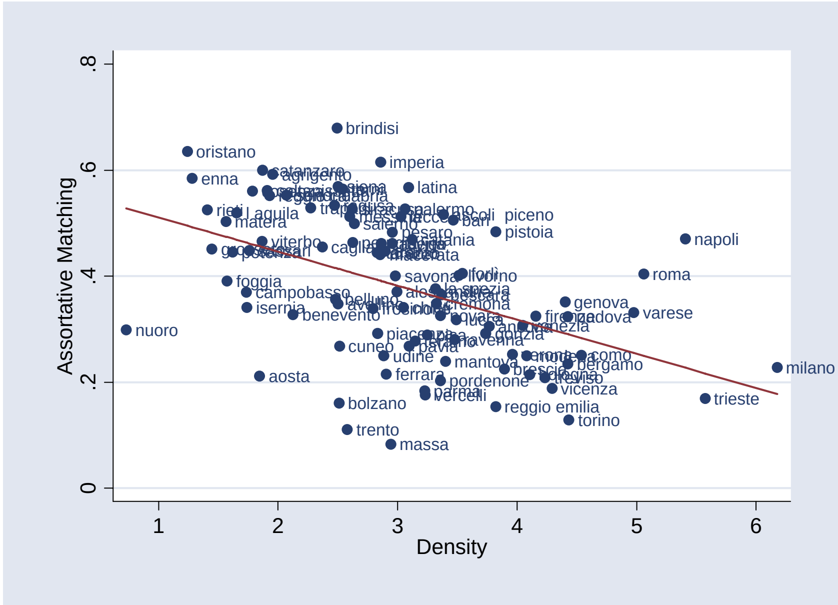

Figure 1: Assortative matching and market density

Appendix

Table A1. First stage regressions - IV estimates

Dependent variable	Specialization	Density	Market pot.
Age	0.0580**	0.0020	0.0014
	(0.014)	(0.800)	(0.638)
Age^2	0.0000	-0.0000**	0.0000
	(0.535)	(0.012)	(0.566)
Firmsize	0.01012***	0.0079***	0.0015***
	(0.000)	(0.000)	(0.000)
Bc Dummy	0.0501***	0.01456***	0.0003
	(0.000)	(0.000)	(0.869)
Wc Dummy	-0.0329***	0.0101***	-0.0017
	(0.000)	(0.001)	(0.160)
Specialization 1951	0.2585***	0.0080***	0.0063***
	(0.000)	(0.000)	(0.000)
Density 1861	0.0559*	0.0199***	-0.1744***
	(0.072)	(0.060)	(0.000)
Density 1881	-0.2813***	-4.2938***	0.9331***
	(0.000)	(0.000)	(0.000)
Density 1901	0.1538**	5.3926***	-0.7133***
	(0.005)	(0.000)	(0.000)
Market potential 1861	0.4065	-5.9075***	-5.7568***
	(0.519)	(0.000)	(0.000)
Market potential 1881	3.3422***	30.6436***	11.4181***
	(0.001)	(0.000)	(0.000)
Market potential 1901	-3.552***	-25.9742***	-4.5686***
	(0.000)	(0.000)	(0.000)
Time & Sector Dummies	Yes	Yes	Yes
Partial R^2 excluded instruments	0.058	0.824	0.190
Value F test excluded instrument	212.93	6724.44	279.71
Test excluded instrument (p-value)	0.000	0.000	0.000
N. of observations	175700	175700	175700

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