Urbanization Externalities, Market Potential and Spatial Sorting of Skills and Firms*

Giordano Mion† and Paolo Naticchioni‡

February 2005

Abstract

Using an individual panel data on Italian workers we look at the spatial distribution of wages among provinces. Even when controlling for individual characteristics and differences in local production structure we find evidence of both urbanization externalities and market potential. Although these two forces are, to some extent, not separately identifiable our results suggest that market potential has a stronger impact on wages. Moreover, spatial sorting of skills is at work in the sense that "good" workers are attracted by location with high density and/or good access to consumers' markets. Sorting actually dampens both estimates. Another issue we deal with is the relation between the firm-size premium, individual skills and location. Our results suggests, coherently with Abowd, Kramarz, and Margolis (1999), that the correlation between the the size of the employing firm and skills is very strong. However, this correlation is not simply the outcome of a co-location phenomenon, suggesting that there is a deeper underlying economic relation like the one traced in Yeaple (2005).

Keywords: Spatial Externalities, Panel-data, Skills, Firms' heterogeneity, Sorting.

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 "La Sapienza" of Rome for the access to the INPS Italian database. We are also grateful to Pierre Philippe Combes, Bart Cox, Keith Head, Emilano Rustichelli, Eric Strobl and AIEL Conference participants for helpful comments and suggestions, and to Andrea Ichino, Antonietta Mundo, Renato Santelia. for providing us with data on spatial variables.

†Giordano Mion: University of Bologna, Address: Dipartimento di Scienze Economiche, Strada Maggiore 45, 40125 Bologna, Italy. Email: mion@spbo.unibo.it.

‡Paolo Naticchioni, Università di Roma "La Sapienza", Universitè Catholique de Louvain (IRES). 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 space. 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 first in 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 country-based studies, like for instance Combes, Duranton and Gobillon (2004), show that wages vary considerably across space. Moreover, looking at manufacturing sectoral earnings, Amiti and Cameron (2004) provide evidence of large and persistent spatial disparities in Indonesia.

So far, many explanations have been put forward to explain such imbalances, concerning for instance the availability of natural resources, infrastructures, technology, the impact of crime, cultural and social factors, etc.[2] In this paper we focus, although not exclusively, on the role of spatial externalities and skills using information on individual wages coming from a matched employer-employee database for Italy. More specifically, we work on panel version of an administrative database provided by INPS (the Italian Social Security Institute), in which we can follow workers overtime, merge information of individuals and firms characteristics, and follow workers when they migrate from one location to another. The advantage of using such a rich database is that we can control for both observable and non-observable individual characteristics that may (and actually they do) interact with spatial externalities.

The starting point of our analysis is a spatial equilibrium model in which two types of externalities are considered simultaneously and subsequently estimated. The first one, which is more rooted in the literature, refers to the positive impact of density on local economic performances (urbanization externalities). The idea that the agglomeration of economic activities fosters local productivity and growth comes back to Marshall (1890). These ideas have been formalized (among others) by Abdel Rahman and Fujita (1990) and have been the object of an intensive applied research.[3] The second source of spatial disparities that we consider here are the (relatively new) pecuniary externalities stemming from increasing returns to scale, transportation costs and

[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 Beeson and Eberts (1989) and Moretti (2004) among others.

[3]See Sveikauskas (1975), Ciccone and Hall (1996) and Ciccone (2002) among others.

proximity to demand. The new economic geography literature (NEG), started with the work of Krugman (1991), has provided a collection of general equilibrium models, explicitly dealing with space, and capable to account for many salient features of the economic landscape.[4] In particular, a core-periphery structure endogenously emerges in these models leading to higher wages in those locations that provide a better access to demand (usually referred as market potential).

One contribution of this paper is the assessment of the absolute and relative importance of density and market potential externalities within a unified model-based framework. In fact, the existing literature either focus on one single externality,[5] or use some proxy for the other (like in Amiti and Cameron, 2004) just as a control.[6] To this respect our results suggest that, even after controlling for individual characteristics and endogeneity, both density and market potential have a positive and significant impact on wages. Although, as we will show later on, these two externalities cannot be fully disentangled, our estimates further indicate that market potential has the strongest effect. Both the nominal and the standardized elasticity of density are in fact considerably lower than those of market potential. This suggest that, at least for Italy, pecuniary externalities play a crucial role in the spatial distribution of wages. These findings are coherent with those of Mion (2004) who finds evidence of sizeable agglomeration externalities in Italy.

The interplay between individual skills and spatial externalities is also a noticeable contribution of our paper. That fact that skills' distribution may be at the hearth of systematic wage differences is well-known in labour economics. However, most of the research has focused on wage differential across industries rather than across space.[7] Glaeser and Mare (2001), and Duranton and Monastiriotis (2002) are relevant exceptions. However, the closest reference to our framework is certainly the paper of Combes, Duranton, and Gobillon (2004) that has been extremely inspiring for much of our analysis. Using a very similar dataset, the authors show that skills explain a great deal of the observable variation in French wages. Moreover, they are highly correlated with location-specific variables and in particular with economic density. Skills are thus sorted in space and this dampens estimates of the elasticity of wages with respect to density. Our results confirms the fact that accounting for skills reduces the impact of spatial externalities, and further identify a positive link with market potential.

[4] See Fujita, Krugman and Venables (1999), and Fujita and Thisse (2002) for a review of the literature.

[5] See Combes and Lafourcade (2001), Mion (2004), Redding and Venables (2004), and Hanson (2005)

[6] Combes, Duranton and Gobillon (2004) try to put some NEG features in their analysis using the mean of the log of density in adjacent areas as a proxy for market potential. However, the equation they estimated is not a reduced form of a structural model and the proxy they use can arguably be related to market potential.

[7] See Gibbons and Katz (1992) and Abowd, Kramarz, and Margolis (1999).

The only framework that deals with market potential and education (which is the closest proxy to skills) is Redding and Schott (2003). The authors develop a NEG model in which they show that, 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 reduces incentives for human capital accumulation. The positive correlation between market potential and skills that we find is thus coherent with their analysis. The issue of skilled workers and cities has instead received more attention from the profession. The fact that skilled workers are disproportionately distributed in big cities can be explained in at least three ways. The first possibility is that cities provide higher returns to education (Moretti, 2004) so that residents find it profitable to invest more in the accumulation of human capital. Second, Glaeser and Mare (2001) discuss a second dynamic mechanisms in which cities foster the accumulation of human capital. Finally, there 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 valuable assets for them, they would prefer cities. It is important to stress the first two explanations of the spatial sorting do not necessarily require migrations of skilled workers in cities. However, when looking at migrants from small to large cities in our panel, we do not find any strong evidence that they have higher skills than the average. Heterogeneous returns to education and faster human capital accumulation in cities thus seems to be a better explanation for spatial sorting.

A final issue that we deal with is the relation between firm-size, skills and location: our paper is, to our knowledge, the first framework dealing with the connection between these three features. According to the recent literature on heterogeneous firms, and in particular with Melitz and Ottaviano (2004), bigger and more productive firms should be attracted by local market size. Indeed, in our data the size of firms is positively correlated with density; a result that has already be documented by Campbell and Hopenhayn (2002). However, when considering workers' heterogeneity, Yeaple (2005) shows that firms who choose to be high tech are endogenously bigger and have a skill-biased technology. Consequently, big size firms should be observed to hire more skilled workers and, even after controlling for skills, to have a residual productivity premium. This in indeed the case in our data and a similar result have been found in Abowd, Kramarz, and Margolis (1999). However, the fact that we specifically deal with space allow us to go a bit further. Actually, the link between skills and firm size may be in principle due to a simple co-location effect of skilled workers and big firms in cities. However, we show that there exist a strong correlation between skills and plant-size *conditionally* on spatial characteristics like density and market potential (that are precisely those that are linked to co-location) suggesting that there is a deeper underlying economic relation like the one traced in Yeaple (2005).

The rest of the paper is organized as follows. Section 2 presents our theoretical model, which comes form the analytical framework of the NEG, while in section 3 we point out how to estimate density and market potential. In section 4 we present the data and descriptive statistics of our data. Section 5 is devoted to the econometric analysis, in which we also deepen the endogeneity issue. In section 6 some additional results are presented. Finally, conclusions are reported in section 7.

2 The Model

In this Section, we present an enriched version of Krugman (1991) and Helpman (1998) models in which we introduce urbanization economies and consider more than two locations. The model represents the theoretical ground on which we will construct the econometric analysis.

Imagine an economy consisting of Φ locations, two sectors (the manufacturing sector M and the housing sector H), and one production factor (labor). The M-sector produces a continuum of varieties of a horizontally differentiated product under increasing returns to scale, using labor as the only input. Each variety of this differentiated good can be traded among locations incurring in iceberg-type transportation costs.[8] Referring to two generic locations as j and k ($j, k = 1, 2, ..., \Phi$), we thus have that for each unit of good shipped from j to k, just a fraction $v_{j,k} = T(d_{j,k})$ of it arrives to destination, where $d_{j,k}$ is distance between the two locations and $T(.)$ is a decreasing function. The H-sector provides instead a homogeneous good, housing, that cannot be traded and whose amount in each location (H_j) is supposed to be exogenously fixed. Its price $P_{H,j}$ can therefore differ from one place to another and is determined by the equilibrium between local supply and demand.

Labor is supposed to be freely mobile, and its (exogenous) total amount in the economy is equal to L. The equilibrium spatial distribution of our workers-consumers is thus determined by both wages (w_j), and prices prevailing in each location. We will denote L_j, with $\sum_{j=1}^{\Phi} L_j = L$, as labor in location j, and $\lambda_j = L_j/L$ as the corresponding share of total workers.

[8] The term transportation costs does not simply refer to shipment costs but in general to all costs and impediments of doing business in different markets, like information costs, language differences, etc.

Preferences do not directly depend upon the location where consumption and production take place, but only indirectly through prices. As usual in NEG models, preferences are described by the standard Cobb-Douglas utility function with CES type sub-utility for the differentiated product, i.e.:

$$U = (C_M)^{\mu} (C_H)^{1-\mu} \qquad 0 < \mu < 1 \tag{1}$$

where C_M stands for an index of the consumption of the M-sector varieties, while C_H is housing consumption. We assume that the modern sector provides a continuum of varieties of (endogenous) size N, the consumption index C_M is thus given by:

$$C_M = \left[\int_0^N c_m(s)^{\rho} ds \right]^{1/\rho} \qquad 0 < \rho < 1 \tag{2}$$

where $c_m(s)$ represents the consumption of variety $s \in [0, N]$. Hence, each consumer has a love for variety and the parameter $\sigma \equiv 1/(1-\rho)$, varying from 1 to ∞, represents the (constant) elasticity of substitution between any two varieties. If Y denotes the consumer income, then the demand function for a variety s coming from utility maximization is:

$$c_m(s) = p_m(s)^{-\sigma} \mu Y (P_M)^{\sigma-1} \qquad s \in [0, N] \tag{3}$$

where $p_m(s)$ is here the consumer-price (or delivered price) of our generic variety and P_M is the price-index of the differentiated product given by:

$$P_M \equiv \left[\int_0^N p_m(s)^{-(\sigma-1)} ds \right]^{-1/(\sigma-1)} \tag{4}$$

Technology is, by contrast, not the same across locations. Each variant of the differentiated product needs labor to be produced. The relation between the amount of labor used ($l_j(s)$) and the quantity of variant s produced ($c_j(s)$) is given by:

$$l_j(s) = f + \beta_j c_j(s) \tag{5}$$

where f and β_j are, respectively, the fixed and the marginal labor requirements. The fixed component is identical across space while, contrary to the standard formulation of Krugman (1991) and Helpman (1998), the marginal one is supposed to depend on the density of economic activities: $\beta_j = L_j^{-\eta}$. The idea that market size has a positive impact on local productivity comes back to Marshall (1890) and have been formalized by Abdel Rahman, and Fujita (1990) and Schulz and Stahl (1996) among others.[9] Ciccone and Hall (1996) and Combes (2000) provide a strong evidence in favour of the positive role of density, and these externalities are often referred to urbanization economies. We decided to introduce a density effect in the production function in order to derive the implication of this assumption for the reduced-form equation of the model that we will subsequently estimate.

Firms know consumers' demand and choose prices in order to maximize their profits given by:

$$\pi_j(s) = p_{m,j}(s)c_j(s) - w_j[f + \beta_j c_j(s)] \tag{6}$$

where w_j is wage paid by our generic firm and $c_j(s)$ is its output. However, when they look at demand structure, i.e. equation (3), they consider Y_j and $P_{M,j}$ as given. Since each of them has a negligible influence on the market, it may accurately neglect the impact of a price change over both consumers' income and the price index. Consequently, (3) implies that each firm faces an isoelastic downward sloping demand with elasticity given by our parameter σ. Solving first order conditions yields the usual equilibrium relation between the optimal price, elasticity of demand, and marginal cost:

$$p_{m,j}(s) = \frac{w_j \beta_j}{1 - (1/\sigma)} \tag{7}$$

Under free entry, profits are zero. This implies, together with equation (7), that the equilibrium output is:

$$c_j(s) = (\sigma - 1)f/\beta_j \tag{8}$$

In equilibrium a firm's labor requirement is unrelated to firms' distribution. In fact, using equation (5) one gets:

$$l_j(s) = l = \sigma\ f \tag{9}$$

[9] There are 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. See Duranton and Puga (2004) for a discussion of the microfoundations of agglomeration economies.

Now, combining equations (3), (7), and (9) we finally obtain the following reduced-form equation for wages that will be the theoretical basis for our estimations (in logarithm):

$$\ln(w_j) = \ln(\kappa) \; + \; \frac{(\sigma-1)\eta}{\sigma} \ln L_j \; + \; \frac{1}{\sigma} \ln \left(\sum_{k=1}^{\Phi} Y_k (P_{M,k} \, v_{j,k})^{\sigma-1} \right) \tag{10}$$

with $\kappa \equiv \rho \, [\mu/(\sigma-1)f]^{1/\sigma}$.

3 Market Potential vs Urbanization Externalities

Equation (10) states that nominal wages in location j depends positively on the local economic density (L_j) and on the weighted sum over space of incomes (Y_k) and prices ($P_{M,k}$) of all locations with weights inversely related to distance ($v_{j,k} = T(d_{j,k})$). The density related component stands for urbanization externalities. It has thus a fully local nature, and can be easily measured by employment or population density like in Ciccone and Hall (1996) and Combes (2000). The spatially weighted part is the counterpart of those pecuniary externalities, stemming from transportation costs, product differentiation and increasing returns to scale, which leads to agglomeration of economic activities in NEG models. This term is much more tricky to deal with as it contains the local price variables $P_{M,k}$ for which proper data do not generally exist. One possible solution, used by Head and Mayer (2004) and Redding and Venables (2004), is to use data on bilateral trade flows and location dummies. In equilibrium the total value of exports ($Exp_{j,k} = n_j p_j c_{j,k}$) equals imports and, using consumers' demand (3) to get $c_{j,k}$, one obtains:

$$Exp_{j,k} = n_j p_j x_{j,k} = n_j \; {p_{m,j}}^{1-\sigma} \; \mu Y_k (P_{M,k})^{\sigma-1} v_{j,k}^{(\sigma-1)} = s_j \; m_k \; v_{j,k}^{(\sigma-1)} \tag{11}$$

where $s_j = n_j \; {p_{m,j}}^{1-\sigma}$ ($m_k = \mu Y_k (P_{M,k})^{\sigma-1}$) depends only on $j(k)$ and can be estimated as a location specific fixed effect in a gravity regression of export flows over a measure of transportation costs $v_{j,k}^{(\sigma-1)}$. Given the availability of trade flows, it would thus be possible to recover the m_k corresponding to each province, which gives the information we need on the average prices and income. Unfortunately, this methodology cannot be applied in our case because, as far as we know, there are no data on commodity flows among Italian provinces even at a very (sectoral) aggregate level. However, even if these data were available, intra-national commodity flows does not necessarily correspond to actual trade among regions because goods shipped to border areas may just reflect *international* trade.[10]

An alternative strategy, used by Hanson (2005) and Mion (2004), is to assume that mobility of labor is perfect so that real wages equalize:

$$\frac{w_j}{(P_{M,j})^{\mu}(P_{H,j})^{1-\mu}} = \frac{w_k}{(P_{M,k})^{\mu}(P_{H,k})^{1-\mu}} \qquad \forall j,k = 1,2,...,\Phi \tag{12}$$

from which $P_{M,j}$ can be expressed as function of w_j - which is observable - and $P_{H,j}$ that -further using the equilibrium on the housing market- is a function of the local housing stock (H_j) that is also observable. However, this strategy - and in particular underlying assumption (12) - is notably inappropriate in our case for at least two reasons. On the one hand, perfect mobility is at odds with the persistent wage and unemployment rate spatial differentials that characterize the Italian landscape.[11] On the other hand, contrary to Mion (2004), the identification of the role of spatial variables by means of the time variation in the data requires here to properly capture the dynamics of migrations. For example, 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' migration from one location to another.[12] The problem is that migration choices are made not only upon wages but also depending on local amenities, provision of public goods, working opportunities, etc. Therefore, even admitting that mobility is perfect, condition (12), which is based on wages and prices only, is probably too far from reality.

The solution we will actually implement is based on the concept of market potential as originally introduced by Harris (1954), which was developed by the author to measure the "potential" demand for goods and services produced in a location $j = 1,2,..,\Phi$ with an index of a location proximity to consumers' markets given by:

$$MP_j = \sum_{k=1}^{\Phi} Y_k d_{jk}^{-1} \tag{13}$$

where Y_k is an index of purchasing capacity of location k (usually income), $d_{jk}$is the distance between two generic locations j and k. By comparing equations (10) and (13) one can notice that (13) is actually a particularization of the spatially weighted term in (13) where $P_{M,k} = 1 \; \forall \; k$, and

[10] See Anderson and van Wincoop (2003).

[11] See Eichengreen (1993), and Overman and Puga (2002).

[12] 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 (migration free) within variance of density. Finally, we compare this within variance with the no-restricted variance that includes workers' movements.

$v_{j,k}^{\sigma-1} = T(d_{j,k}) = d_{jk}^{-1}$.[13] The main advantage of the Harris' formulation compared to the more rigorous one is that data on prices are not needed. If the scope of the analysis is to structurally estimated the model, like in Hanson (2005) and Mion (2004), then this simplification would be certainly unacceptable. However, our goal here is to give a measure of the magnitude of agglomeration economies, as well as to link them with the heterogeneity of individuals and firms, in the more possible rigorous way without pretending to interpret estimate as the parameters of our underlying model. Furthermore, when comparing Harris' market potential with a more structural measure, Head and Mayer (2004) did not find any strong evidence in favour of the latter in terms of predictive power and magnitude. All this being said, our solution thus seems fairly acceptable for our purposes.

One of the main goal of our paper is to compare the relative strength and magnitude of market potential compared to urbanization externalities. In particular we measure density like in Combes (2000) and Ciccone and Hall (1996) as:

$$Dens_{j,t} = \ln\left[\frac{empl_{j,t}}{size_t}\right] \tag{14}$$

where $empl_{j,t}$ is employment in location j at time t, while $size_t$ is a location surface in square km. As standard, we consider here the log, and we do the same for all the other location variables, in order to interpret parameters as elasticities and to ease the comparison. As for market potential, we have to decide between two alternatives. The first solution consists in taking the log of (13). However, this solution has a couple of drawbacks that come from the fact that, when taking the sum across space, one has to consider for the market potential of j the income coming from region j itself. When considering such a term, there is a first problem of calculating internal distances. One solution, introduced by Head and Mayer (2000), is to assume that locations can be approximated by a disk of area A in which production occurs at the center while consumers are uniformly distributed throughout the rest of the area. This leads to an average distance among consumption and production equal to 0.376 $A^{0.5}$. However, this solution is not universally accepted in literature because it implicitly assumes that the shape on

[13] 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 and Krugman (1995), and 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.

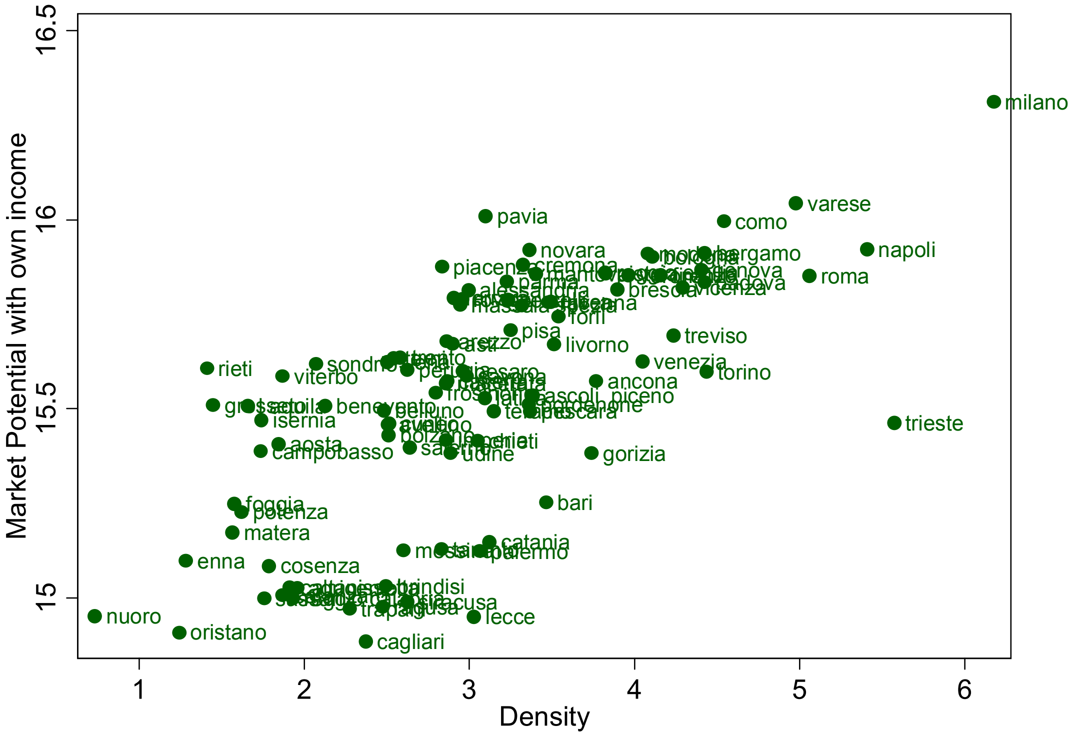

Figure 1: Correlation between density and market potential with own income.

the transportation technology is the same between long (external) and short (internal) distances. Indeed, as shown in Disdiez and Head (2004), the elasticity of trade with respect to distance is higher for intra-continental compared to inter-continental flows.

A second complication arises because of the high correlation between the internal-distance corrected income (that is a kind of density) and the density of employment. Since the two are very close to each other and the weight $(d_{j,j}^{-1} / \sum_{k=1}^{\Phi} d_{j,k}^{-1})$ given to Y_j in equation (13) is fairly high (roughly 0.25 in our data) then our measure of market potential turns out to be strongly correlated with economic density.

This strong link is clearly evident in Figure 1, in which we plot the (time) average density and market potential of the 95 Italian provinces which reports a correlation of around 0.80. One direct consequence of this problem is that including market potential, as defined in equation (13), in regressions makes the coefficient of density insignificant in many cases. Both to avoid such multicollinearity problem and to weaken endogeneity problems due to the spatial nature of data, we consider (as in Mion, 2004 and Hanson, 2005), the following measure of Harris market

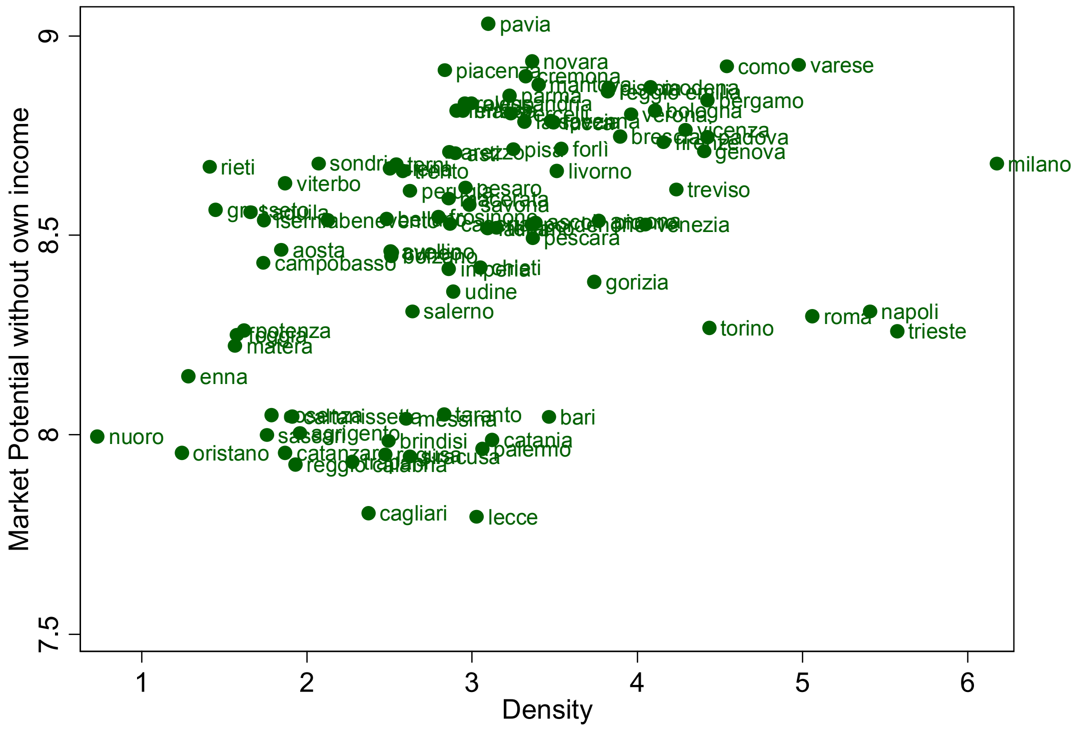

Figure 2: Correlation between density and market potential without own income.

potential:

$$MP_{j,t} = \ln \left[\sum_{k \neq j} Y_k d_{jk}^{-1} \right] \tag{15}$$

where we exclude the income of region j in the computation of $MP_{j,t}$. By dropping own income, the correlation between the two variables falls down to 0.25. Figure 2 shows the new scatter plot. As one can see, densely populated cities like "Milano", "Torino", "Roma" and "Napoli" are now characterized by having a moderate market potential, while small neighbor locations like "Pavia", "Novara", and "Rieti" largely benefit from the proximity to such big markets (so having a high $MP_{j,t}$).

Obviously, although this strategy avoids the multicollinearity problem, it does not provide us with a "pure" measure of market potential and economic density. This comes from the fact that, at very local level, these two forces are virtually impossible to disentangle. Consequently, we will use the following "unstructural" counterpart of equation (10) in our estimations:[14]

$$\ln(w_{j,t}) = \delta_t + \gamma_1 Dens_{j,t} + \gamma_2 MP_{j,t} = \delta_t + \gamma_1 \ln\left[\frac{empl_{j,t}}{size_t}\right] + \gamma_2 \ln\left[\sum_{k\neq j} Y_k d_{jk}^{-1}\right] \quad (16)$$

In the next Section we will present the datasets used, the way we have constructed our regressors and instruments as well as some simple descriptive statistics that gives a flavor of the spatial variability of wages.

4 Data Description and Variables

4.1 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[15], which matches employer and employee data - like the one used by Kramarz, Abowd, and Margolis (1999) - and that has been recently used for instance by Garibaldi and Pacelli (2004). The sample units are salaried full-time workers[16] in all private sectors but agriculture. The panel is an employer-employee database, constructed merging INPS employee information database with the employer information database and covers 14 years from 1985 to 1998. 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.

As far as workers' information is concerned, the database contains many individual information like age, gender, qualification, place and date of birth, province where the job takes place, 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 yearly wage, and the number of worked weeks and days.

For firms this database contains the following information: plant location (province), headquarter location (province), the average number of employees, the sector and the date of start

[14] We call equation (16) "unstructural" because, although it comes almost directly from the structural equation (10), the proxies we use for the variable of interest are different from the model for the coefficients γ_1 and γ_2, and cannot be interpreted as structural parameters.

[15] For a detailed explanation of this database see Centra and Rustichelli (2005).

[16] Apprenticeships and part time workers are excluded from the dataset, since the attention is focused on standard labour market contracts (blue collar, white collar and managers). Further, self-employed are not included in our sample.

up and the one of shut down (if any). This means that, contrary to other datasets, we are able to exactly identify here where a job takes place since the headquarter and plant location are two separate information. This is particularly important for us because job location is a crucial element in our analysis.

As far as job location is concerned, we use data on the 95 Italian provinces.[17] 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. Furthermore, Mion (2004) uses the same spatial disaggregation in its structural analysis of agglomeration externalities. 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). This choice should provide us with sufficient information to properly account for sectoral heterogeneity. Data on households' disposable income at province level are instead provided by the "Istituto Tagliacarne" and cover the period 1991-2000.

As for historical data, employment density and specialization for the year 1951 come from the "Ateco51-91" database provided by ISTAT. Disposable Income and GDP at province level in 1961 have been reconstructed from regional accounts data provided by ISTAT by assigning to each province a share of regional income corresponding to its share of population.[18] The same methodology have been used to reconstruct province unemployment rates in 1961 with regional data again provided by ISTAT. Finally, data on surface crow-fly distances among provinces comes from Arcview GIS software.

4.2 Variable 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 at year t. 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 from 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

[17] Actually, in 2005 the Italian provinces are 103. The transition between 95 and 103 took place gradually during the nineties. In this paper we consider the initial classification of 95 provinces, converting subsequent changes in definitions on our data.

[18] There are 20 regions in Italy and a unique correspondence with each of the 95 provinces.

vs reported working days vs weeks which can systematically vary across space.[19] We further eliminate those extreme observations below (above) the 1^{st} (99^{th}) percentile of the yearly wage distribution and consider only workers with at least two observations in order to be able to perform a within transformation on all our data. This lead us to an unbalanced panel of 92,579 individuals corresponding to 560,040 observations over the period 1991-1998. Table 1a contains some descriptive statistics of the variables used in our analysis that refers to such a sample.

However, in the regressions we extensively use a smaller dataset that refers to male prime age workers only. In particular only male individuals with age between 24 and 39 (when they first enter in the database) are considered, *i.e.* 24,353 workers and 175,700 observations.[20] The choice to consider only male workers is quite standard in the wage equation setting.[21] Women wage dynamics is in fact often affected by non-economic factors, meaning that standard economic and spatial covariates are less important in explaining their carriers. Further, workers aged 24-39 are the ones most interested in economic and spatial incentives both for carrier and personal (familiar) reasons. Consequently, their location choice should better reflect the kind of forces we want to investigate.[22] Descriptive Statistics on this subsample are given in Table 1b.

The dependent variable in our regression is the (log) of before tax 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: gender dummy (one for female), age, age^2, log of firm size, and two other dummies for blue collar and white collar with the residual category being managers, and time and sectoral dummies.

Moreover, the relation between wages, productivity and firm size has also been extensively studied in labor economics. The seminal papers are probably the ones of Krueger and Summers (1988) and Brown and Medoff (1989), and more recently the well-known paper of Burdett and Mortensen (1998) that derives this result using a general equilibrium matching model with frictions. In particular, this literature points out a persistent positive effect of firm size on wages,

[19] 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 (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.

[20] Moreover, in order to carry out very general dynamic panel data estimations, we further select only workers for which we have at least 5 observations. We also use datasets in which we consider only individual with at least 3 observations: results do not change.

[21] See for instance Topel (1991) and Altonji and Shakotko (1987).

[22] It is certainly true that male prime-age workers are more mobile. In the full sample of 92,579 workers 10.49% of them change location at least one time in the observation period against the 13.54% of male prime-age. For instance, also Dustmann and Meghir (2005) use a similar sample, considering male workers aged under 35.

identifying several different explanations that have been put forward in order to explain such a relation. 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 (see Podgursky, 1986). Further, another explanation concerns the fact that big size firms make use of a better screening device in order to select high skilled workers.[23]

According to these insights, it is of particular interest for the present analysis the literature on heterogeneous firms/workers and the costs of trade, started with the work of Melitz (2004). In Bernard, Eaton, Jensen, and Kortum (2003), Melitz (2004), and Melitz and Ottaviano (2004) big size firms are actually more productive and locate disproportionately in thick markets. To the extent that higher firm productivity is transferred to higher wages, this literature suggests a positive relation between firm size and wages. Yeaple (2005) further identifies a link between firm size and individual skills and we will come back to these issues later on.

As spatial variables we consider both employment density and market potential as defined (respectively) in (14) and (15). Both variables are in log (to be interpreted as elasticities), and stand for urbanization economies as apposed to pecuniary externalities stemming from transportation costs, product differentiation and increasing returns to scale. Although we focus on the comparison among these two forces, in the urban literature there is also a substantial interest for the so-called Marshall-Arrow-Romer (MAR) externalities, which concern the productivity growth in a specific sector due to the concentration (or specialization) of this sector in the related location. Both the theoretical models of Henderson (1974) and Duranton and Puga (2004) and the empirical findings of Glaeser et al (1992), Henderson (2003) and Rosenthal and Strange (2003) suggest that these externalities play an important role in local growth and productivity. Consequently, we decide to include them in the analysis even though, compared to density and market potential, we are less able to tackle the related endogeneity issues. Nevertheless, as we will see later on, neglecting them does not alter substantially the results. As a proxy for such externalities, we use both a battery of industry dummies (to control for sectoral heterogeneity), as well as a measure of local sectoral specialization as in Combes (2000):

$$Spec_{j,s,t} = \ln \left[\frac{empl_{j,s,t}/empl_{j,t}}{empl_{s,t}/empl_{t}} \right], \tag{17}$$

[23] For a survey concerning all this literature see Oi and Idson (1999).

where the specialization index for sector s in province j is defined as the ratio of the employment share of sector s in province j divided by the same ratio at the national level.

4.3 Some basic stylized facts

From descriptive statistics some basic elements concerning the goals of our paper come out quite clearly. First of all, spatial distribution of wage does not seem uniform across Italian provinces, meaning that location matters. This does not have to be taken for granted, since Italy is a country characterized by a very important centralize wage setting, where each sectoral national contract has to respect several constraints, like a minimum wage. For this reason we were expecting a more uniform distribution of wages. However, it is worth noting that firms are allowed to integrate the national sectoral contract with a company specific contract, in which for example the minimum wage can be increased. Besides, since several standard economic theories suggest that fixing wages above the minimum wage level might represent an efficient solution for the firm (for instance the efficiency wage approach, the insider outsider and/or the wage setting in presence of unions etc.), it is not surprising that wage distribution is affected by economic location.

First of all, it is interesting to remark that the ratio between the province with highest average wages and the lowest one is (considering time averages over the period) 1.52 and this ratio is increasing overtime (1.46 in 1991 and 1.56 in 1998). This result still holds even 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 while collar and 2.82 for managers. Even considering a less extreme indicator than the max/min as the $90^{th}/10^{th}$ percentile ratio we still derive interesting results. This ratio for all workers is 1.24, 1.22 for blue collar, 1.17 for white collar and 1.33 for managers, confirming that there is a substantial wage spatial distribution.

When considering the relation between province average wages (across individuals) and density (mean 1991-1998), it is possible to derive a clear positive correlation leading to an R^2 of 0.36. Market potential is also a powerful explanatory variable, with the R^2 of the regression on province average wages being 0.28. However, coherently with Combes, Duranton and Gobillon, we find that individual skills are by far the most important factors behind spatial wage disparities. In fact, the province average of our estimates of individual fixed effects,[24] explains alone 96% of the differences among aggregated province wages. The spatial literature has paid little attention to the issue

[24] Estimated fixed effects are those referring to IV estimation of column (1) Table 7.

of skills and much more on spatial externalities. In this paper we try to mind this gap thanks to individual data and in the following Section we will show, using more rigorous econometric techniques, the implication of the unequal skills' distribution on the estimated magnitude of spatial externalities.

5 Econometric Analysis

5.1 The Importance of Individual Panel Data

The goal of our empirical analysis is the estimation of (10) and in particular of its "unstructural" counterpart in equation (16). We use a panel data on individual wages and this certainly represents one of the main contributions of this paper. Most of the studies that has dealt with the measurement of location-specific externalities, like Glaeser et al (1992), Ciccone and Hall (1996), and Mion (2004), use in fact aggregate data on labor and/or productivity. Compared to them, we can thus control here for a possible composition effect due to individual characteristics like age, gender and qualification. 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. In particular, female workers can be found more easily in big cities with the working population being a little older. Since female workers earn less while wage is positively correlated with age, the sign of the bias coming from omitting these variables on economic density is (in this case) undetermined a priori.

Another important variable in our analysis is the employer' firm size. On the one side, as already underlined there is in fact 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, then omitting this variable would have no impact on the estimates of spatial externalities. To this respect, the recent literature on heterogeneous firms/workers and the costs of trade (started with the work of Melitz, 2004) has something to say about that. In particular, Melitz and Ottaviano (2004) build a model in which the size of the local market is positively related to firm size and productivity. The bottom line of their argument is that local competition lead to a self-selection of firms in which less productive ones exit the market. 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 is important in our spatial analysis.

Least but certainly not last the issue of individual skills. First of all, higher skilled workers earn higher wages. This is certainly a clear-cut statement and there is an old literature in labor economics focusing on the returns of education.[25] However, the focus of this paper is not to measure "the value of education", but to control for those individual characteristics that are unobservable and related to skills. As argued by Glaeser and Mare (2001) those skills are, to some extent, a stock variable that is subject to accumulation. However, in our short time period (1991-1998) one can reasonably consider them as fixed. All this being said, the use of individual fixed effects to proxy for individual skills seems to be a reasonable choice.[26]

As shown in the previous Section, individual effects explain a great deal of the (aggregate) spatial variability of wages. However, the goal of our empirical exercise is the estimation of the relative and absolute magnitude of spatial externalities. In this light, the need for individual effects is thus (again) conditional on the presence of a significant correlation with our spatial regressors that (if neglected) would bias our results. Glaeser and Mare (2001), Moretti (2004), and Redding and Schott (2004) provide the rationale for an interplay between skills and space. In particular, these frameworks suggest that skilled workers should locate in location characterized by high density and market potential. Yeaple (2005) further identifies a link between individual skills and firm size. In particular, the author shows that firms who choose to be high tech are more productive, bigger and have a skill-biased technology. Consequently, bigger firm should be observed to hire more skilled workers and, even after controlling for skills, to have a residual productivity advantage. The link between firms' size and skills have already be found empirically by Abowd, Kramarz, and Margolis (1999), but we are able here to go a bit further by assessing whether the observed correlation is just due to co-location of both agents in big cities or not. Furthermore, the interaction of skills with firm size can create (if not accounted) a feedback bias on the spatial variables whose sign and magnitude cannot be predicted a priori. These considerations thus suggest that individual skills certainly needs to be included in the analysis.

Once established the importance of using panel data on individuals in our spatial analysis we can proceed to the estimation of equation (16). Characterizing this equation by adding individual and time labels (i, t), as well as a set of controls for individuals and sectoral characteristics (both variables refers to vectors):

[25] See for instance Mincer (1974) and Card (1999).

[26] This type of choice is usually made in labour economics, see for instance Krueger and Rouse (1998) and Moretti (2004).

$$w_{i,t} = \mathbf{B}_1^{'}\mathbf{I\_Char}_{i,t} + \mathbf{B}_2^{'}\mathbf{Sect\_Char}_{j(i,t),s(i,t),t} + \gamma_1 Dens_{j(i,t),t} + \gamma_2 MP_{j(i,t),t} + \delta_t + u_i + \varepsilon_{i,t} \quad (18)$$

where the dependent variable is the logarithm of before tax monthly wage, u_i is an individual effect (skills) , δ_t is a time effect, the term $\mathbf{I\_Char}_{i,t} = \{Gender_i,\ Age_{i,t},\ Age_{i,t}^2,\ \ln(FirmSize_{i,t}),$ $Bc\ dummy_{i,t},\ Wc\ dummy_{i,t}^{'}\}$ is a battery of individual characteristics, $\mathbf{Sect\_Char}_{j,s,t} = \{\mathbf{i}_{s(i,t)},$ $Spec_{j(i,t),s(i,t),t}^{'}\}$ contains a full set of industry dummies ($\mathbf{i}_{s(i,t)}$) and our measure of specialization $Spec_{j(i,t),s(i,t),t}$ as defined in (17), and finally $Dens_{j(i,t),t}$ and $MP_{j(i,t),t}$ indicate density and market potential as defined in (14) and (15). It is worth noting that in our notation both the sectoral index s (referring to the 52 Ateco81 sectors) and the location index j (referring to the 95 provinces) depend upon the couple (i, t) because they vary when an individual change sector and/or province at time t. As already mentioned, these "migrations" represent most of the effective within time variation in the data and are crucial to identify our spatial externalities.

5.2 First Results

In order to give a comprehensive overview of the relation between wages, skills and spatial externalities we present in this Subsection estimations of (18) based on OLS, GLS and Within estimators. More sophisticated techniques that deal with the endogeneity issue are implemented in the next Subsection.

Table 2 shows the results obtained using the sample of male prime age workers for all sectors. Columns (1) to (3) contains (respectively) OLS, GLS, and within estimates of (18) in which firm size (which is a relevant variable for us) has been excluded in order to better evaluate its contribution in the within estimations in column (4). In all specifications a complete set of time and sectoral dummies have been added.

First of all, our estimates on the impact of $Age_{i,t}$ and $Age_{i,t}^2$, and the two dummies for blue and white collar are in line with previous findings (Naticchioni and Panigo, 2004). Moreover, the impact of MAR externalities, as proxied by our specialization measure, is always very low (between 0.55% and 0.15%) and weakly significant in within estimations of column (3). This is also consistent with other works and in particular with those of Cingano (2003) and de Blasio and Di Addario (2002) who did not find any strong evidence in favour of a positive wage differential in highly specialized areas (Industrial Districts). These variables are not of direct interest in our analysis and thus we will not discuss them further.

The Spatial Sorting of Skills and Firms

As for density and market potential, going from column (1) to (3) it is quite straightforward to see that taking into account individual effects dampens simple OLS results. Nevertheless, these variables are always strongly 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 in Italy there is a sectoral bottom floor that is fixed at national level and this can reasonably limit spatial variation of wages. However, taking into account individual effects further reduce this estimate. When considering (uncorrelatd) random effects the effect of density drops to 1.87%, while allowing this effects to be correlated with regressors in the within estimations of column (3) leads to only 0.74%.[27] This simple estimates suggest a strong positive correlation between individual skills, as measured by u_i, and density. Indeed, our within estimations give a (significant) correlation of 0.20, which suggest that sorting of skills in space is at work.

These findings confirm those of Combes, Duranton and Gobillon (2004) and lead to interesting theoretical considerations. The fact that skilled workers are disproportionately distributed in big cities can be explained in at least three ways. The first possibility 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. Moretti (2004) actually finds some evidence of a positive relation between density and returns to education (or what he calls "social returns to education"). As for the second explanation, Glaeser and Mare (2001) discuss a different 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 favour of their conjecture which does not necessarily imply that those who initially moves to cities are particularly skilled. Third, there 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 valuable assets for them, they

[27] Although this elasticity might seem really low, Di Addario and Patacchini (2005) 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.

would prefer cities as the destination of their migrations. We will come back to this issue in the next Subsection.

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 (2005) find a very similar result. However, taking into account individual skills push down elasticity to 5% in the within estimates. The (significant) correlation between the u_i and market potential is 0.08 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). The authors develop a NEG model in which they show that, 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 reduces incentives for human capital accumulation. Empirically, they further exploit the structural relationships of their model to demonstrate that countries with lower market potential have lower levels of educational attainment.

In column (4), we further add to our within estimation the variable firm size. The elasticity with respect to wages is 1.94%, 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 impact on spatial variables, considering firm size lowers 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. The fact that big firms should be found disproportionately in large markets is consistent with the model of Melitz and Ottaviano (2004). According to the authors bigger firms are in fact more productive and more capable to survive to the strong competition faced in thick markets. Consequently, as further showed by Baldwin and Okubo (2004), part of the measured added value created by agglomeration forces is in fact due to a spatial sorting of firms that closely resemble that of workers. In its theoretical framework, Yeaple (2005) combines these two elements drawing some conclusions on which we will come back in the next Subsection where we will present our preferred IV estimations.

Sample and Sectoral Robustness of the Spatial Sorting

In Table 2 we consider only male aged 24-39 (when they first enter in the database) because we believe that their location choice should better reflect the kind of forces we want to investigate. However, one can reasonably wonder to what extent our results still hold in the full sample of workers. We explore this issue in Table 3. The first conclusion one can draw is that sorting of firms and workers still applies. Within estimates are in fact systematically lower than OLS and GLS and the correlation of u_i and firm size with density and market potential are still significant and of the same sign of the sample of prime age workers. However, both elasticities are lower than those of Table 2. For instance, the potential market potential coefficient is almost twice higher (4.53%) than the one related to all workers (2.60%) in the within estimate, and a similar remark applies to density. This suggest that non male prime age workers, and in particular women, are less sensitive to those spatial wage imbalances that lead to migrations. As already mentioned, the change of locations is in fact the main source of spatial data variation in our analysis, and male prime-age workers for sure are more mobile. In the full sample of 92,579 workers 10.49% of them change location at least one time in the observation period against the 13.54% of male prime age. However, in the full sample estimates are still positive and significant and suggest that agglomeration externalities play an important role in the wage profile of all workers.

Another issue we deal with is the sectoral scope of our analysis. One can in fact reasonably argue that using dummies to account for sectoral heterogeneity may be too restrictive and that all results may be driven by manufacturing activities only. In Table 4 and 5, which are the counterpart of Tables 2 and 3, we show estimations of (18) obtained on the subsample of manufacturing workers. These estimates are based on a set of economic activities that are more directly comparable and should thus provide a more robust measure of agglomeration externalities. As one can see, however, all elasticities are still positive and significant with magnitudes comparable to those referring to all activities with two interesting exceptions. On the one hand, MAR externalities seems to be stronger and more significant for manufacturing and this is somehow expected since the idea that specialization fosters growth and productivity is historically related to such activities. However, the difference between comparable estimates is in most cases not significant and caution is needed. The other interesting difference is with respect to market potential. In the subsample of manufacturing, market potential seems to matter less than for other activities. Punctual estimates are in fact (in within estimations) almost half of their counterpart in the sample of all activities. This may suggest that other sectors, and in particular services -that usually display higher transportation costs- are more sensitive to market centrality. However, the difference between comparable estimates is again not significant and caution is needed.

5.3 Endogeneity

In this Subsection we explore the issue of endogeneity and we restrict our analysis to the sample of male prime age workers[28]. Although within estimations can give 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, in the within estimation 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 technology 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. Combes, Duranton, and Gobillon (2004) show that the bias is much reduced in a dynamic context when workers make their employment decision on the basis of both current and future (expected) wages. Nevertheless, the issue of endogeneity of density, market potential and, to some extent, also of the sector choice with the related specialization variable remains open.

We deal with endogeneity in two ways. The first one, inspired by the structural analysis of market potential externalities of Mion (2004), is based on dynamic GMM panel regressions à la Arellano and Bond (1991) and Arellano and Bover (1995). The basic idea is to exploit the time variability of the data and the assumption of weak exogeneity in order to overcome the endogeneity problem. The outcome of such estimations is presented in Table 6. An alternative and much more satisfying solution based on IV is instead presented in Table 7. The problem with GMM regressions is in fact that "pure" time variability is, contrary to Mion (2004), rather low in our sample, while variability induced by migrations is much more important. However, as long as migrations are concerned, a time span of 8 years is probably too short for lagged spatial variables to be valid (exogenous) instruments. Our specification test confirm this idea, since the Sargan test always rejects GMM moments' restrictions. In our preferred IV estimations we instead exploit the idea of Ciccone and Hall (1996) and Combes, Duranton and Gobillon (2004) of using deeply lagged values of the endogenous variables, as well as other "past" variables, as instruments. Crucially, in IV estimations the test on over-identifying restriction now accepts the validity of such instruments while parameters are in line with economic meaningful values.

[28]Unreported estimations on the full sample of workers actually yeld qualitatively similar results.

Dynamic Panel GMM Estimations

In Table 6 we show the results of several GMM dynamic panel estimations of (18). We start in column (1) with a simple estimation of the model in difference with all leads and lags of the levels of specialization, density and market potential used as instruments[29]. This estimation exploits almost all the moments conditions implicit in the within estimation (from t-4 to t+4) and in fact we derive very similar results to those of Table 2 column (4). Nevertheless, both the Hansen test and the test on differenced residuals autocorrelations suggest that this specification is not valid, meaning that within estimations cannot be fully trusted.[30] In columns (2) we consider a more general model in which only weak exogeneity of specialization, density and market potential is assumed, that is Cov($\varepsilon_{i,s}, \mathbf{X}^*_{i,t}$)=0 only for $\forall\ s >\ t$ and $\mathbf{X}^*_{i,t} = \left\{Spec_{j(i,t),s(i,t),t},\ Dens_{j(i,t),t},\ MP_{j(i,t),t}\right\}'$, while in column (3) we further introduce an autoregressive term ($w_{i,t-1}$) as an additional (predetermined) regressor.[31] As a consequence, all estimates of our spatial variables in columns (2) and (3) takes implausible negative and significant values with a doubling of market potential being for example associated to a 65% decrease of wages. As already said, it is likely that the weight of migrations in the time variability of the data makes GMM inappropriate to tackle with endogeneity. In fact migration is the outcome of a long lasting and sticky choice in which a long time horizon approach is needed. Consequently, using a three-four years lagged values of density in a province is probably a too close information to be exogenous to current re-location decisions. Indeed, specification tests for estimates of columns (2) and (3) suggest that this is actually the case. Moreover, even when experimenting with system GMM in column (4) and with the introduction of lags of the spatial variables in column (5) the quality of tests and estimates do not improve. We thus conclude that GMM dynamic panel data are not a useful tool in our short panel.

IV Estimations

In Table 7 we show our IV results, which we believe are the most reliable estimates of spatial externalities we can provide. In particular, column (1) represents our preferred specification. In

[29] More precisely we always report, as suggested by Blundell and Bond (1998), one-step GMM robust estimates and (robust) two-step test estimates.

[30] In particular, if time dynamics is properly captured, differenced residuals are expected to be negatively correlated at one lag and not correlated at two lags.

[31] In order to get more robust results, we further assume that $\varepsilon_{i,s}$ follows a MA(1) process, and so we consider lagged levels of endogenous regressors starting from t-3 only. Unreported estimations in which we consider higher order MA processes or independent errors yield qualitatively identical results.

this IV within estimation we actually use data on specialization and density in 1951, market potential calculated with disposable income in 1961, as well as unemployment rates in 1961, as instruments for our spatial variables.[32] The idea of using unemployment rates comes from the consideration that, as long as one wants to capture the determinants of workers' migrations, local availability of jobs should certainly be considered. Italy has in fact experienced, especially in the past, massive migrations of job-seekers from the South to the North. By contrast, 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 pattern 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, and world war two, and the consequent return to democracy in Italy, would be a good candidate.

Unfortunately, data before world war two may be obtained only for density. Nevertheless, 1951 and 1961 are quite close to world war two, and the growth take-off that has changed the Italian economy has essentially started only in the 60'. Moreover, as a matter of fact, the oldest workers in our sample were one year old in 1961. These considerations suggest that, although precaution is needed, our instruments should be sufficiently reliable. Crucially, the Sargan test on the one over-identifying restriction does 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 very high. Furthermore, the effect of accounting for endogeneity is, for comparable variables, close to the results of Combes, Duranton and Gobillon (2004) that use pre-world war two population data. In fact, our IV estimates suggest that the elasticity of density should be reduced by almost 50% like in Combes, Duranton and Gobillon (2004). Compared to within estimation in column (4) of Table 1, density goes from 0.56% to 0.26%. By contrast market potential, which is not properly measured in Combes, Duranton and Gobillon (2004), is just slightly affected going from 4.53% to 4.25%. Nevertheless, these differences are not significant at 1% and caution is needed. Finally, the elasticity with respect to specialization is still not significant.

As for the relative importance of density and market potential all our estimations, except GMM, suggest that the latter is more important in explaining spatial wage disparities. The elasticity corresponding to market potential is in fact always higher (with a gap that is statistically

[32] Data on disposable income, GDP, and unemployment rates at province level for 1961 have been reconstructed from available regional data by assigning, when applicable, the corresponding province share on the basis of population data which are available at province level for 1961.

significant), and this apply in particular to our preferred estimation in column (1) of Table 7. When considering "standardized" elasticities, this result still holds with the one corresponding to density (market potential) being 0.0087 (0.0335). 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.[33] Both absolute and beta elasticity thus suggest that, at least for Italy, pecuniary externalities play a crucial role in the spatial distribution of wages. These findings are coherent with those of Mion (2004) who finds evidence of sizeable agglomeration externalities in Italy.[34]

Robustness Checks

In column (2) we perform the same estimation as in column (1) except from the fact the we now exclude both the specialization variable $Spec_{j(i,t),s(i,t),t}$ and the sectoral dummies $\mathbf{i}_{s(i,t)}$. Although we are able to reasonably instrument for the endogeneity linked to location specific variables, we cannot say the same for the sectoral variables. The sector choice is in fact certainly endogenous and, although we can use specialization in 1951, we do not really have something to instrument sectoral dummies. Therefore, we have performed in column (2) a restricted estimation that intentionally avoids the choice of sector. The results we got for density and market potential are nevertheless almost identical. Furthermore, the Sargan test on over-identifying restrictions does not detect any bias from omitted variables. We interpret these results as an evidence that location and sector choice are quite independent from each other, implying that our estimates of agglomeration externalities are robust to the misspecification of the sector choice.

In column (3) we replicate, for comparability, the same estimation methodology used by Combes, Duranton and Gobillon (2004). The authors assume in their estimation equation, which

[33] See Wooldridge (2003, Section 6.1) for a further description of this transformation.

[34] It is worth noting that in our database migration decisions are always generated by job change decisions: a worker decides to change job in order to achieve an improvement in working conditions, and in some cases this implies a migration. In this framework, it would be possible to argue that the reasons underlying job-to-job decisions could entail an impact on the identification of the spatial coefficients. In order to avoid such a critic we have generated three dummy variables, to captures the following effects: 1) general decisions to change a job (no matter if voluntary or involuntary); 2) a voluntary decision to change a job (defined as absence of unemployment spells between two jobs, as in Abowd, Kramarz and Margolis (1999); 3) a decision to move in another provinces to work. We have added these dummies to our standard within estimations, both one dummy at a time and jointly. Coefficients are significant and coherent with our expectations: a general job change entails negative impact of wages (because involuntary job changes are included), while voluntary movements are linked to gains. Further, dummy migration is also positive and significant, even if very small in magnitude. More importantly, previous estimates for our variables of interest (density, market potential and firm size) do not change, using both within estimates and IV, meaning that their identification is weakly concerned by job change decisions.

is very similar to (18), that there is a further time-location specific error component ($v_{j(i,t),t}$) that can be thought as an idiosyncratic technological shock. In order to control for this additional source of heterogeneity, the authors perform a first-step within estimation in which they include a full set of time-location dummies ($\beta_{j(i,t),t}$) that capture all the variation in the time-space dimensions. Subsequently, they recover the parameters of density and their proxy for market potential in a second step regression $\beta_{j(i,t),t} = \gamma_1 Dens_{j(i,t),t} + \gamma_2 MP_{j(i,t),t} + v_{j(i,t),t}$ using a two-steps least squares estimator and deeply lagged values as instruments. Compared to our strategy, their methodology has the advantage of accounting for the heteroschedasticity that comes from time-location technological shocks $v_{j(i,t),t}$. In other words, although our IV estimates would still be unbiased, the standard errors may not. However, when they first recover their dummies without instrumenting, the endogeneity of spatial variables is still at work and can seriously bias their estimates of $\beta_{j(i,t),t}$. In order to get some insights on the relative advantages of the two procedures we have fully implemented their estimation techniques with the exception that we have just considered location dummies in the first stage (i.e. $\beta_{j(i,t),t} = \beta_{j(i,t)}$). We do not have in fact enough time span and migrations to be able to identify all possible time-locations effects. However, comparing results of column (1) and (3) reveals that the difference between the two sets of estimates is actually very small. Moreover, the increase in standard errors that is expected as a consequence of heteroschedasticity is very small suggesting that location heterogeneity is very small compared to that of individuals. However, the Sargan test does not accept the validity of instruments in the second step. This may be due to the fact that first step estimates in the procedure of Combes, Duranton, and Gobillon (2004) suffer from endogeneity. However, this may also come from the restriction of the time-invariance of $\beta_{j(i,t),t}$ that we were forced to impose.

In columns (4) and (5) we perform additional estimations of (18) with our IV method. In particular, in column (4) we partitioned Italian provinces in four subsamples according to four macro areas (North-East, North West, Center, South).[35] For each subsample, we have thus performed separate estimations. Coefficients and standard errors reported are actually the average of the corresponding values of the four regressions. The reason of this exercise is twofold. On the one hand, we want to check whether spatial effects are possibly due to different returns on age, profession, or firm size across space. On the other hand we want to be sure that the North

[35] The four macro areas are made by the following regions (according to the official classification): 1) Northwest: Valle d'Aosta, Piemonte, Liguria, Lombardia; 2) North-east: Veneto, Trentino, Friuli Venezia Giulia, Emilia Romagna; 3) Center; Toscana, Marche, Umbria, Lazio; 4) South; Abbruzzo, Molise, Campania, Puglia, Basilicata, Calabria, Sicilia and Sardegna.

is not driving our results and particularly the one on market potential. As for the first issue, our estimations confirms that the positive impact of density and market potential on individual wages is robust even after introducing heterogeneity in individual characteristics' returns. In particular, both elasticities increases although the difference with pooled estimations is not significant. Moreover, looking at macro-area specific estimates, while the coefficient of density is very stable across space around its mean of 0.62% the same is not true for market potential that ranges from 2.27% for the North-East to the 10.45% of the South. This relatively unstable effect of market potential may be due to the fact that the Southern Italian economy heavily relies on the richer North that is (by contrast) much more export oriented.[36] In fact, we did not consider international demand in the construction of market potential and so the proximity of the North to the economic core of Europe is actually neglected. However, in all four cases estimates are positive and significant suggesting that market effect is a pervasive spatial force. In column (5) we instead split space according the (time) average quartiles of the distribution of density. Once again, estimates suggest that heterogeneity of returns on age, profession, and firm size do not kill spatial effects. However, in this case coefficients vary considerably depending on the size of provinces suggesting the existence of strong non-linear relationship between spatial externalities and economic density/market potential. We will further explore this issue in the next Subsection.

The Interaction between Skills, Firm Size and Space

As for spatial sorting of workers' skills, our IV estimation of Table 7 column (1) confirms the strong link with spatial variables. Individual fixed effects are in fact positively and significantly correlated with both density (0.21) and market potential (0.08), suggesting that the arguments we developed when discussing simple within estimations are robust to controlling for endogeneity. However, with more reliable measures of individual skills in our hands, we can now go a step further trying to investigate the reasons why such skills are spatially sorted. As already said, there are at least three competing explanations concerning the link with the density of economic activities. However, a peculiar feature of the "bright lights" hypothesis, compared to the other two, is that it needs workers migrating to cities to have above average skills.

In order to provide some insights on this issue we have analyzed the skills (as measured by u_i) of workers moving from big (small) to small (big) provinces. Big (small) provinces are those with a (time average) density above (below) the median. The average skills of workers in big (small)

[36] Interestingly, Mion (2004) also finds that market potential has a bigger impact for the South of Italy.

cities workers is 0.0501 (-0.0467) with a standard error of 0.0022 (0.0024). Big cities have thus a significant higher average skill level of workers. However, looking at the 1,154 (1,153) migrations from big (small) to small (big) provinces of our sample we have that the average level of migrants' skills is -0.0109 (-0.0039) with a standard error of 0.0079 (0.0079), which are both not significantly different from the overall mean of zero. This means that the skills of people moving from and to large cities are not particularly high and certainly far below the skills of people already living in those cities. To sum up, human capital accumulation in cities and heterogeneous returns to education seems to be a more robust explanation for spatial sorting across Italian provinces.

Another interesting issue, linked to the theoretical framework of Yeaple (2005), is the relation between workers skills and productivity. In particular, the author shows that firms who choose to be high-tech are endogenously bigger and have a skill-biased technology. Consequently, big size firms are expected to hire more skilled workers and, even after controlling for skills, to have a residual productivity premium. This in indeed the case in our regressions where the firm-size effect is always significant. Furthermore, the correlation between the individual effects of our IV estimations and firm size is very strong (0.35) and significant. A similar result have been found in Abowd, Kramarz, and Margolis (1999). However, the fact that we specifically deal with space in our framework allow us to go a bit further. Actually, the link between skills and firm size may be in principle due to a simple co-location effect. Melitz and Ottaviano (2004) suggest that big and more productive plants should locate in big 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. However, the partial correlation between individual effects and plant size *conditional* on spatial characteristics (density and market potential) is 0.33 which is just slightly smaller than the unconditional one and still highly significant. This suggest that there is a deeper underlying economic relation between skills and firms' characteristics like the one traced in Yeaple (2005).

6 Additional results

In this paragraph we want to further investigate some issues of the previous sections.

First of all, we want to evaluate whether the impact of spatial variable are actually linear or there are some non-linearity at work. Since in the previous section we have showed that endogeneity issues do not seem to be a relevant issue in our analysis, we do not use neither IV nor GMM estimators. Further, in order to take into account the correlation between individual effects and covariates we use within panel estimates.

Of course we focus our attention on the variables of interest of this paper, *i.e.* the relevance of both density and potential demand variables in the determination of wages. So far a linear specification has been assumed. However, the impact of agglomeration and potential demand might be of a different shape, entailing non-linearity. For this reason we have derived dummies variable, one for each decile of the frequency distribution of density and potential demand. From table 9 it is possible to note that non-linearity seems to play a role.

As far as density is concerned the impact on wages seem to be very low for the first three deciles (the first decile is the omitted one), while it is stronger from the fourth deciles on. A threshold level seems to be at work: under the fourth decile the agglomeration process is not enough developed to guarantee a wage premium, after this level there is a jump that is quite constant afterwards. The fact that after this threshold the returns from agglomeration do not increase in a relevant way probably means that there might be some other forces, such as congestion costs or decreasing returns, which set a sort of upward constraints to the positive impact of agglomeration on wages.

The non linear impact of potential demand on wage is more clearly positive and increasing, displaying also in this case a discontinuity threshold at the fourth deciles. Two remarks: this discontinuity is less relevant than in the density case; after this level the impact of potential demand continue to be increasing. Note also that moving from linear to decile dummy specification the differences in magnitude between potential demand and density coefficients tend to decrease, especially due to the increase in the magnitude of density.

	Density		Potential Demand	
	Coeff	p-val	Coeff	p-val
2	0.006	0.013	0.005	0.136
3	0.008	0.017	0.005	0.110
4	0.023	0.000	0.018	0.000
5	0.028	0.000	0.017	0.000
6	0.024	0.000	0.015	0.001
7	0.028	0.000	0.017	0.000
8	0.025	0.000	0.019	0.000
9	0.026	0.000	0.024	0.000
10	0.029	0.000	0.024	0.000

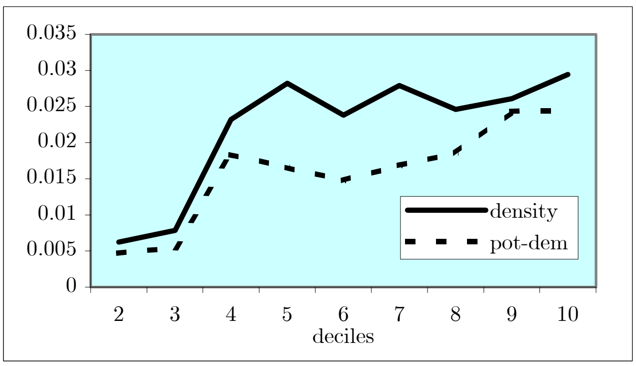

Table 8: Non linear effect of density and market potential, in deciles, on wages

*within estimates. Control variables: age, age^2, firmsize, sector, qualification.

The last remark concerns the fact that so far the industrial specialization at the province level has turned out to be not significant with respect to wages. At first sight it represents an unexpected result, especially if we think that the Italian case is well known for the industrial districts, which can be defined as an homogenous local productive system, characterized by a high concentration of specific manufacturing enterprises, mostly small and medium, and by a high productive specialization. Why is our analysis unable to capture this phenomenon? And Should it be possible that the industrial district system do not imply an effect on wages? Let us briefly investigate this issue. First of all, note that our analysis is carried out at the national level. This basically means that we are putting together provinces in which industrial district are actually at work and other provinces in which high degree of specialization are not necessary related to virtuous industry and dynamic sectors. For instance, in the south there are some provinces specialized in specific sectors -such agriculture or other traditional activities- not necessarily entailing a significant effects on wages. In order to test this interpretation we have selected only the supposed provinces known for the relevance of the Italian districts. More precisely, we consider three regions, Veneto, Emilia Romagna and Marche, for a total of 19 provinces. Using this sub-sample of the INPS database we are able to derive the expected positive impact of specialization on wages. More specifically, using dummies for specialization level we point out that only for high degree of specialization there is a positive impact on wages (table 9), even if not very sizeable. For low level of specialization coefficients become non significant.

quintiles	specification coefficient	p-value
2	0.002	0.355
3	0.006	0.095
4	0.014	0.001
5	0.024	0.000

Table 9: Impact of industrial specialization on wages in the Italian district areas

*within estimates. Control variables: age, age^2, firmsize, sector, density, potential-demand, qualification.

7 Conclusions

In this paper we use wages from a panel data on Italian workers that matches individual and firm-level information (employer-employee) in order to estimate the absolute and relative magnitude of urbanization externalities and market potential. Our results suggest that, after controlling for individual and firm characteristics, skills, and endogeneity, both externalities are positive and significant. However, both absolute and strandardized elasticities suggest that market potential has the stronger impact. These findings are coherent with those of Mion (2004) who finds evidence of sizeable agglomeration externalities in Italy.

The analysis of skills and firms heterogeneity is one of the major contributions of this paper. In particular we provide evidence that spatial sorting is at work in the sense that "good" workers and firms are disproportionately located in provinces characterized by high density and/or good access to consumers' markets. In the paper, we provide several theoretical arguments that can justify the sorting. In particular, Melitz and Ottaviano (2004) suggest that big and more productive plants should locate in big 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. Our data confirm these arguments and further suggest that, as suggested by Baldwin and Okubo (2004), sorting dampens previous estimates of spatial externalities.

Another issue we deal with is the relation between the firm-size premium, individual skills and location. Our results suggests, coherently with Abowd, Kramarz, and Margolis (1999), that the correlation between the size of the employing firm and skills is very strong. However, this correlation is not simply the outcome of a co-location phenomenon, suggesting that a deeper underlying economic relation is at work, like the one traced in Yeaple (2005).

References

[1] H. M. Abdel Rahman, and M. Fujita (1990). "Product variety, Marshallian externalities, and city size", Journal of Regional Science 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] Altonji, J., R. Shakotko, 1987. Do wage rise with job seniority? Review of Economic Studies 54, 437-459.

[4] Amiti M., and Cameron L. (2004). “Economic Geography and Wages”. CEPR Discussion Paper, 4234.

[5] Anderson J. E., and van Wincoop E. (2003). “Gravity with Gravitas: a solution to the border puzzle”. American Economic Review, Vol. 93, 170-192.

[6] M. Arellano, S. R. Bond, Some tests of Specification for panel data: Monte Carlo evidence and an application to employment equations, Review of Economic Studies 58 (1991) 277-297.

[7] Arellano M, and Bover O. (1995). “Another look at the instrumental-variable estimation of error components models”. Journal of Econometrics Vol. 68, 29-52.

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

[9] Beeson P., and R. Eberts (1989). “Identifying Productivity and Amenity Effects in Interurban Wage Differential”. Review of Economics and Statistics Vol. 71, 443-452.

[10] Bernard A.B., Eaton J., Jensen J.B., and Kortum S. (2003). “Plants and Productivity in International Trade”. The American Economic Review Vol. 93(4), 1268-1290.

[11] Blundell R.W., and Bond S.R. (1998). “Initial conditions and moment restrictions in dynamic panel data models”. Journal of econometrics Vol. 87, 115-143.

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

[13] Burdett, K., and Mortensen, D.T. (1998), "Wage Differential, Employer Size, and Unemployment", International Economic Review, vol.39, n.2, pp. 257-273.

[14] Campbell, J. and H. Hopenhayn (2002). “Market size matters” NBER Working paper n.9113.

[15] Card, D., (1999), "The Casual Effect of Education on Earnings", in Ashenfelter, O., and Card, D. (eds.), Handbook of Labor Economics, vol.3A, chap. 30, Amsterdam: Elsevier Science, North-Holland.

[16] Centra, M. and Rustichelli, E. (2005), "La costruzione del panel longitudinale ISFOL su microdati INPS (1985-1999)", forthcoming monografie ISFOL.

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

[18] Ciccone A. (2002). “Agglomeration effects in Europe”. European Economic Review Vol. 46, 213-227.

[19] Cingano, F. (2003). “Returns to Specific Skills in Industrial Districts”. Labour Economics, 10, 149-164.

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

[21] P. P. Combes, M. Lafourcade, Transport Costs Decline and Regional Inequalities: Evidence from France, CEPR Discussion Papers 2894 (2001).

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

[23] de Blasio G., and Di Addario S.(2002). “Labor Market Pooling: Evidence from Italian Industrial Districts”. Temi di discussione No. 453, Bank of Italy.

[24] Di Addario S., and Patacchini E. (2005). “Is There an Urban Wage Premium in Italy?”. Bank of Italy, mimeo

[25] A. C. Disdiez, K. Head (2004). “Exaggerated Reports on the Death of Distance: Lessons from a Meta-Analysis” Mimeo, University of Paris I.

[26] Duranton, Gilles and Vassilis Monastiriotis (2002). “Mind the gaps: The evolution of regional earnings inequalities in the UK 1982–1997”. Journal of Regional Science 42(2), 219-256.

[27] Duranton G., and Puga D. (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.

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

[29] B. Eichengreen, Labor markets and European monetary unification, in P. R. Masson, M. Taylor (Eds.), Policy issues in the operation of currency unions, Cambridge University Press, Cambridge, 1993, pp. 130–162.

[30] M. Fujita, P. Krugman, When in the Economy Monocentric? von Thunen and Chamberlin Unified, Regional Science and Urban Economics 25 (1995) 505-528.

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

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

[33] Garibaldi, P. and Pacelli, L. (2004), "Do Larger Severance Payments Increase Individual Job Duration?", CEPR Discussion Paper no.4607.

[34] Gibbons, Robert and Lawrence Katz (1992). "Does unmeasured ability explain inter-industry wage differentials?". Review of Economic Studies 59(3), 515–535.

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

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

[37] A. Ginzburg, G. Solinas, M. Scaltriti, R. Zoboli, Un nuovo autunno caldo nel mezzogiorno? Note in margine al dibattito sui differenziali territoriali, Politica Economica, vol.14, no.3, 1998.

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

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

[40] Head K. and Mayer T. (2000). "Non-Europe : The Magnitude and Causes of Market Fragmentation in Europe". Weltwirtschaftliches Archiv Vol. 136(2), 285-314.

[41] Head K., and Mayer T. (2004). "Market Potential and the Location of Japanese Investment in the European Union". *Review of Economics and Statistics*, forthcoming.

[42] K. Head, T. Mayer (2005). "Regional Wage and Employment Responses to Market Potential in the EU". University of Paris I, mimeo.

[43] E. Helpman, The Size of Regions, in D. Pines, E. Sadka, I. Zilcha (Eds.), Topics in Public Economics, Cambridge University Press, Cambridge 1998, pp. 33-54.

[44] Henderson, J. Vernon, The sizes and types of cities. American Economic Review 64(4) (1974) 640-656.

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

[46] Krueger; A.B. and Summers, L.H. (1988), "Efficiency Wages and the Inter-Industry Wage Structure", American Economic Review, Vol. 56, No. 2, pp. 259-293

[47] Krueger, A.B., and Rouse, C. (1998), "The Effects of Workplace Education on Earnings, Turnover and Job Performance". Journal of Labor Economics, vol. 16, no.1, pp.61-94.

[48] P. Krugman, Increasing returns and economic geography, Journal of political economy 99 (1991) 483-499.

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

[50] Melitz M. (2004). "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity". Econometrica Vol. 71(6), 1695-1725.

[51] Melitz M., and Ottaviano G. I. P. (2004). "Market size, trade, and productivity". Mimeo.

[52] Mincer, J. (1974), *Schooling, Experience and Earnings*, New York,: National Bureau of Economic Research.

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

[54] 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.

[55] Naticchioni, P. and Panigo, D. (2004) ""Employment protection, job-tenure and short term mobility wage gains: A New Explanation for the Italian Case", CEPREMAP Working Papers (Couverture Orange) no.02.

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

[57] H. G. Overman, D. Puga, Unemployment clusters across European regions and countries, Economic Policy 34 (2002) 115-147.

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

[59] Redding S., and A. Venables, (2004). "Economic Geography and International Inequality". Journal of International Economics Vol 62(1), 53-82.

[60] S. Rosenthal, W. Strange, Geography, Industrial Organization, and Agglomeration, Review of Economics and Statistics 85 (2003) 377-393.

[61] Nobert Schulz and K. Stahl, Do consumers search for the highest price? Equilibrium and monopolistic optimum in a differentiated products markets, Rand Journal of Economics, 27, 542-562 1996

[62] Sveikauskas L.A. (1975). "The Productivity of Cities". Quarterly Journal of Economics Vol. 89(3), 393-413.

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

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

[65] S. Yeaple (2005). "A Simple Model of Firm Heterogeneity, International Trade, and Wages" Journal of International Economics, forthcoming.

Table 1a. Sample Statistics On All Workers

Variable	Observ.	Mean	Std. Dev.	Min	Max
ln(wage)	560040	6.4087	0.3918	1.9810	8.4789
Sex	560040	0.2965	0.4567	0.0000	1.0000
Age	560040	36.5851	10.4533	14.0000	81.0000
Age^2	560040	1447.7390	814.8257	196.0000	6561.0000
Firmsize	560040	4.3774	2.7605	0.0000	12.2699
Bc Dummy	560040	0.6264	0.4838	0.0000	1.0000
Wc Dummy	560040	0.3601	0.4800	0.0000	1.0000
Specialization	560040	0.0962	0.9955	-8.7156	4.9706
Density	560040	3.9837	1.2118	0.6903	6.2398
Market Potential	560040	8.5627	0.2825	7.7637	9.0872
Specialization in 1951	560040	-0.0568	0.8630	-7.2878	3.6926
Density in 1951	560040	3.6322	1.1554	0.7297	5.8864
M.P with Disposable Income in 1961	560040	5.2751	0.1337	5.0608	5.6345
Unemployment rate in 1961	560040	0.0451	0.0157	0.0225	0.0861
Economic Infrastructures	560040	4.6856	0.3557	2.8154	5.2407

All variables (except Sex, Age, Age^2 , 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 for 1951 is the log of current billion liras.

Table 1b. Sample Statistics On Male Prime-Age Workers

Variable	Observ.	Mean	Std. Dev.	Min	Max
ln(wage)	175700	6.4824	0.3625	3.1180	8.2036
Age	175700	34.1257	5.0713	21.0000	48.0000
Age^2	175700	1190.2820	350.7040	441.0000	2304.0000
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 in 1951	175700	3.6073	1.1574	0.7297	5.8864
M.P with Disposable Income in 1961	175700	5.2744	0.1332	5.0607	5.6344
Unemployment rate in 1961	560040	0.0454	0.0156	0.0225	0.0861
Economic Infrastructures	175700	4.6383	0.3172	3.8662	5.1933

All variables (except Sex, Age, Age^2 , 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 for 1951 is the log of current billion liras.

Table 2. Regression results for male prime-age workers: all industries. Dependend valiable ln(wage)

	(1)	(2)	(3)	(4)
Age	0.0481***	0.0485***	0.0456***	0.0466***
	(0.0015)	(0.0010)	(0.0080)	(0.0080)
Age^2	-0.0005***	-0.0005***	-0.0005***	-0.0005***
	(0.0000)	(0.0000)	(0.0001)	(0.0000)
Firmsize				0.0194***
				(0.0004)
Bc Dummy	-0.7619***	-0.3567***	-0.2132***	-0.2149***
	(0.0049)	(0.0037)	(0.0041)	(0.0041)
Wc Dummy	-0.4891***	-0.1771***	-0.1452***	-0.1466***
	(0.0049)	(0.0031)	(0.0031)	(0.0031)
Specialization	0.0055***	0.0030***	0.0015*	0.0008
	(0.0006)	(0.0008)	(0.0008)	(0.0008)
Density	0.0221***	0.0187***	0.0074***	0.0056***
	(0.0005)	(0.0008)	(0.0010)	(0.0011)
Market Potential	0.1088***	0.0912***	0.0500***	0.0453***
	(0.0021)	(0.0039)	(0.0058)	(0.0058)
Estimation mehtod	OLS	GLS	Within	Within
Time & Sector Dummies	Yes	Yes	Yes	Yes
Corr(u_i, Xb□			0.2559	0.3075
R^2	0.5249	0.4974	0.3596	0.4412
N. of individuals	24353	24353	24353	24353
N. of observations	175700	175700	175700	175700

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

Table 3. Regression results for all workers: all industries. Dependend valiable ln(wage)

	(1)	(2)	(3)	(4)
Sex	-0.2038***	-0.1894***		
	(0.0008)	(0.0018)		
Age	0.0305***	0.0293***	0.0201***	0.0195***
	(0.0002)	(0.0003)	(0.0029)	(0.0030)
Age^2	-0.0003***	-0.0003***	-0.0003***	-0.0003***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Firmsize				0.0174***
				(0.0003)
Bc Dummy	-0.8350***	-0.3682***	-0.1724***	-0.1753***
	(0.0031)	(0.0026)	(0.0029)	(0.0029)
Wc Dummy	-0.5412***	-0.1595***	-0.1060***	-0.1073***
	(0.0031)	(0.0023)	(0.0022)	(0.0023)
Specialization	0.0076***	0.0056***	0.0046***	0.0036***
	(0.0004)	(0.0005)	(0.0005)	(0.0005)
Density	0.0232***	0.0211***	0.0054***	0.0039***
	(0.0003)	(0.0005)	(0.0007)	(0.0007)
Market Potential	0.0814***	0.0727***	0.0290***	0.0260***
	(0.0013)	(0.0023)	(0.0041)	(0.0042)
Estimation mehtod	OLS	GLS	Within	Within
Time & Sector Dummies	Yes	Yes	Yes	Yes
Corr(u_i, Xb□			0.1957	0.2748
R^2	0.5318	0.5196	0.2607	0.3478
N. of individuals	92579	92579	92579	92579
N. of observations	560040	560040	560040	560040

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

Table 4. Regression results for male prime-age workers: manufacturing only. Dependend valiable ln(wage)

	(1)	(2)	(3)	(4)
Age	0.0539***	0.0446***	0.0453***	0.0468***
	(0.0019)	(0.0013)	(0.0073)	(0.0074)
Age^2	-0.0006***	-0.0005***	-0.0004***	-0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Firmsize				0.0228***
				(0.0007)
Bc Dummy	-0.8263***	-0.3927***	-0.2404***	-0.2417***
	(0.0070)	(0.0054)	(0.0060)	(0.0060)
Wc Dummy	-0.5322***	-0.1997***	-0.1625***	-0.1653***
	(0.0071)	(0.0046)	(0.0045)	(0.0046)
Specialization	0.0084***	0.0058***	0.0046***	0.0033**
	(0.0009)	(0.0011)	(0.0013)	(0.0013)
Density	0.0207***	0.0190***	0.0081***	0.0060***
	(0.0007)	(0.0012)	(0.0015)	(0.0016)
Market Potential	0.1084***	0.0875***	0.0251***	0.0233***
	(0.0031)	(0.0058)	(0.0095)	(0.0095)
Estimation mehtod	OLS	GLS	Within	Within
Time & Sector Dummies	Yes	Yes	Yes	Yes
Corr(u_i, Xb)			0.1915	0.2229
R^2	0.5146	0.4879	0.3525	0.4245
N. of individuals	13149	13149	13149	13149
N. of observations	87056	87056	87056	87056

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

Table 5. Regression results for all workers: manufacturing only. Dependend valiable ln(wage)

	(1)	(2)	(3)	(4)
Sex	-0.2177***	-0.2075***		
	(0.0010)	(0.0022)		
Age	0.0275***	0.0236***	0.0147***	0.0137***
	(0.0003)	(0.0004)	(0.0029)	(0.0030)
Age^2	-0.0003***	-0.0002***	-0.0002***	-0.0002***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Firmsize				0.0225***
				(0.0005)
Bc Dummy	-0.8984***	-0.4342***	-0.2002***	-0.2002***
	(0.0042)	(0.0038)	(0.0045)	(0.0045)
Wc Dummy	-0.5932***	-0.1946***	-0.1237***	-0.1254***
	(0.0043)	(0.0034)	(0.0034)	(0.0035)
Specialization	0.0076***	0.0076***	0.0081***	0.0069***
	(0.0005)	(0.0007)	(0.0009)	(0.0009)
Density	0.0234***	0.0220***	0.0038***	0.0024**
	(0.0004)	(0.0007)	(0.0011)	(0.0011)
Market Potential	0.0948***	0.0864***	0.0194***	0.0148**
	(0.0018)	(0.0034)	(0.0071)	(0.0071)
Estimation mehtod	OLS	GLS	Within	Within
Time & Sector Dummies	Yes	Yes	Yes	Yes
Corr(u_i, Xb□			0.1719	0.2657
R^2	0.5545	0.5355	0.2502	0.3541
N. of individuals	48038	48038	48038	48038
N. of observations	272470	272470	272470	272470

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

Table 6. GMM regression results for male prime-age workers: all industries. Dependend valiable ln(wage)

	(1)	(2)	(3)	(4)	(5)
Age	0.1009***	0.1007***	0.0316***	0.0073***	0.0072***
	(0.0019)	(0.0042)	(0.0036)	(0.0018)	(0.0019)
Age^2	-0.0005***	-0.0005***	-0.0003***	-0.0001***	-0.0001***
	(0.0001)	(0.0001)	(0.0001)	(0.0000)	(0.0000)
Firmsize	0.0157***	0.0186***	0.0159***	0.0141***	0.0108***
	(0.0012)	(0.0014)	(0.0017)	(0.0015)	(0.0017)
Bc Dummy	-0.1406***	-0.1277***	-0.1211***	-0.1904***	-0.1837***
	(0.0085)	(0.0091)	(0.0105)	(0.0103)	(0.0111)
Wc Dummy	-0.0821***	-0.0708***	-0.0572***	-0.0866***	-0.0815***
	(0.0058)	(0.0061)	(0.0065)	(0.0069)	(0.0075)
Specialization	0.0011	-0.0181*	-0.0172*	0.0030	-0.0114
	(0.0015)	(0.0106)	(0.0097)	(0.0041)	(0.0070)
Density	0.0042*	-0.1602***	-0.1034***	0.0264***	0.1717***
	(0.0024)	(0.0188)	(0.0175)	(0.0089)	(0.0366)
Market Potential	0.0422***	-0.0901	-0.6503***	-0.0176	0.0164
	(0.0137)	(0.1804)	(0.2004)	(0.0246)	(0.1717)
Lag $w_{i,t}$			0.3164***	0.8035***	0.7944***
			(0.0567)	(0.0124)	(0.0137)
Lag Specialization					0.0113**
					(0.0046)
Lag Density					-0.1156***
					(0.0299)
Lag Market Potential					-0.0255
					(0.1574)
Model estimeted in	Diff	Diff	Diff	Diff&Lev	Diff&Lev
Instruments Diff.	T-4 to T+4	T-4 to T-3	T-4 to T-3	T-4 to T-3	T-4 to T-3
Instruments Lev.				T-3 to T-2	T-3 to T-2
Time & Sector Dummies	Yes	Yes	Yes	Yes	Yes
Hansen test degr. freed.	138	24	32	63	60
Hansen test	671.44***	108.17***	155.25***	587.96***	483.81***
Test for AR(1)	-29.41***	-28.33***	-10.11***	-32.59***	-25.21***
Test for AR(2)	-6.32***	-7.05***	1.27	8.21***	7.62***
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 7. IV within regression results for male prime-age workers: all industries. Dependend valiable ln(wage)

	(1)	(2)	(3)	(4)	(5)
Age	0.0455***	0.0466***	0.0452***	0.0598***	0.0571***
	(0.0079)	(0.0080)	(0.0079)	(0.0067)	(0.0074)
Age^2	-0.0005***	-0.0005***	-0.0005***	-0.0004***	-0.0004***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Firmsize	0.0193***	0.0194***	0.0191***	0.0185***	0.0208***
	(0.0004)	(0.0004)	(0.0004)	(0.0010)	(0.0014)
Bc Dummy	-0.2114***	-0.2148***	-0.2112***	-0.1944***	-0.1832***
	(0.0041)	(0.0041)	(0.0041)	(0.0091)	(0.0126)
Wc Dummy	-0.1461***	-0.1466***	-0.1456***	-0.1302***	-0.1265***
	(0.0031)	(0.0031)	(0.0031)	(0.0072)	(0.0103)
Specialization	-0.0037		0.0009	0.0059	0.0213
	(0.0035)		(0.0008)	(0.0122)	(0.0226)
Density	0.0026***	0.0026***	0.0038***	0.0062**	0.0286*
	(0.0010)	(0.0010)	(0.0014)	(0.0031)	(0.0156)
Market Potential	0.0425***	0.0433***	0.0650***	0.0642***	0.1444**
	(0.0080)	(0.0080)	(0.0082)	(0.0229)	(0.0686)
Estimation mehtod	IV - Within	IV - Within	Two-Steps	IV - Within	IV - Within
Time & Sector Dummies	Yes	Time only	Yes	Yes	Yes
R^2	0.4380	0.4071		0.4012	0.3923
Endog. Test (df=1)	1.5729	1.6001	39.3681***	0.4326	0.3218
Separate Estimations	No	No	No	Geography	Density
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. Data in Columns (4) and (5) are averages of the separate regressions.