# Job Search in Thick Markets

Sabrina Di Addario*

15th September 2006

**Abstract**

I analyze empirically the effects of both urban and industrial agglomeration on men's and women's search behavior and on the efficiency of matching. The analysis is based on on a unique panel data set from the Italian Labor Force Survey micro-data, which covers 520 randomly drawn Local Labor Market Areas (66 per cent of the total) over the four quarters of 2002. I compute transition probabilities from non-employment to employment by jointly estimating the probability of searching and the probability of finding a job conditional on having searched, and I test whether these are affected by urbanization and/or industry localization. The main results indicate that both urbanization and industry localization raise job seekers' chances of finding employment (conditional on having searched), but neither of them affects non-employed individuals' search behavior.

JEL classification: J64, R00, J60.

Keywords: Labor market transitions, search intensity, urbanization, industry localization.

*Bank of Italy and University of Oxford. E-mail address: sabrina.diaddario@bancaditalia.it. This work owes much to extensive discussions with Eliana Viviano. I am particularly grateful for her invaluable help with the attrition analysis and with some aspects of programming. I also thank Erich Battistin, Simon Burgess, Massimo Caruso, Harald Dale-Olsen, Mary Gregory, William Greene, Vernon Henderson, Barbara Petrongolo, Alfonso Rosolia, Margaret Stevens, William Strange and two anonymous referees for their helpful comments and suggestions. Finally, I thank the participants to: the CEPR conference "Integration and Technological Change: Challenges for European Regions" (11-13 June 2004, Paris); the EALE/SOLE 2nd World Meeting (2-5th June 2005, San Francisco); the Society for Economic Dynamics Annual Meeting (23-25th June 2005, Budapest); and the EEA/ESEM Annual Congress (24th-28th August, Vienna). The views expressed in this paper are those of the author and do not involve the responsibility of the Bank.

# 1 Introduction

Matching models are widely used to analyze the process of job formation in the presence of labor market frictions. These models are typically taken to operate, and empirically estimated, at the national level (see Petrongolo and Pissarides (2001) for a survey). In a context of slow mobility of labor, however, the matching of workers and jobs may occur instead at a much more localized level (e.g., at the local labor market level), and in particular, it may be affected by the degree of urban or industrial agglomeration.[1] Furthermore, the majority of the literature analyzes labor market dynamics by focusing on the unconditional hazard rate into employment. However, since the latter is the product of the probability of searching and the probability of finding a job conditional on having searched, it would also be interesting to explore whether transitions to employment are due to the effort individuals devote to job seeking and / or to the employment chances per unit of search.[2] This distinction is even more important in the context of this study, as local hazard rates and job seekers' propensity to search are likely to be differently affected by agglomeration externalities: the former through changes in labor market tightness (i.e., the ratio between the amount of vacancies and the number of job seekers) and in the technology of matching; the latter through individual resources, search costs and returns, and hazard rates.

In this paper I empirically analyze the impact of agglomeration on both the individual's search intensity and the hazard rate into employment. Even though the final impact is not *a priori* obvious, the majority of the transmission channels have a positive effect on both the two stages of the search process (see Section 2.1 for more details on the predictions of the theory and Table 1 for a summarizing scheme). Indeed, a shorter distance to job interviews, more frequent "face-to-face contacts", and the presence of thicker informal networks lowering information asymmetries may reduce both commuting and information-gathering costs, increasing search intensity.[3] Another factor on the cost side that may induce individuals to search more intensively in the most agglomerated areas is the higher cost of living (e.g., housing costs), which raises the opportunity cost of staying unemployed. On the return side, agglomeration may increase job seekers' search intensity by raising local wages or improving hazard rates. The latter, in turn, depend on the intensity of job

[1] For instance, local markets may differ in the presence of skill heterogeneities: agglomeration may lower the degree of mismatch between the skills required by firms and those offered by workers, improving the quality of the match. Also, denser markets may be characterized by a lower degree of information imperfection. Finally, congestion depends on population and firm density, which may vary to a great extent across local markets.

[2] Peracchi and Viviano (2004) are one of the few exceptions in the literature exploiting this relationship.

[3] On the other hand, but perhaps less importantly, congestion might increase search costs (e.g., time spent in traffic jams) and hence lower the intensity of search.

advertising, the thickness of the labor market, and the technology of matching. While there is some empirical evidence of higher wages in agglomerated areas, the net effect of agglomeration on labor market tightness and on the technology of matching is less clear-cut. Indeed, agglomeration may raise both the demand and the supply of labor, so that it is not obvious whether it would make markets more or less tight. With regards to the technology of matching, whether the size of the market improves or depresses the contact rate (per unit of search) depends on whether "thick market" externalities dominate over congestion effects (see Petrongolo and Pissarides, 2001). Finally, the matching process may be more efficient in the areas where specialized workers with similar skills and firms of the same type are pooled together (Marshall's "labor pooling hypothesis"). However, the expectation of higher wage offers might increase individuals' choosiness, lowering the probability of job offer acceptance and therefore hazard rates. Which of these effects will prevail is thus a matter of empirical investigation.

In the empirical analysis I use the Italian Labor Force Survey micro-data to estimate the effects of agglomeration on employment probabilities and job search intensity. First, to measure the effects of *urban agglomeration* I use a dummy for "large city", equal to one if the individual resides in a local labor market system (LLM) with a population above $404,526$ inhabitants. In contrast to the majority of the studies that use arbitrary cut-off points, I adopt the same threshold value devised by Di Addario and Patacchini (2006) on the basis of spatial autocorrelation analysis applied to Italian LLMs. However, since the spatial unit of analysis is crucial to determine the existence and extent of agglomeration externalities (Arzaghi and Henderson, 2005), I also use a continuous variable: the LLM population size.[4] Second, to measure the effects of *industrial agglomeration* I use, alternatively, an "industrial district" and a "super-district" dummy, denoting the LLMs with a high presence of small and medium sized manufacturing firms.[5] Since all (but one) super-districts have a population below the $404,526$ inhabitants, I am able to compare the labor market dynamics of the non-employed people living in urban or industrially agglomerated areas to those living in the rest of the country by partitioning the Italian territory into three sets of LLMs: large cities, small towns containing super-districts, and the rest of the economy. To my knowledge, the comparison

[4] According to Rosenthal and Strange (2004) the size of the area may matter, as externalities decay quickly over space (within 10 miles). However, the logarithm of LLM area is rarely significant in my regressions. While in theory both population size and density may generate agglomeration externalities on search behavior, in practise this does not seem to be the case in Italy and in the UK (for the latter, see Petrongolo and Pissarides, 2006).

[5] Industrial districts are spatially concentrated productive systems characterized by a large number of small firms specialized into one or few stages of a main manufacturing production. Specialization and inter-firm division of labor enable a district to achieve economies of scale that are external to the single firm but internal to the cluster as a whole. Super-districts, in turn, are a subset of industrial districts with the highest incidence of small and medium sized manufacturing employment (see Section 4.2 for further details).

of urbanization and industry localization effects[6] on search behavior and employment probabilities has not been analyzed before.

Overall, my results indicate that both urban and industrial agglomeration affect job seekers' hazard rates, but neither of them influences their search behavior. In particular, residing in a large city increases men's (women's) chances of finding a job by 6 percent (8 percent), while each $100,000$-inhabitant increase in LLM population raises job seekers's probability of employment by 1 percent (but only below the $2,400,000$-inhabitant threshold). With respect to industrial agglomeration, living in a super-district increases a man's (a woman's) probability of finding a job by 8 percent (5 percent). These results are robust to the use of an alternative econometric model correcting for sample selection. In this case, the positive externalities generated by localization appear only beyond the super-district threshold (i.e., there is no effect in industrial districts). As a robustness check, I run two separate regressions: one on the foreigners who have been resident in Italy for more than five years, the other one on the Italian residents, on the basis of the idea that under certain conditions (i.e., after five years of residence immigrants face same constraints and opportunities as Italians; labor mobility across LLMs is low) the former can be considered as movers and the latter as stayers. While results on stayers confirm the previous findings, those on the sub-sample of movers do not show any evidence of significant differences in the employment chances nor in search behavior between the non-employed individuals living in the most agglomerated areas and those residing elsewhere. Thus, the available data does not enable me to reject the hypothesis of the presence of agglomeration externalities in the matching process.

These findings suggest that the magnitude of the externalities generated by agglomeration on employment probabilities varies according to both the type and the degree of agglomeration considered. This has two main important policy implications.

First, if the spatial concentration of small and medium sized industrial firms improves the efficiency of matching, it might be advisable to favor the emergence or the development of industrial clusters.[7] However, my results indicate that not all industrial districts reduce frictions, as the probability of finding a job per unit of search is significantly higher in super-districts but not in

[6] Similarly to Rosenthal and Strange (2004), I use the term *urbanization* to mean urban agglomeration, and the term *localization* as a synonymous of industrial agglomeration.

[7] Although this is a controversial issue. According to some authors (e.g., Putnam, 1993) the genesis of Italian industrial districts has been a slow process, with roots in historical events that took place centuries ago, and thus cannot be fostered by any policy. Nevertheless, since the 1990s Italy provides subsidies to promote and sustain industrial districts. The Budget Law for the year 2006 (22nd December 2005; articles $366-372$), for instance, establishes that firms belonging to industrial districts can choose to pay taxes through the District as an institution (rather than individually). In this case, the District is also entitled to provide private banks guarantees to lower the capital adequacy that each firm has to fulfil in order to meet the Basle requirements when applying for a loan.

the other industrial districts. While the super-districts subset has been identified out of industrial districts on the basis of statistical criteria (namely, firm size and sector concentration), it would be important to study more in detail whether they also differ along other lines (e.g., product quality, organization of the production process, etc.).

Second, the absence of urbanization effects on job seekers' hazard rates beyond the $2,400,000$-inhabitant threshold might imply that the largest cities (i.e., Rome, Milan and Naples) are "too big", possibly because of decreasing returns in the local matching function. Knowing whether these cities are over-sized is an important issue, since reducing their dimension (for a given industrial composition) would generate productivity gains.[8]

The paper is structured as follows. The next section presents the theoretical framework; Section 3 reports the empirical model, Section 4 the data set and the variables; Section 5 discusses the estimation results; and Section 6 concludes.

# 2 The theoretical framework

## 2.1 A simple model

In this section, I am going to present a simple model in order to identify the factors affecting search behavior and hazard rates that could differ between the more and the less agglomerated areas.

In the standard search and matching literature (for instance, Pissarides, 2000), the number of matches $M$ is expressed as an increasing and concave function of the amount of workers searching for employment and the number of vacant positions. Often, the matching function is assumed to be homogeneous of degree 1. To study the effects of agglomeration on search, I assume that: 1) the number of matches is an increasing function $\Psi$ of the job-finding rate; 2) the rate of job-finding is homogenous of degree 1 in both its arguments;[9] and 3) the national labor market is geographically segmented. Thus, every geographical unit or local labor market $j$ has a matching function specific to the area, both in terms of arguments (as in Patacchini and Zenou, 2006) and in terms of technology:

$$M_j = \Psi(m_j(s_j J_j, a_j V_j)) \tag{1}$$

where $J_j$ is the number of searchers in local labor market $j$, $s_j$ the area's average search intensity,

[8] In any case, being "too small" would be worst than being "too big", as the loss of real output per worker generated by under-sized cities is larger than that originating from oversize (Au and Henderson, 2006).

[9] Thus, the matching function is homothetic in both the number of vacancies and the number of job seekers.

$V_j$ the amount of vacancies, $a_j$ the area's intensity of job advertising, and $m_j$ the job-finding rate.

The rate of job-finding for an individual $i$ searching with intensity $s_{ij}$ is:

$$m(s_{ij}, a_j\theta_j) = s_{ij}\frac{\Psi(m_j(s_jJ_j, a_jV_j))}{s_jJ_j} = s_{ij}\frac{\Psi(m_j)}{m_j}h_j(a_j\theta_j) \tag{2}$$

where $h_j$ is the rate of matching per unit of search,[10] and $\theta_j = V_j/s_jJ_j$ is a measure of the area's labor market tightness.

Let a job seeker's budget constraint be:

$$b = C_j(s_{ij}) + p_jz_{ij} \tag{3}$$

with:

$$C_j(s_{ij}) = d_js_{ij}^{\gamma}, \gamma > 1 \tag{4}$$

where $b$ denotes the income of a non-employed person, $C_j(s_{ij})$ the cost of search, $z_{ij}$ a real consumption good bundle, and $p_j$ the area cost of living (e.g., housing costs). I assume that agents' utility from consumption $u(z_{ij})$ is an increasing and concave function of $z_{ij}$. The expected intertemporal utility (in steady state) achieved by an unemployed agent is therefore:

$$rW_{ij}^U = u\left(\frac{b - C_j(s_{ij})}{p_j}\right) + s_{ij}\frac{\Psi(m_j)}{m_j}h_j(a_j\theta_j)(W_{ij}^E - W_{ij}^U) \tag{5}$$

where $W_{ij}^E$ is her expected lifetime utility when currently employed and $r$ the discount rate.

The optimal level of search intensity $s_{ij}^*$ a job seeker will exercise is that which maximizes (5): $\partial W_{ij}^U/\partial s_{ij} = 0$, or (at an interior solution):

$$u'(z_{ij})\frac{C_j'(s_{ij})}{p_j} = \frac{\Psi(m_j)}{m_j}h_j(a_j\theta_j)(W_{ij}^E - W_{ij}^U) \tag{6}$$

Job seekers are hence faced with a trade-off between the marginal cost of increased search effort in terms of current consumption and the marginal increase in their chances of finding a job that it induces. Thus, whether search is more intense in agglomerated areas depends on whether labor

[10] That is, the rate at which a worker searching with unit intensity will find a job, if $s_{ij}$ is normalized to be between 0 and 1. Under this normalization, in the empirical part of the paper (Section 4) I take $s_{ij}$ to be the probability of searching and $h_j$ to be the hazard rate (i.e., the probability of finding a job conditional on having searched). Since I do not intend to estimate specifically this structural model (which I am only using to understand the predicted dependencies), there does not need to be complete consistency between this and the empirical section.

market size lowers the costs of search and/or increases its returns. The *market size effect* $\Psi(m)/m$ determines whether the matching function does or does not exhibit increasing returns to scale.[11]

In the next section, I will thus take this simple model as the starting point to discuss the mechanisms through which agglomeration may affect individuals' search behavior.

### 2.2 The effects of agglomeration

As Section 2.1's model shows, on the cost side there are two channels through which agglomeration may increase search: search costs and the cost of living (see also Table 1).

With respect to the former, a shorter distance to job interviews or more frequent face-to-face contacts due to physical proximity may reduce both transportation costs and the costs of acquisition of information on vacancies.[12] In denser areas, search costs may be lower also because of the presence of thicker formal and informal networks facilitating the diffusion of information on job opportunities.[13] It should be noted, however, that congestion (e.g., more intense traffic jams, crowded buses, etc.) may, on the contrary, increase search costs and thus reduce individuals' search propensity.

With regards to the cost of living, the most congested areas are likely to suffer from higher house prices and rents, which, by increasing the cost of staying unemployed with respect to lower-density areas, should induce job seekers to search more intensively (Smith and Zenou, 2003). This effect occurs whenever the unemployment benefit $b$ is either fixed or less responsive to the local cost of living $p_j$ than local nominal wages; in fact, there is evidence that wages are actually higher in denser areas, and $b$ will include some nationally determined benefits that are not indexed for local cost-of-living.

On the return side (the hazard rate and the market size effects), there are four main channels through which agglomeration can intensify search: wages, labor market tightness, vacancy advertisement, and the technology of matching (see Section 2.1's model).

[11] Plausibly, $\Psi$ is first convex, then concave (i.e., logistic shaped), in which case the matching function would first show increasing and then decreasing returns to scale (the latter avoids that in equilibrium all people and firms are agglomerated in one single city).

[12] In Wasmer and Zenou (2002) for instance, a longer distance to jobs lowers the quality of information on vacancies, and thus individuals' search efficiency and unemployment. Note that the presence of inner-city unemployment generated by spatial segregation from the major urban zones of employment is usually referred to as the "spatial mismatch hypothesis", which arises when search is costly and firms' entry costs are lower in the locations (e.g., the suburbs) far away from the zones of residence of the unemployed workers (e.g., the center; Coulson, Laing and Wang, 2001).

[13] Coulson, Laing and Wang (2001), for instance, show that the higher is the number of neighbors employed in the area where vacancies are located, the higher are the chances for unemployed job seekers to find a job there ("neighborhood externalities"). See also Wahba and Zenou (2005) and Ioannides and Loury (2004).

First, job seekers may search more intensively in agglomerated areas because they have a higher utility from employment than elsewhere. Indeed, according to the literature on agglomeration, in larger labor markets wages may be higher than average because of the productivity gains generated by the Marshallian externalities.[14]

Second, agglomeration might increase labor market tightness and thus raise both hazard rates and individuals' search intensity. Indeed, even if there would be reasons to expect the number of applications to be higher than in non-agglomerated zones,[15] it can be shown that their increase is likely to be smaller than that of vacancies.[16]

Third, agglomeration may increase job seekers' propensity to search by intensifying firms' job advertising, for mainly three reasons.[17] Firstly, because the existence of thicker networks[18] may reduce the cost incurred by firms in advertising their vacant positions. Secondly, because the higher number of job seekers may allow employers to more easily cover any fixed costs of advertisement.[19] Thirdly, because of a greater average labor productivity.[20] In all these cases, job seekers exercise more effort simply because they have better chances to find a job and are hence more encouraged to search than elsewhere.[21]

Finally, search intensity depends on the technology of matching. Agglomeration may improve both the chances and the quality of matching.[22] With respect to the former, a higher probability

[14] For empirical results on higher urban wages see, for instance, Glaeser and Mare' (2001) for the US and Di Addario and Patacchini (2006) for Italy, though de Blasio and Di Addario (2005) find no evidence of different average earnings in the Italian industrially agglomerated areas (i.e, industrial districts and super-districts).

[15] Thus, whether markets are more or less tight in agglomerated areas is itself a question of empirical investigation. Since there are no reliable data on vacancies in Italy, I cannot empirically test the existence of differentials in local labor market tightness due to agglomeration. These can only be inferred from the impact of urbanization and localization on individual hazard rates, which are increasing in market tightness and can be measured directly (see Section 5).

[16] Helsley's and Strange's (1990) model, for instance, shows that the competition externality that firms generate when locating in a city (due to the fact that other firms' profits are reduced) prevails on the productivity externality (due to the fact that the productivity of all workers is enhanced). Under free entry, this leads to "too many" firms in cities, which implies, other things being equal, a higher vacancy-to-unemployment ratio.

[17] However, note that if the most agglomerated areas were characterized by tighter labor markets they would also exhibit less intense job advertising, since in this case a lower chance of filling their vacancies would discourage firms from advertising their positions (a sort of "discouraged-job" effect).

[18] Networks can either be informal (e.g., Marshall's "industrial atmosphere") or real network agencies (as in Arzaghi and Henderson, 2005). For a survey on job information networks, see Ioannides and Loury (2004).

[19] In Wheeler (2001), for instance, per-worker firm recruitment costs decrease with population density, as the frequency of interactions enhances the arrival rate of potential workers for a job opening, which has a fixed cost.

[20] See Pissarides (2000) for a partial equilibrium analysis of job advertising and Ciccone and Hall (1996) – among others – for the evidence on higher labor productivity in denser areas.

[21] As Pissarides (2000) notices, this is the reverse of the discouraged-worker effect.

[22] Note that agglomeration may also affect the elasticities of the matching function with respect to job seekers and vacancies, so as to generate increasing returns to scale. As a matter of fact, the majority of the empirical studies (see Petrongolo and Pissarides (2001) for a review) finds constant returns to scale in the aggregate matching function, possibly because reservation wages adjust to offset the scale effects generated in the contact technology or in the

of matching may derive from the greater concentration and / or specialization of matching agents in agglomerated areas, which could increase the effective job contact rate, and thus the hazard rate.[23] With respect to the quality of matches, according to Marshall's "labor pooling hypothesis" agglomeration improves the efficiency of matching between jobs and workers, as the areas where many specialized firms concentrate tend to attract the job seekers with the specific skills required. Thus, the better expected quality of matches may raise the job seekers' probability of acceptance as firms make more attractive offers.

In principle, all the positive effects on hazard rates could be partially or completely offset by higher reservation wages, which increase job seekers' choosiness, lower their acceptance probability, the hazard rates and thus their intensity of search.[24] Reservation wages could increase because of higher expectations of future earnings or because of improved contact rates (per unit of search). Petrongolo and Pissarides (2006) suggest that when agglomeration improves the quality of matches and/or the mean of the wage offer distribution increases, job seekers raise their reservation wages so as to offset the scale effects generated in the contact technology or in the productivity of job matches.[25] Conversely, when agglomeration raises the arrival rate of job offers (for instance, through a higher vacancy-to-unemployment ratio), hazard rates tend to increase while individual wages do not.

In conclusion, even though the equilibrium generating these externalities is rather complex, agglomeration is likely to reduce the costs of search and to increase its benefits, generating a positive effect on both hazard rates and search intensity.

---

productivity of job matches (Petrongolo and Pissarides, 2006).

[23] Note, however, that a higher density may actually lower the meeting rate if congestion effects dominate over thick markets externalities (see Petrongolo and Pissarides, 2001). Besides the negative externality generated by a job seeker on the other, other sources of congestion may derive from local "dis-amenities" such as more traffic jams, crowded subways, pollution, etc. Which type of external (dis)economy will prevail is, ultimately, a matter of empirical investigation. For a survey on agglomeration externalities see Rosenthal and Strange (2004) and Duranton and Puga (2004).

[24] Note, however, that the other side of the coin is that firms become less choosey about whom they hire as their difficulties in filling vacancies raise. Moreover, a greater selectiveness increases the efficiency of the match. In Andersson, Burgess and Lane (2004), for instance, agglomeration increases the job offer arrival rate, enabling a more (positive) assortative matching between workers and jobs, and thus a higher productivity (given the complementarities in production). Similarly, in Berliant, Reed III and Wang (2006), the more the matched agents are complementary (i.e., each endowed with a different type of knowledge from the other), the higher the match efficiency. In dense areas agents are more selective (because they are more likely to meet with a potential partner), and thus matching more efficient.

[25] This might explain why, as a matter of fact, the majority of the empirical studies (see Petrongolo and Pissarides (2001) for a review) finds constant returns to scale in the aggregate matching function.

# 3 The empirical model

As I showed in the previous section (equation (6)), the transition probabilities from non-employment into employment depend on two elements, one determined by agents' search behavior and the other one by the matching process. In order to empirically examine the impact of agglomeration on the transition probabilities between labor market states, thus, one needs to find measures of both the individual's propensity to search and the effectiveness of matching.

A non-employed at time $t$ can be in one of the possible three states at time $t+1$ :

1. they sought employment between $t$ and $t+1$ and found a job ($E_{t+1}$);

2. they sought employment between $t$ and $t+1$ but did not find a job ($U_{t+1}$);

3. they did not seek employment between $t$ and $t+1$ ($O_{t+1}$).

I shall define $s_{it}$ as the probability that a non-employed person looks for a job at time $t$,[26] and $h_{it}$ as the probability that she finds employment at time $t+1$, conditional on having searched. Let $\tilde{s}_{it}$ be the latent variable determining whether a non-employed person looks for a job at time $t$ (i.e., the difference in her expected utility from searching and not searching) and $\tilde{h}_{it}$ the variable determining whether a job seeker finds employment at time $t+1$ (incorporating both the likelihood of her meeting a prospective employer and the sign of the surplus generated by that match). Even though $\tilde{h}_{it}$ and $\tilde{s}_{it}$ are not observable, I can express them as a function of two non-coincident sets of individual and location-specific variables, $X_{it}$ and $Z_{it}$ (detailed in Section 5), available in the Labor Force Survey micro-data on labor market transitions:[27]

$$\tilde{h}_{it} = \beta' X_{it} + \epsilon_{1t} \tag{7}$$

and

$$\tilde{s}_{it} = \gamma' Z_{it} + \epsilon_{2t} \tag{8}$$

The probability of observing a person who has searched at time $t$ is thus $Pr(\gamma' Z_{it} + \epsilon_{2t} > 0 |\ Z_{it})$, which I assume to be a probit $\Phi(\gamma' Z_{it})$. Similarly, the probability of observing a job seeker finding a job at $t+1$ is $Pr(\beta' X_{it} + \epsilon_{1t} > 0 |\ X_{it}) = \Phi(\beta' X_{it})$.

[26] Note that in the theoretical model presented in Section 2.1, $s_{it}$ was a continuous variable greater of equal to zero denoting the number of search units supplied by the individual $i$. Here, without loss of generality, I am normalizing search intensity to be between zero and one.

[27] Even though in the estimations I allow for location-specific effects, in this exposition I take the geographic area indexes $j$ as implicit in the individual characteristics of agent $i$.

My econometric methodology will consist in the joint estimation of $s_{it}$ and $h_{it}$ by maximum likelihood. To ensure robustness, two alternative econometric specifications will be estimated.

I first consider a simple search model where (after controlling for observable characteristics) individuals can be treated as identical, in the sense of being randomly matched to vacancies. In this framework, the transition probability from non-employment into employment is the product of the probability of searching $s_{it}$ and the probability $h_{it}$ that a job seeker finds a job. Thus, I will estimate $s_{it}$ and $h_{it}$ by maximizing the following likelihood function (as in Peracchi and Viviano, 2004):[28]

$$L = \prod_{i \in \{E_{t+1}\}} [\Phi(\beta' X_i)][\Phi(\gamma' Z_i)] \prod_{i \in \{Ut+1\}} [1 - \Phi(\beta' X_i)][\Phi(\gamma' Z_i)] \prod_{i \in \{Ot+1\}} [1 - \Phi(\gamma' Z_i)] \tag{9}$$

If there was unobservable heterogeneity among workers, however, the probabilities of searching and finding a job (conditional on the $X_i$ and $Z_i$'s) would not be independent. I therefore correct the above maximum-likelihood estimation to take into account the fact that the hazard-rate equation can be estimated only on the censored sample of the agents who search ($Z_{it}\gamma + \varepsilon_{i2} > 0$). To do so I adopt the method proposed by van de Ven and van Praag (1981) for bivariate probit models with sample selection. In this case, the likelihood function is:

$$L = \prod_{i \in \{Et+1\}} \Phi_2(\beta' X_i, \gamma' Z_i, \rho) \prod_{i \in \{Ut+1\}} \Phi_2(-\beta' X_i, \gamma' Z_i, -\rho) \prod_{i \in \{Ot+1\}} [1 - \Phi(\gamma' Z_i)] \tag{10}$$

where $\Phi_2$ is the bivariate standard normal cumulative distribution of the joint probability of $s_{it}$ and $h_{it}$, and $\rho$ is the correlation between the error terms. This method corrects the bias that arises from using (9) when the error terms in equations (7) and (8) contain some common omitted variable.

The results of the two estimation methods are reported in Section 5.

# 4 The data

## 4.1 The data set

For the empirical estimation I use the Labor Force Survey (LFS), conducted in the year 2002 by the Italian National Statistical Office (Istat). This survey is the main source of information on individuals' working condition, unemployment and job search behavior, in addition to their personal

[28] A large part of the empirical literature on hazard functions (see Devine and Kiefer (1991) for a review) assumes that the error terms are distributed according to a logistic function. I adopt here a normal distribution to be consistent with the second econometric model (see below). In any case, I also tested all the specifications reported in Section 5 assuming a logistic distribution and obtained very similar results (available upon request).

characteristics. The survey is conducted quarterly in two stages: about $1,300$ municipalities are sampled at the first stage, and about $70,000$ households at the second one. The LFS follows a rotating scheme according to which each family is interviewed for two successive rounds, and then again for two other consecutive waves after two quarters of interruption, for a total of four times. So, theoretically 50 per cent of the sample is kept constant between two consecutive rounds.

The LFS has a natural longitudinal dimension with people followed up to fifteen months, but the (yearly) longitudinal files constructed by Istat on the basis of a stochastic matching algorithm[29] (recovering 90 percent of the potential sample) do not contain information on individuals' place of residence, and therefore cannot be used to study the effects of agglomeration on labor market dynamics. However, even though the linkage of individual records across surveys is made problematic by the lack of a personal identifier, I was able to reconstruct the longitudinal quarterly transitions with a deterministic method linking individuals' records on the basis of their place of residence, their family identifier and some time-invariant information (i.e., the date of birth and sex; see the Appendix for further details). This method enables me to recover 75 percent of the potential sample. In principle, the loss of the remaining observations could be a potential source of bias for my estimates in case it was not randomly distributed. However, when I test whether this loss is due to random reporting errors in the key variables or to the non-random exit of some individuals from the LFS (i.e., "attrition"; see the Appendix for the methodology adopted and the test outcome), the results confirm that the matching procedure I used to construct the panel dataset is appropriate for an analysis of labor market dynamics.

### 4.2 The agglomeration variables

In this paper most agglomeration variables are defined at the "local labor market" (LLM) level. LLMs are clusters of municipalities aggregated on the basis of the residents' daily commuting flows to their place of work.[30] LLMs are relatively self-contained, in that, by definition, they offer employment to at least 75 per cent of their residing workers, both with respect to the total number of workers in the area and with respect to the total number of residents. Exhaustive partitions of the territory based on worker commuting have been devised in many OECD countries,[31] since they

[29] For a thorough explanation of the differences between stochastic and deterministic methods, see Paggiaro and Torelli (1999).

[30] The flows are obtained from the 1991 Population Census data. I assigned each LFS observation to a LLM with an Istat's algorithm matching LLMs to municipalities. Note that LLMs are computed by Istat; since I do not have access to Census data, I am not able to construct alternative indexes (a' la Gautier and Teulings (2003a), for instance).

[31] The UK, for instance, has been divided into 308 "Travel-To-Work Areas" (OECD, 2002).

reflect local labor market conditions better than administrative areas do. The literature on matching is increasingly basing the empirical analysis on LLMs, in order to avoid a geographical aggregation bias in contexts of imperfect labor mobility. The geographical reach of agglomeration externalities is itself at the center of the literature debate, and may depend on the specific phenomenon analyzed.[32] In this respect, the characteristic of self-containment makes LLMs particularly suited to be my spatial unit of analysis, since it enhances, by construction, the likelihood that a job seeker searches within the boundaries of the labor market where he resides.

Various measures of agglomeration, both urban and industrial, are examined.

*Urbanization* is measured with the LLM population size.[33] Since the absolute level of population increases very gradually across LLMs, with the largest variations occurring only at the upper end of the distribution, I also use a large-city dummy to test whether agglomeration economies manifest themselves only beyond a certain threshold value. Nevertheless, the choice of a threshold defining a large city is not a straight-forward issue; it should not be arbitrary and should plausibly be country-specific.[34] Thus, this paper adopts the threshold level of $404,526$ inhabitants devised by Di Addario and Patacchini (2006) on the basis of spatial autocorrelation analysis applied on Italian LLMs.[35] The intuition behind this methodology is that in order for a LLM to be classified as a large city, its population: 1) must be above the national average, and 2) must not be uniformly distributed (i.e., it must show a significant correlation with that of the neighboring LLMs).[36] Finally, in order to check the sensitivity of the results to the presence of outliers I replicate all the estimations on the sub-sample excluding the three largest LLMs (those with a population above $2,400,000$ inhabitants).[37]

*Industry localization* is measured by two alternative dummies denoting the incidence of LLM small-firm manufacturing employment: "industrial districts" and "super-districts". Industrial dis-

[32] See Arzaghi and Henderson (2005) for a discussion on this issue and Petrongolo and Pissarides (2001) for a review of matching studies based on LLMs.

[33] I also tested the joint effect of logarithm of LLM population size and logarithm of LLM area, but the latter was never significant. Also Petrongolo and Pissarides (2006) make a case for using the UK's Travel-To-Work Areas' size rather than their density, in contrast with the earlier literature (e.g., Ciccone and Hall (1996), Ciccone (2002), or Coles and Smith, 1996), which stated that density is more important than population or employment size in generating externalities.

[34] The Italian population, for instance, is much more dispersed over the territory than the US one, suggesting the use of different threshold values in the two countries.

[35] More specifically, the authors define a LLM as a large city if it lies in either the HH or in the HL quadrant of the Moran Scatterplot and if it is associated to a significant local Moran's $I$ statistic. The $404,526$-inhabitant threshold corresponds to the lower bound of the LLM population distribution in the large-city set.

[36] Note that the surrounding LLMs, chosen on the basis of a $k$-nearest neighbor weight matrix, are not part of the large city itself.

[37] That is, the LLMs containing Rome, Milan and Naples (the three largest municipalities in the Center, North, and South of the country). The population level of the remaining LLMs is below $1,500,000$ inhabitants.

tricts are identified by an Istat's algorithm that associates to each LLM a dummy variable equal to one if the area shows both a dominant sectoral specialization and a higher-than-average share of small and medium enterprises and manufacturing employment.[38] As the threshold values used to single out industrial districts are somewhat arbitrary, I also use a stricter definition: the super-districts, which are simply an industrial district subset with a higher share of both manufacturing and small and medium enterprises employment (see Cannari and Signorini (2000) for the identification criteria).

### 4.3 Italy: a good case study

The LLM characteristic of self-containment together with a very limited mobility of labor, make Italy a good case study for analyzing agglomeration effects, as under these conditions LLMs can conceivably be considered as separated markets, and this minimizes the possible problems of self-selection. If, on the contrary, the urbanization and localization variables were endogenous (e.g., because correlated to some omitted unobservable factor), the agglomeration effects on hazard rates and search intensity would not be correctly detected. For instance, if it were the case that the most able job seekers moved to the largest cities,[39] the urbanization effect on hazard rates would be biased upwards (provided that the probability of finding a job increased with city size and that ability could be observed by the employer before forming the match). In contrast, if the more generous government support or the presence of a stronger informal labor market in the largest cities attracted particularly the less able or lazier people, the urbanization coefficients on hazards would be biased downwards.

However, the risk that either the most or the least able people move to the most agglomerated areas is relatively little in Italy, since labor mobility is, in general, particularly low. Indeed, even the unemployed job seekers, who are generally the most likely to migrate (Dohmen, 2005), are unwilling to move out of their town of residence to find a job. As Table 2 shows, almost 80 percent of the non-employed Italians who look for a job are ready to accept an offer only in their LLM of

[38] More specifically, an LLM is an industrial district if: (1) the share of LLM's manufacturing employment in total non-farm employment is higher than the corresponding share at the national level; (2) the LLM's share of small and medium enterprises manufacturing employment in total non-farm employment is higher than that at the national level; (3) for at least one sector, the ratio between the LLM's share of sector employment in total manufacturing employment and the corresponding share at the national level is greater than one; (4) in at least one sector for which the LLM's specialization index is greater than one, the LLM's share of small and medium enterprises employment in total employment is higher than the corresponding share at the national level (see Istat (1997) for further details).

[39] In a context where people have a preference for urban consumption amenities this phenomenon could occur because the most able individuals, who can command higher wages, might be better capable of affording the large cities' higher cost of living (in Venables (2002), for instance, big cities' crowding costs select the high quality workers).

residence, and more than 41 percent do not intend to move from their own municipality.[40] The table also indicates that just 1.1 percent of the non-employed individuals in working age interviewed by the LFS in the four 2002 waves had been absent from their household of residence at the time of the interview for more than a year, and a merely 0.2 percent was also looking for a job. Moreover, none of the people interviewed changed municipality of residence between two consecutive 2002 quarters. This result is supported by Di Addario and Patacchini (2006), who, using data from the biannual Bank of Italy's Survey of Household Income and Wealth, find that none of the (about) $1,500$ employees present in the panel Section of the Survey changed residence between 1995 and 2002.[41]

Labor mobility has been decreasing over time, especially with respect to long-distance movements (Cannari, Nucci and Sestito, 2000): between 1960s and 1990s the share of inter-town changes of residence in total population fell from 0.3 to 0.2 percent. The authors show that a large part of this reduction is explained by a house price increase over the period in the areas with better employment perspectives relatively to the rest of the country (namely, the North versus the South).

Indeed, the rigidities in the Italian housing market can certainly discourage geographic mobility.

First of all, the presence of rent controls down-sizes the private rented sector, rationing rents and increasing workers' moving costs. The degree of imperfection of the Italian rental market is apparent from the figures on the distribution of rent contract types, reported in Table 3. In 2000, the share of non-liberalized rents was still surprisingly low: only 16 percent of rent contracts were in derogation from the rent-control law,[42] 35 percent of households were still under controlled rents ('equo canone' law), up to a quarter of contracts were informal, more than 16 per cent regarded council housing, and almost 5 percent were subsidized.

Secondly, the large transaction costs for buying and selling a house raise migration costs further and discourage owner-occupiers from becoming renters when relative price change,[43] thus increasing the bias towards owner-occupation. The share of owner-occupying households is indeed rather high in Italy (more than 70 percent of the total) and has been increasing over time, hampering mobility further (see Henley, 1998).[44] As a matter of fact, homeowners have a lower propensity to move

[40] In Italy there are about $8,100$ municipalities, amounting to an average of 10.3 municipalities per LLM.

[41] The figure on mobility amongst the employed individuals is rather low also according to the LFS (Table 2), which reports that 7.5 percent of the employees interviewed in 2002 were actually working in a province different from the one of their residence (this might include commuting).

[42] Before 1992 the 'equo canone' law put ceilings on rents. Afterwards rents were liberalized for new contracts, in derogation from the rent-control law (L.359/1992).

[43] In Italy tenure choices may be less responsive to prices than in the US, where the housing market is characterized by a high residential mobility across States.

[44] Note that according to Dohmen (2005): 1) high homeownership rates lead to greater unemployment, and

than renters (after controlling for individual observable characteristics; Di Addario, 2002).[45] The propensity to change house is generally low even within the same city: figures from the 2000 Bank of Italy's Survey of Household Income and Wealth indicate that only 7 percent of households are planning to change house in the next two years.[46]

Finally, the sub-optimal size of the market rented sector together with the high transaction costs for buying and selling a house may also bias people's choices towards daily commuting rather than change of residence. However, this would not raise endogeneity issues in my agglomeration variables, since they are defined on the basis of LLMs, which are self-contained precisely in terms of workers' daily commuting flows.

Thus, my estimates are unlikely to be substantially affected by the endogeneity problems deriving from selective migration to the most agglomerated areas. Nevertheless, this potential problem will be dealt with more directly in Section 5.2.3.

### 4.4 Descriptive statistics

In 2002 LFS surveyed $777,248$ individuals. In order to analyze transition probabilities I restricted the sample to the people who were surveyed for at least two consecutive waves. Since my analysis concerns the labor market dynamics of non-employed persons, I also excluded the individuals already employed at time $t$ and those either below the age of 15 or above that of 64. After excluding the persons for whom there were missing observations on the relevant variables, the data set comprises $71,247$ non-employed individuals, $11,276$ of which job seekers.

Table 4 presents the descriptive statistics for the whole sample and for the large-city, super-district and industrial district sub-samples. From a comparison across the columns, it is apparent that the sample's characteristics do not vary much between the most agglomerated areas and the rest of the economy. For instance, the sample composition in terms of human capital is very similar across the columns, with education being slightly higher in large cities and age in the industrial-denser areas. As expected, industrial districts and super-districts exhibit a more intense labor turnover than the rest of the country: a higher-than-average share of non-employed people

2) migration is more sensitive to wage than to unemployment differentials. Indeed, after controlling for individual characteristics, the probability of owner-occupying is higher in the South of Italy (Di Addario, 2002), where migration rates are low in spite of the presence of higher unemployment rates than in the North (see Table 6). Also in line with Dohmen's (2005) theory, in Italy wage differentials over the territory are rather small in size.

[45] The author also shows that immigrants are less likely to buy the house of residence, confirming a greater difficulty or reluctance to settle in a province different from one's own.

[46] The data does not enable me to tell whether people intend to change house within or across LLMs, but since the most frequently reported motivation for moving is the purchase of a house, I presume that the majority of the expected moves would be within the same municipality.

had previous work experience (while in large cities it is the reverse) and a greater percentage of job seekers had been searching for less than one year (93 percent in super-districts, 85 percent elsewhere). Finally, in the most industrially agglomerated areas individuals have a higher (lower) than average number of employed (non-employed) family members, which might imply an easier access to thicker and better-quality informal networks.

In Italy there are 784 LLMs. LLM population size, density and area vary greatly. The mean population size is 73,424 inhabitants, ranging from 2,901 in Limone sul Garda to 3,311,431 in Rome. Density ranges from a minimum of 10 inhabitants per square Kms. (Crodo) to a maximum of 3,250 (Naples), with a mean of 184.6. Finally, the mean of the LLM area distribution is 384 square Kms., ranging from 10.4 (Capri) to 3,539 (Rome). Nineteen of the 784 LLMs have a population above the 404,526 inhabitant threshold, 199 are classified as industrial districts, and 99 as super-districts. My sample includes 518 LLMs (66 percent of the total; see Table 4) and comprises an average of 138 individuals per LLM. Since the LFS is stratified to represent Italian regions and municipalities, all the 19 large cities are always sampled (for a total of 20,335 observations).[47] Furthermore, even though the LFS was not designed to represent the industrial district or super-district population, the sample distribution reflects that found at the national level: in my sample, 28 percent of LLMs are classified as industrial districts (25 percent in Italy) and 13 percent as super-districts (same at the national level).[48]

Table 5 reports the quarterly transition probabilities and flows both at the aggregate level and for men and women separately. The transition matrix shows that in Italy there is a high unemployment persistence, as almost 64 percent of the people unemployed in the quarter preceding the interview are still unemployed in the successive quarter. While these numbers are very similar for men and women, significant gender differences can be found in other respects. First, in the average probability of finding a job, conditional on being non-employed at time $t$: the transition probability from unemployment into employment is more than 16 percent for men and only 11 percent for women, and the respective probabilities of finding a job for those recorded as inactive at time $t$ are 5 and 3 percent respectively.[49] Second, the transition probability from unemployment into inaction, greater than that into employment for both sexes, is much larger for women than for men (in line with other empirical results, e.g., Broersma and Van Ours, 1999). Finally, Table 5

[47] These are (in descending order of population levels): Rome, Milan, Naples, Turin, Bari, Florence, Genoa, Palermo, Bologna, Catania, Venice, Padua, Desio, Taranto, Verona, Bergamo, Cagliari, Como and Lecce.

[48] For a total of 12,863 individuals sampled in industrial districts and 5,285 in super-districts.

[49] However, when expressed in percentage of the working age population, the flows from inactivity to employment are slightly larger for women than for men.

shows that the flows from inactivity to employment as a percentage of the working age population are generally more substantial than those from unemployment into employment (1.4 versus 0.8 percent; in line with previous results, e.g., Petrongolo and Pissarides, 2001). In light of this fact, and consistently with the most recent literature (Broersma and Van Ours (1999); Brandolini *et al.*, 2004), I shall estimate hazards from non-employment to employment rather than from unemployment.

The Italian labor market is known to be segmented with respect to territory (see, for instance, Peracchi and Viviano, 2004). While, traditionally, labor market conditions are analyzed at the macro-area level (North, Center, and South),[50] I examine whether they also differ along the degree of urban and / or industrial agglomeration. Table 6 reports Istat's statistics for the year 2002 on the employment, unemployment and activity rates for all the agglomeration units considered in this paper (large cities, industrial districts, super-districts, and industry-thin small-sized towns). It also shows the hazard rate into employment and the share of job seekers in the total non-employed population. The former is the probability that a job seeker finds a job between successive quarters; in the table it is computed as the mean of a dummy variable equal to one if the individual looks for a job at time $t$ and moves into employment at time $t+1$. The latter is computed as the mean of a dummy variable equal to one if a non-employed person at time $t$[51] is either employed or non-employed in the following quarter. Note that in this paper the pool of job seekers is larger than the set of the people recorded as unemployed according to the ILO definition. This is because, having only quarterly data (higher frequency data do not exist in Italy), I have to assume that each search period (the time interval between $t$ and $t+1$) lasts three months – in line with a large part of the empirical literature on matching (see Petrongolo and Pissarides (2001) for a survey). Thus, to ensure temporal consistency between stock and flow data (transitions to employment) the job seekers' pool must comprise all non-employed people, willing to start working immediately, whose last search action took place in the previous quarter – rather than in the previous month, as it is in the ILO definition (see Brandolini *et al.* (2004), and Peracchi and Viviano (2004) for a discussion).

In 2002 the unemployment rate ranged from a minimum of 4 percent in super-districts to a maximum of 12 percent in the least agglomerated areas. Conversely, employment rates were lowest in the small non-industrial towns and highest in super-districts (41 percent against about

[50] In 2002, for instance, unemployment rates ranged from 3 percent, on average, in the North-East to 14 percent in the South, while employment rates ranged, respectively, from 64 percent to 50 percent (see Table 6).

[51] That is, someone who at time $t$: a) undertook at least one search action in the previous 30 days (including the individuals searching for the first time); or b) searched, even if not actively; or c) did not search, but was willing to work.

50 percent). These patterns are largely confirmed at the macro-area level, so that they cannot be explained by the fact that most industrial districts or super-districts are located in the regions of the Center-North-East of the country. With regards to labor market dynamics, the industrially denser areas show the lowest share of job seekers and the highest hazards to employment from non-employment (respectively, 10 and 46–53 percent). In contrast, large cities show the lowest hazards to employment (22 percent), probably because of the greater stock of job seekers concurring for available jobs. These offsetting effects are mostly confirmed in all the Italian macro-areas.

The descriptive statistics of Table 6 would thus indicate that agglomeration is associated with specific labor market dynamics. In particular, these results suggest that search intensity is highest in large cities and hazard rates are highest (lowest) in the industrially agglomerated areas (large cities). The impact of agglomeration, however, can be better analyzed in a more comprehensive model where the features of the local labor markets and the characteristics of individuals are taken into account.

# 5 Empirical analysis

I now turn to the empirical estimation of the determinants of individual search intensities and hazard rates, examining in particular whether these probabilities differ between agglomerated and non-agglomerated areas. The estimations were conducted separately for men and women and, unsurprisingly, labor market dynamics turned out to be substantially different for the two groups.

## 5.1 The empirical specification

The empirical models proposed in Section 3 can be used for this purpose. In the remainder of this section, I will first examine a baseline model estimating the parameters of the log-likelihood functions (9) and (10) on the basis of individual and local labor demand characteristics, then test the existence of agglomeration effects on both hazard rates to employment and search intensity.

The hazard rate to employment depends first of all on the variables affecting local labor demand conditions and the individual's productivity. The former are proxied with two set of indicators. First, two indexes meant to capture contemporaneous labor demand shocks: the share of employees working overtime in total workers and the average number of extra-hours worked.[52] The coefficients on these variables should be either significantly positive or zero, depending on whether demand

[52] I am aware that these indexes are imperfect proxy for demand, as they could also reflect supply-side conditions. Ideally, I should control for vacancies (even though the majority of hazard studies does not; Petrongolo and Pissarides, 2001), but there are no data for Italy.

expansion is or is not fully compensated by overtime work increases. In the latter case, a rise of overtime work would be accompanied by an increase in the number of vacancies, which, other things being equal, would improve the hazard rate. In contrast, if all the demand increase was entirely compensated by overtime work, my indicators should not affect the hazard rate. The second local labor market variable I consider is the geographical density of job seekers (similarly to Petrongolo, 2001).[53] Since, as shown in Section 2.1, hazard rates are increasing in local labor market tightness, I expect job seeker density to have a negative sign. The personal characteristics that I use to control for the individual's productivity are age, age squared, and educational attainment (first degree, high school, middle school). I also control for search duration (0–1 month, 1–5 months, 6–11 months), expecting it to be inversely related to the chances of finding a job, for a dummy denoting whether the individual had previous work experience, as well as for seasonal and geographical dummies. Finally, I control for the number of employed household members, which could be taken as a proxy of network quality. The idea is that family networks are important to find employment and that employed individuals have access to better quality networks than unemployed ones, as they presumably have more information on job offers.[54]

As seen in the theoretical model (equation (6)), an agent's optimal search intensity $s_{it}$ depends on the hazard rate $h_{it}$ into employment that he anticipates facing if he searches. In estimating the equation for search intensity, I therefore include all the individual and labor-market explanatory variables used in the hazard-rate equation. In order to identify the propensity to search, I also add the number of non-working people in the household,[55] and two proxies for the value (monetary and other) of non-search activities, which I expect to lower the probability of participation in any given application round (i.e., search intensity). These are: a) the individual's position within the household (single living alone, household head, and spouse); and b) the self-perceived work status (housewife, student, or retired).[56]

[53] Alternatively to the logarithm of job seekers, I also tested the effect of the logarithm of the total labor force and that of the population above the age of 15, with no different results.

[54] This is similar to Wahba and Zenou (2005), who proxy network quality with the number of family members in the labor force and consider it an agglomeration variable. The validity of this variable clearly relies on the absence of unobserved characteristics (such as ability) shared among family members.

[55] Using data at the provincial level from the *Consulente Immobiliare*, I also controlled for house prices and rents, but these were never significant. I used data for 2002, the oldest year available (1965 for house prices and 1993 for rents), and the average of the entire period.

[56] Since the household decisions are linked by a budget constraint, the position in the household may matter. Note that the sum of the three self-perceived work status dummies equals to being inactive at time $t$.

## 5.2 The results

### 5.2.1 Baseline model

Tables 7 and 8 present the results of the baseline model for men and for women, respectively. To show the robustness of my results, in each table I report the outcomes of both the econometric models discussed in Section 3 ((9) and (10)). In spite of the fact that the Wald-test always rejects the null hypothesis of zero correlation between the error terms, confirming the presence of a selection bias, the two estimation methods provide the same signs and statistical significance levels for almost all the regressors considered in the hazard rate equation (which is the one subject to the selection problem).

**a) Hazard rates**

In the baseline model for men (Table 7), hazard rates are higher for the individuals with previous work experience and better-quality family networks; they are lower in the South, for the more educated people and for the older population.[57] As expected, the probability of moving from non-employment into employment decreases with search duration (see, among others, Lancaster, 1979). In particular, individuals who have been searching for less than one month have a chance of finding a job twice as large as those who have been searching for more than one year.[58] Moreover, a higher LLMs' job seeker density reduces the individual's probability of finding a job, probably because of the congestion that unemployed workers create on each other (see Burgess (1993) or Petrongolo and Pissarides, 2001). Finally, neither the LLM share of overtime workers in total workers nor the LLM average extra-hours worked have any significant impact on hazard rates, possibly because demand increases are fully compensated by overtime work. In contrast to the male population, women have a higher chance to find a job when they are younger, when they have a University degree,[59] and when they live in the North-East, while the thickness of family networks does not affect their likelihood of finding a job (Table 8).[60]

[57] Even though these two last results are in contrast with some empirical studies on the UK (e.g., Lancaster, 1979), they are in line with previous findings on Italy (see, for instance, Peracchi and Viviano, 2004).

[58] Throughout the paper, marginal effects have been computed at the mean for the continuous variables and for a discrete change from 0 to 1 for the dummy variables.

[59] These results are less surprising than those for men, which could possibly derive from a different composition of the non-working population (e.g., a higher incidence of old women difficult to employ, such as long-term unemployed, or people with health problems), and/or from a greater choosiness of the most educated men (which could completely offset the positive effect of higher meeting rates).

[60] This could occur either because networking is a more male-oriented search channel, or because female networks are of a lower quality. It is also possible that women living in families where more members work have a higher reservation wage, as they can benefit from a higher income (in contrast, men might not "afford" to be choosey because of the different role they have in the household). In passing, note that the fact that the number of employed household members has an opposite effect for men and women contrasts with the hypothesis that this variable

**b) Search propensities**

Search intensity increases with age, education, past work experience, and with residing in the North-East. In contrast, students, retired workers and housewives search less intensively, probably because these categories of job seekers assign a higher value to non-search activities than those who perceive themselves as unemployed. Interestingly, the position in the household matters differently for the two sexes, as being a household head or a spouse increases the probability of searching for men but decreases it for women (with respect to being an offspring or having other positions within the household). This different behavior probably reflects the tendency for wives and mothers to stay at home, and a greater need for non-employed husbands and fathers, who are most often the primary earners in the household, to increase their search effort. The hypothesis that men and women differ in search behavior because the traditional household division implies that they face different (opportunity) costs of search is consistent with the finding that when the number of non-working individuals in the family increases only men raise their search effort.[61] Moreover, consistently with having higher chances of finding employment, the men who have better-quality family networks search more intensively, while women's behavior is not affected by the thickness of family networks. Finally, the LLM job seeker density is non-significant for either men nor women, implying that non-employed individuals do not exercise more effort when competition for vacant jobs raises.

### 5.2.2 Effects of agglomeration

To examine the effects of agglomeration on $s_i$ and $h_i$, I add the variables discussed in Section 4 to the baseline specification. Tables 9 and 10 summarize the results on hazard rates and search intensity for the two econometric models (9) and (10). In both tables, I first consider the joint effect of the large city and the super-district dummies (first specification).[62] I then substitute the large city variable with LLM population size, and test its effect with either the super-district or the industrial district dummy (second and third columns). In the last three columns I replicate the

captures, rather than network quality, unobservable ability shared by the members of the same family.

[61] Note that this may be due to child care, as Italy lacks of policies aimed at supporting mothers' employment. In order to examine this hypothesis further, I also ran the same regressions (not reported here) on the parent subsample, controlling for the number of children below the age of six. I find that a marginal increase in this variable lowers women's probability of searching by 1 percent (at 1 percent statistical significance), but does not affect men's behavior (for similar outcomes, see Del Boca, 2001).

[62] I also considered the effect of each of these variables separately, with no substantially different results. Note that whether the signs and the statistical significance of the urbanization and localization dummies can correctly identify agglomeration differentials in employment probabilities and search behavior clearly relies on LLMs to be separated markets (see, for instance, Coles and Smith (1996) or Duranton and Monastiriotis, 2002), as discussed in Section 4.3.

former specifications on the sub-sample excluding the three largest LLMs.[63]

Thus, after controlling for LLM job seekers' density, which captures the negative congestion externality exercised by the unemployed workers on each other (see Petrongolo, 2001), I find that urban agglomeration has an overall positive effect on the probability of finding a job. Indeed, as Table 9 shows, residing in a large city improves men's employment possibilities by 6 percent (at the 6 percent statistical significance level) and women's chances by 8 percent (at the 1 percent statistical significance level), both in the full and in the restricted samples (columns (9.1), (9.4), (9.7), and (9.10)). In contrast, the level of population is significant only once I exclude the three largest LLMs from the sample (at the 4-6 percent level for men and at the 1 percent level for women; columns (9.5)-(9.6) and (9.11)-(9.12)). In particular, each $100,000$-inhabitant increase raises both men's and women's probability of employment by 1 percent. This result implies that job seekers benefit from agglomeration externalities only below the very top of the population distribution. There are various reasons for why this could be the case. First, positive externalities may predominate over crowding effects only below the $2,400,000$-inhabitant threshold.[64] Second, the three largest cities may be over-sized with respect to employment possibilities.[65] Third, it is possible that in Rome, Milan and Naples the positive effect of agglomeration on meeting rates is fully compensated by a lower acceptance probability,[66] cancelling-out the final impact on hazard rates.

With respect to localization, searching in more industrially agglomerated areas raises mens' chances of finding a job by 8 percent in super-districts, by 4-5 percent in industrial districts (respectively, at the 4 and 8-11 percent statistical significance level; columns (9.1)-(9.6)). In contrast, women have a higher probability of finding a job only in super-districts (by 5 percent, at the 8-10 percent statistical significance level; columns (9.7)-(9.12)).

The positive externalities deriving from (sufficiently thick) industry localization are robust to controlling simultaneously for all the urbanization variables. When comparing the urbanization

[63] The number of observations drops from $25,116$ to $22,332$ in the men's sub-sample and from $46,131$ to $40,885$ in the women's case. The non-employed individuals residing in the excluded LLMs amount to $2,848$ for Rome, $1,835$ for Milan, and $3,530$ for Naples.

[64] Positive externalities could be due to the presence of tighter markets (more intense job advertising or more vacancies), urban wage premia, higher meeting rates, or better quality of matches; negative externalities might be generated by congestion (see Section 2.1).

[65] This may occur if job seekers chose to reside in the largest cities because of the amenities that these offer (e.g., cultural events, better quality of services, presence of infrastructures not available elsewhere, etc.), independently of the labor market conditions (so that they do not move elsewhere even if the chances of finding employment are reduced).

[66] This could happen if the three largest cities: a) exhibited a higher quality of matches than the rest of the country, b) job seekers expected firms to make more attractive offers than those located elsewhere, and c) job seekers' higher choosiness lowered their acceptance probability so as to offset their greater probability to meet a vacancy.

effects on hazard rates to those of industry localization, it is evident that in the men's sample the super-district coefficient is greater than the large-city one, while for women it is the reverse. This finding is even more apparent in Table 10, which examines the hazard rates per unit of search with the econometric model correcting for sample selection ((10)). In this case, for localization to create significantly positive net externalities a minimum degree of firm thickness is necessary. Indeed, searching in more industrially agglomerated areas raises the probability of finding employment (per unit of search) only above a certain threshold of manufacturing small-sized firm concentration. Thus, while residing in an industrial district has no effect on hazard rates, other things being equal, living in a super-district increases men's probability of finding a job (at the 4 percent statistical significance level; columns (10.1)-(10.2) and (10.4)-(10.5)), while the super-district localization effect on women's employment chances is only significant at the 11-13 percent level (specifications (10.7)-(10.8) and (10.10)-(10.11)). In contrast, the positive impact of urbanization is more significant for women than for men (respectively, at the 1-2 and 10-13 percent statistical significance level). A possible explanation of why industry localization (urbanization) improves more the matching of men (women) than that of women (men), is that women might apply for jobs (e.g., in the tertiary sector rather than in industry, in administration rather than in the production process, etc.) that benefit less (more) from industrial (urban) agglomeration externalities than those preferred by men.[67]

I now turn to the effects of agglomeration on men's and women's search behavior. The bottom part of Tables 9 and 10 shows the results.

In general, agglomeration does not affect either men's or women's behavior in any of the samples considered. Indeed, in spite of the fact that urbanization and localization improve their employment chances per unit of search, job seekers do not search more intensively in large cities nor in the more industrially agglomerated areas (columns (9.13)–(9.24) and (10.13)–(10.24)). This may seem

[67] Of course, it is also possible that the employers in the more industrially agglomerated areas segregate women (which would lower the probability of finding a job per unit of search; see Black, 1995). Apart from the most commonly reported reasons, this could occur if in super-districts, where the mastery of production is both accumulated over a lifetime and transmitted from generation to generation, the old-generation-male employees passed on their knowledge to their sons rather than to their daughters. Indeed, while not finding any presence of wage discrimination in Italian industrial districts, de Blasio and Di Addario (2005) find some evidence of vertical segregation (i.e., after controlling for observable individual characteristics, industrial-district female employees do not earn any differently than their male counterparts, but have a lower probability of becoming entrepreneurs than men). Alternatively, it is possible that super-district women have higher reservation wages and thus accept job offers less frequently than super-district men. This could occur if in super-districts, where the traditional division of labor in the household is likely to be more persistent than in large cities, women tend to decide the amount of labor to offer in the market on the basis of the whole family income rather than on that of their own (see Del Boca, 2001). However, neither of these two hypotheses would help explaining why urban agglomeration effects are less important for men than for women.

somewhat surprising, as job-seekers should increase their propensity to search when their chances of finding a job rise. This result could be explained either by the fact that people do not need to exert a higher level of search effort to find a job precisely because they have greater chances of employment,[68] or by the fact that in the most populated areas search cost increases offset the higher chances of employment. Indeed, the large commuting costs due to congestion (travelling on crowded public transportation, spending time in traffic, etc.) may discourage people from searching even though they have a higher probability of finding a job.

### 5.2.3 Robustness checks

Previous section's results would be unbiased and consistent if there was no omitted variable correlated with agglomeration. In particular, if the most agglomerated areas were – to a greater extent than the others – composed of selected job seekers (i.e., the most easily employable because of characteristics observable by the employer but not by the analyst), previous section's agglomeration estimates would be upward biased. Even though this is unlikely, given Italy's low labor mobility (see Section 4.3), there is always some small chance that the positive effect of agglomeration on hazard rates could in fact be due to LLMs' job-seeker-composition. To test whether there is a bias it would be necessary to run separate regressions on the "mover" and on "stayer" sub-samples, and test whether the two results are statistically different.[69] Thus, if the agglomeration coefficients were significantly positive (negative) only in the regressions on the movers from a less (more) into a more (less) agglomerated area and non-significant in the regressions on stayers, it would not be correct to conclude that agglomeration generates positive externalities in the matching process. Clearly, the sign of the correlation between a possible omitted variable and agglomeration could in principle be negative too. However, in this case Section 5.2.2's results would be reinforced rather than weakened, because the most agglomerated areas would exhibit higher employment chances, while being composed of a greater–than–average share of worst-quality job seekers.

The main issue is then the identification of the movers and the stayers, which can be done in two ways. Firstly, by considering as movers all the people who change municipality of residence from time $t$ to time $t+1$. Unfortunately, my data is too high a frequency and too short a time span to register these movements: none of the individuals interviewed by the LFS in 2002 changed municipality between two following quarters.[70] Secondly, by considering as movers all the people

[68] Although in the model presented in Section 2.1 the causality runs only from search intensity to hazard rates (and not viceversa).

[69] Under the assumption that ability at birth is randomly distributed across LLMs.

[70] The LFS would also allow to identify the individuals whom, at the time of the interview, were living in a munic-

who were born in a municipality different from that of residence. Regrettably, the LFS does not provide information on where the individual was born, which makes the identification of movers rather difficult. The only exception is for foreigners, for whom it is known the number of years of residence in Italy but not the country of origin (this information exists but is not available).

In this section I am going to consider as movers all the foreigners who have been living in Italy for more than five years and as stayers all the Italian residents. This restriction is due to the fact that during the first years of permanence immigrants might suffer from severe segregation (e.g., in terms of language barriers, housing, cultural differences, etc.), making their search behavior and employment chances rather peculiar.[71] Thus, I assume that after five years of residence immigrants face the same constraints and job opportunities as Italians, and that the latter's movements across LLMs are negligible (see Section 4.3 for evidence in support to this hypothesis).

The sample of movers (stayers) is composed of 297 $(70,711)$ non-employed individuals, 70 $(11,167)$ of whom job seekers. Since the number of movers is not large enough, I cannot differentiate between men and women. Results, reported in Table 11, are similar to those of Table 9 for the group of stayers (columns (11.7)-(11.12)), while show no evidence of agglomeration externalities for the mover sub-sample. Indeed, the immigrants living in the most agglomerated areas do not search differently from those living elsewhere, nor have significantly higher chances of finding employment per unit of search (columns (11.1)-(11.6)). Thus, the available data provides no evidence in favor of the hypothesis according to which the individuals (i.e., foreign immigrants) who move into the most agglomerated areas have non-observable characteristics that make them more easily to employ than those who live there.

# 6 Conclusions

In this paper I analyze agglomeration effects on both individual search intensity and hazard rates from non-employment (rather than from unemployment) into employment for Italian men and women. More specifically, I empirically examine whether population size and small-sized manufac-

ipality different from that of residence. Unfortunately, the available data does not disclose which is the municipality of presence (nor its size), nor whether the reason of the individual's absence is or is not related to work (e.g., it could as well be for holidays, health reasons, military service, etc.).

[71] I also run Table 9's specifications on the sub-sample of the 170 recent immigrants (i.e., the people who have been resident for less than five years). The number of job seekers (35) is too small to estimate hazard rates, so I can only estimate search intensity. Results (available upon request) show that the behavior of new immigrants does indeed differ from that of the rest of the sample, since they search less intensively than elsewhere in large cities, super-districts and, progressively, as the size of LLMs grows. This finding would be in line with the *spatial mismatch hypothesis* if immigrants were segregated in areas far away from the places of work. Unfortunately, I cannot directly test whether this is the case because the LFS does not provide information on intra-urban job and home location.

turing firm concentration generate overall net positive or negative externalities.

From the descriptive statistics, it would seem that hazard rates are highest (lowest) in the most industrially agglomerated areas (in large cities), and search intensity in the most highly populated LLMs. However, after controlling for individuals' observable characteristics, I find that only the matching process is affected by agglomeration. As to search intensity, on average it is not affected by either urbanization nor industrial agglomeration. A possible explanation of why the intensity of search does not increase in spite of higher hazard rates is that job seekers are discouraged from bearing the higher commuting costs produced by the presence of a large population mass (i.e., travelling on congested public transportation, spending time in traffic, etc.). Moreover, hazard rates increase not only with industry localization, but also with urbanization.

While these findings hold on average, it is interesting to analyze whether they occur at any level of agglomeration or only above certain threshold values. In this paper I show that results are sensitive to both the type and the degree of agglomeration of the local labor market. In particular, industry localization creates positive net economies mainly in super-districts, that is, in the subset of industrial clusters with the highest concentration of small and medium firms in the manufacturing sector; for "regular" districts the effect is less significant. Furthermore, job seekers' employment chances raise with the degree of urbanization, but only below to the $2,400,000$-inhabitant threshold. It would be useful, in terms of policy recommendations, to know whether this is due to an excessive congestion in Rome, Milan and Naples, or to the choosiness of these cities' job seekers.

Finally, while agglomeration effects are usually studied either at the urban or at the industry level, I am able, by using an Istat algorithm that identifies the more industrialized LLMs, to compare the magnitude of urbanization and localization effects on job seekers' probability of finding a job. Surprisingly, I find that the relative importance of the two effects depends on gender, as the urbanization (localization) differential in hazard rates is larger for women (men) than for men (women). While it is well known that labor markets dynamics are gender-specific, it is less obvious this is also the case for agglomeration externalities (even though this result is not new in the literature: see, for instance, Rosenthal and Strange, 2002). A possible explanation can be found in a difference in men's and women's preferences with respect to sectors (e.g., men might prefer industry, women services) and/or jobs (e.g., production versus administration) that are differently affected by agglomeration. Segregation might help explaining why women do not "prefer" applying for vacancies in the industrial district production process (even though it would be more difficult to explain why men should be segregated in large cities). Only in case of segregation would affirmative action policies be effective (see Flabbi, 2001).

Table 1: **Agglomeration effects on labor market dynamics**

| | | | | |
|---|---|---|---|---|
| Agglomeration factors increasing individual search intensity: | | | | |
| ↓ distance to job interviews | transportation costs | ↓ | search costs | ↓ |
| ↑ face-to-face contacts | job information-gathering costs | ↓ | search costs | ↓ |
| ↑ formal and informal networks | information on vacancies | ↑ | search costs | ↓ |
| ↑ congestion | house prices and rents | ↑ | cost of being U | ↑ |
| productivity gains | wages | ↑ | hazard rates | ↑ |
| ↑ number of vacancies | labor market tightness | ↑ | hazard rates | ↑ |
| ↑ formal and informal networks<br>↑ number of job seekers<br>productivity gains | job advertising | ↑ | hazard rates | ↑ |
| ↑ concentration of matching agents | chances of matching | ↑ | hazard rates | ↑ |
| labor pooling | quality / efficiency of matching | ↑ | hazard rates | ↑ |
| | | | | |
| Agglomeration factors lowering individual search intensity: | | | | |
| ↑ expectations on wages and hazards | reservation wages, choosiness | ↑ | hazard rates | ↓ |
| ↑ number of job seekers | labor market tightness | ↓ | hazard rates | ↓ |
| ↑ labor market tightness | job advertising | ↓ | hazard rates | ↓ |
| ↑ congestion > thick market externalities | chances of matching | ↓ | hazard rates | ↓ |
| ↑ congestion | job information-gathering costs | ↓ | search costs | ↑ |

Note: U = unemployed.

Table 2: **Mobility attitudes**

| Acceptable job location by those unemployed | | | | |
|---|---|---|---|---|
| Own municipality | Daily commuting distance | Anywhere in Italy | Anywhere | |
| 37.0 | 41.6 | 16.5 | 4.9 | |

| Job location of those employed | | | | |
|---|---|---|---|---|
| Own municipality | Other municipality in same province | No fixed place | Other province or abroad | |
| 61.2 | 26.6 | 4.7 | 7.5 | |

| Presence in the household at the time of interview | | | | |
|---|---|---|---|---|
| Present | Absent for less 1 year | Absent for more 1 year and searching | Absent for more 1 year not searching | |
| 98.3 | 0.5 | 0.2 | 0.9 | |

Source: author's elaboration on LFS data.

Table 3: **Frequency of rent contracts by landlord type**

| Contract type: | No. | % |
|---|---|---|
| Rent-controlled | 595 | 35.0 |
| In derogation from rent-control law | 269 | 15.8 |
| Non-resident | 3 | 0.2 |
| Informal/friendship | 422 | 24.9 |
| Subsidized | 81 | 4.8 |
| Council housing | 277 | 16.3 |
| Other | 51 | 3.0 |
| Total | 1,698 | 100.0 |
| Source: elaboration on the Bank of Italy's Survey of Household Income and Wealth data. | | |

Table 4: **Sample statistics**

| | Whole sample | | | Large cities | | | Super-districts | | | Industrial Districts | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Obs. | Mean | Std. Dev. | Obs. | Mean | Std. Dev. | Obs. | Mean | Std. Dev. | Obs. | Mean | Std. Dev. |
| Individual characteristics | | | | | | | | | | | | |
| Age | 71247 | 40.04 | 16.72 | 19849 | 40.49 | 16.66 | 5182 | 43.08 | 17.23 | 12620 | 43.13 | 17.12 |
| University degree or higher | 71247 | 0.03 | 0.18 | 19849 | 0.04 | 0.20 | 5182 | 0.03 | 0.16 | 12620 | 0.03 | 0.17 |
| High school | 71247 | 0.28 | 0.45 | 19849 | 0.29 | 0.45 | 5182 | 0.23 | 0.42 | 12620 | 0.24 | 0.43 |
| Compulsory education | 71247 | 0.41 | 0.49 | 19849 | 0.40 | 0.49 | 5182 | 0.38 | 0.49 | 12620 | 0.38 | 0.49 |
| North-East | 71247 | 0.15 | 0.36 | 19849 | 0.09 | 0.29 | 5182 | 0.35 | 0.48 | 12620 | 0.34 | 0.47 |
| Center | 71247 | 0.17 | 0.37 | 19849 | 0.18 | 0.39 | 5182 | 0.33 | 0.47 | 12620 | 0.22 | 0.41 |
| South | 71247 | 0.48 | 0.50 | 19849 | 0.45 | 0.50 | 5182 | 0.04 | 0.19 | 12620 | 0.07 | 0.26 |
| Past work experiences | 71247 | 0.44 | 0.50 | 19849 | 0.40 | 0.49 | 5182 | 0.62 | 0.49 | 12620 | 0.61 | 0.49 |
| Search duration: < 1 month | 71247 | 0.85 | 0.36 | 19849 | 0.84 | 0.36 | 5182 | 0.93 | 0.26 | 12620 | 0.93 | 0.26 |
| Search duration: 1-5 months | 71247 | 0.02 | 0.15 | 19849 | 0.02 | 0.14 | 5182 | 0.02 | 0.15 | 12620 | 0.02 | 0.15 |
| Search duration: 6-11 months | 71247 | 0.02 | 0.13 | 19849 | 0.01 | 0.12 | 5182 | 0.01 | 0.12 | 12620 | 0.01 | 0.11 |
| Employed family members | 71247 | 0.97 | 0.83 | 19849 | 0.91 | 0.81 | 5182 | 1.16 | 0.93 | 12620 | 1.12 | 0.90 |
| Non-working househ. members | 71247 | 2.50 | 1.08 | 19849 | 2.50 | 1.08 | 5182 | 2.22 | 0.95 | 12620 | 2.22 | 0.96 |
| Single living alone | 71247 | 0.03 | 0.17 | 19849 | 0.04 | 0.18 | 5182 | 0.03 | 0.17 | 12620 | 0.03 | 0.18 |
| Household head | 71247 | 0.20 | 0.40 | 19849 | 0.20 | 0.40 | 5182 | 0.22 | 0.42 | 12620 | 0.23 | 0.42 |
| Spouse | 71247 | 0.38 | 0.49 | 19849 | 0.37 | 0.48 | 5182 | 0.40 | 0.49 | 12620 | 0.40 | 0.49 |
| Student | 71247 | 0.35 | 0.48 | 19849 | 0.36 | 0.48 | 5182 | 0.34 | 0.47 | 12620 | 0.34 | 0.47 |
| Housewife | 71247 | 0.25 | 0.43 | 19849 | 0.23 | 0.42 | 5182 | 0.25 | 0.43 | 12620 | 0.24 | 0.43 |
| Other inactive condition | 71247 | 0.24 | 0.43 | 19849 | 0.24 | 0.43 | 5182 | 0.34 | 0.47 | 12620 | 0.34 | 0.47 |
| LLM characteristics | | | | | | | | | | | | |
| LLM's population | 518 | 1.03 | 2.52 | 19 | 10.18 | 8.78 | 67 | 0.73 | 0.85 | 145 | 0.92 | 1.01 |
| LLM's job seekers (log) | 518 | 7.29 | 1.25 | 19 | 10.09 | 1.02 | 67 | 6.74 | 0.94 | 145 | 6.97 | 0.95 |
| LLM's area (log) | 518 | 5.86 | 0.76 | 19 | 7.03 | 0.66 | 67 | 5.65 | 0.66 | 145 | 5.77 | 0.60 |
| LLM's average extra hours | 518 | 0.04 | 0.05 | 19 | 0.05 | 0.02 | 67 | 0.05 | 0.06 | 145 | 0.05 | 0.06 |
| LLM's share of overtime work/workers | 518 | 35.63 | 5.38 | 19 | 35.73 | 1.43 | 67 | 36.86 | 3.78 | 145 | 36.70 | 4.15 |

Source: author's elaboration on LFS data.

Table 5: **Average Transition Probabilities**

| Quarterly transition probabilities | | | | |
|---|---|---|---|---|
| | Employed$_{t+1}$ | Unemployed$_{t+1}$ | Inactive$_{t+1}$ | Total |
| | | Men and Women | | |
| Employed$_t$ | 96.6 | 0.9 | 2.5 | 100.0 |
| Unemployed$_t$ | 13.3 | 63.8 | 22.9 | 100.0 |
| Inactive$_t$ | 3.6 | 3.6 | 92.8 | 100.0 |
| Population composition$_{t+1}$ | 54.7 | 5.6 | 39.7 | 100.0 |
| | | Men | | |
| Employed$_t$ | 97.3 | 0.8 | 1.9 | 100.0 |
| Unemployed$_t$ | 16.1 | 65.8 | 18.1 | 100.0 |
| Inactive$_t$ | 5.1 | 4.4 | 90.5 | 100.0 |
| Population composition$_{t+1}$ | 68.3 | 5.2 | 26.5 | 100.0 |
| | | Women | | |
| Employed$_t$ | 95.5 | 1.1 | 3.4 | 100.0 |
| Unemployed$_t$ | 10.9 | 62.2 | 26.9 | 100.0 |
| Inactive$_t$ | 2.8 | 3.2 | 94.0 | 100.0 |
| Population composition$_{t+1}$ | 41.4 | 5.9 | 52.7 | 100.0 |

| Quarterly transition flows | | | | |
|---|---|---|---|---|
| | Employed$_{t+1}$ | Unemployed$_{t+1}$ | Inactive$_{t+1}$ | Population composition$_t$ |
| | | Men and Women | | |
| Employed$_t$ | 52.5 | 0.5 | 1.3 | 54.2 |
| Unemployed$_t$ | 0.8 | 3.6 | 1.3 | 5.6 |
| Inactive$_t$ | 1.4 | 1.4 | 37.1 | 40.2 |
| Population composition$_{t+1}$ | 54.7 | 5.6 | 39.7 | 100.0 |
| | | Men | | |
| Employed$_t$ | 66.1 | 0.6 | 1.3 | 67.7 |
| Unemployed$_t$ | 0.8 | 3.4 | 0.9 | 5.2 |
| Inactive$_t$ | 1.4 | 1.2 | 24.3 | 27.1 |
| Population composition$_{t+1}$ | 68.3 | 5.2 | 26.5 | 100.0 |
| | | Women | | |
| Employed$_t$ | 39.2 | 0.4 | 1.4 | 40.9 |
| Unemployed$_t$ | 0.7 | 3.8 | 1.6 | 6.1 |
| Inactive$_t$ | 1.5 | 1.7 | 49.7 | 53.1 |
| Population composition$_{t+1}$ | 41.4 | 6.0 | 52.7 | 100.0 |

Source: elaboration on LFS data. Note: flows are expressed in percentage of the working age population.

Table 6: **Labor market characteristics and agglomeration**

| | Employment rate | Unemployment rate | Job seekers | Activity rate | Hazard into employment |
|---|---|---|---|---|---|
| | | | Italy | | |
| Large city | 44.1 | 10.1 | 16.1 | 48.7 | 21.6 |
| Large city and super-district | 52.6 | 3.6 | 7.3 | 54.6 | 38.7 |
| Small town and super-district | 48.9 | 4.2 | 9.7 | 51.0 | 52.5 |
| Small town - other | 41.2 | 11.8 | 16.4 | 46.4 | 29.7 |
| Industrial district | 48.5 | 4.6 | 9.5 | 50.8 | 45.9 |
| | | | North-West | | |
| Large city | 49.3 | 4.8 | 10.4 | 51.8 | 32.8 |
| Large city and super-district | 52.6 | 3.6 | 7.3 | 54.6 | 38.7 |
| Small town and super-district | 50.1 | 3.8 | 8.8 | 52.1 | 50.6 |
| Small town - other | 48.1 | 4.3 | 9.8 | 50.2 | 41.4 |
| Industrial district | 49.7 | 3.9 | 8.7 | 51.7 | 44.4 |
| | | | North-East | | |
| Large city | 50.1 | 3.4 | 7.4 | 51.9 | 47.4 |
| Large city and super-district | – | – | – | – | – |
| Small town and super-district | 50.9 | 3.2 | 8.4 | 52.6 | 59.1 |
| Small town - other | 51.2 | 3.2 | 9.9 | 52.9 | 58.7 |
| Industrial district | 51.1 | 3.3 | 8.4 | 52.8 | 54.3 |
| | | | Center | | |
| Large city | 46.0 | 6.4 | 13.8 | 49.2 | 22.5 |
| Large city and super-district | – | – | – | – | – |
| Small town and super-district | 47.4 | 4.5 | 11.3 | 49.6 | 49.2 |
| Small town - other | 44.4 | 6.0 | 12.5 | 47.2 | 33.1 |
| Industrial district | 46.8 | 4.7 | 11.3 | 49.2 | 42.9 |
| | | | South | | |
| Large city | 36.3 | 18.7 | 22.1 | 44.7 | 16.5 |
| Large city and super-district | – | – | – | – | – |
| Small town and super-district | 41.4 | 10.1 | 10.7 | 46.1 | 42.9 |
| Small town - other | 35.2 | 18.3 | 21.5 | 43.0 | 23.6 |
| Industrial district | 39.0 | 12.6 | 12.4 | 44.5 | 32.5 |

Source: elaboration on LFS data. Note that the only LLM that is both a large city and a super-district is that of Desio.

Table 7: **Baseline models for men**

| | Hazard to employment | | | | Search intensity | | | |
|---|---|---|---|---|---|---|---|---|
| | Probit | | Heckprobit | | Probit | | Heckprobit | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's job seekers (log) | -0.102 | 0.000 | -0.089 | 0.000 | 0.019 | 0.178 | 0.018 | 0.190 |
| LLM's area (log) | 0.030 | 0.331 | 0.013 | 0.668 | -0.041 | 0.112 | -0.041 | 0.113 |
| LLM's average extra hours worked | 0.001 | 0.999 | -0.180 | 0.791 | -0.540 | 0.348 | -0.544 | 0.338 |
| LLM's share of overtime workers in total workers | -0.009 | 0.196 | -0.010 | 0.133 | -0.006 | 0.180 | -0.007 | 0.171 |
| Quarter I (seasonal dummy) | 0.060 | 0.229 | 0.059 | 0.207 | 0.010 | 0.735 | 0.011 | 0.718 |
| Quarter II (seasonal dummy) | 0.074 | 0.151 | 0.072 | 0.154 | 0.068 | 0.013 | 0.068 | 0.014 |
| North-East | 0.157 | 0.115 | 0.146 | 0.137 | 0.116 | 0.047 | 0.115 | 0.048 |
| Center | -0.033 | 0.694 | -0.034 | 0.677 | -0.013 | 0.802 | -0.016 | 0.745 |
| South | -0.177 | 0.017 | -0.135 | 0.063 | 0.043 | 0.389 | 0.045 | 0.363 |
| Age | 0.016 | 0.124 | 0.053 | 0.000 | 0.090 | 0.000 | 0.091 | 0.000 |
| Age squared | 0.000 | 0.371 | -0.001 | 0.000 | -0.001 | 0.000 | -0.001 | 0.000 |
| University degree or higher | -0.165 | 0.085 | -0.124 | 0.163 | 0.237 | 0.001 | 0.238 | 0.001 |
| High school | -0.151 | 0.012 | -0.168 | 0.004 | 0.041 | 0.316 | 0.041 | 0.320 |
| Middle school | -0.118 | 0.045 | -0.139 | 0.015 | -0.013 | 0.737 | -0.009 | 0.802 |
| Past work experiences | 0.205 | 0.001 | 0.263 | 0.000 | 0.094 | 0.064 | 0.101 | 0.043 |
| Search duration: < 1 month | 1.378 | 0.000 | 0.804 | 0.000 | -1.022 | 0.000 | -1.015 | 0.000 |
| Search duration: 1-5 months | 0.546 | 0.000 | 0.549 | 0.000 | 0.133 | 0.039 | 0.135 | 0.036 |
| Search duration: 6-11 months | 0.310 | 0.000 | 0.293 | 0.000 | -0.047 | 0.502 | -0.046 | 0.515 |
| Employed family members | 0.048 | 0.044 | 0.053 | 0.020 | 0.040 | 0.014 | 0.039 | 0.017 |
| Single living alone | | | | | 0.116 | 0.071 | 0.083 | 0.189 |
| Household head | | | | | 0.121 | 0.037 | 0.086 | 0.137 |
| Spouse | | | | | 0.400 | 0.001 | 0.361 | 0.002 |
| Student | | | | | -0.149 | 0.133 | -0.253 | 0.008 |
| Housewife | | | | | -1.136 | 0.000 | -1.128 | 0.000 |
| Other inactive condition | | | | | -1.348 | 0.000 | -1.369 | 0.000 |
| Number of non-working household members | | | | | 0.021 | 0.127 | 0.023 | 0.090 |
| Constant | -0.408 | 0.260 | -1.134 | 0.002 | -0.447 | 0.103 | -0.471 | 0.082 |
| Number of observations: | 25,116 | | 25,116 | | | | | |
| of which uncensored: | | | 5,545 | | | | | |

Source: author's elaboration on LFS data. Note: White-robust standard errors adjusted for clustering.

Table 8: **Baseline models for women**

| | Hazard to employment | | | | Search intensity | | | |
|---|---|---|---|---|---|---|---|---|
| | Probit | | Heckprobit | | Probit | | Heckprobit | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's job seekers (log) | -0.070 | 0.001 | -0.069 | 0.001 | -0.013 | 0.286 | -0.014 | 0.280 |
| LLM's area (log) | 0.043 | 0.287 | 0.041 | 0.301 | 0.007 | 0.746 | 0.007 | 0.765 |
| LLM's average extra hours worked | -0.458 | 0.490 | -0.461 | 0.474 | 0.041 | 0.931 | 0.044 | 0.925 |
| LLM's share of overtime workers in total workers | 0.007 | 0.281 | 0.006 | 0.350 | -0.005 | 0.166 | -0.005 | 0.160 |
| Quarter I (seasonal dummy) | 0.030 | 0.531 | 0.033 | 0.479 | 0.038 | 0.193 | 0.039 | 0.178 |
| Quarter II (seasonal dummy) | 0.066 | 0.206 | 0.056 | 0.267 | 0.000 | 0.988 | 0.000 | 0.986 |
| North-East | 0.192 | 0.031 | 0.191 | 0.030 | 0.095 | 0.057 | 0.096 | 0.057 |
| Center | -0.074 | 0.299 | -0.070 | 0.317 | -0.035 | 0.404 | -0.035 | 0.402 |
| South | -0.276 | 0.000 | -0.254 | 0.000 | 0.003 | 0.942 | 0.004 | 0.934 |
| Age | -0.035 | 0.000 | -0.025 | 0.018 | 0.063 | 0.000 | 0.062 | 0.000 |
| Age squared | 0.001 | 0.000 | 0.000 | 0.010 | -0.001 | 0.000 | -0.001 | 0.000 |
| University degree or higher | 0.123 | 0.204 | 0.188 | 0.049 | 0.191 | 0.000 | 0.193 | 0.000 |
| High school | -0.037 | 0.623 | -0.020 | 0.790 | 0.079 | 0.014 | 0.080 | 0.013 |
| Middle school | -0.090 | 0.223 | -0.092 | 0.200 | 0.010 | 0.745 | 0.009 | 0.768 |
| Past work experiences | 0.233 | 0.000 | 0.294 | 0.000 | 0.129 | 0.000 | 0.133 | 0.000 |
| Search duration: $<$ 1 month | 1.346 | 0.000 | 0.954 | 0.000 | -1.054 | 0.000 | -1.057 | 0.000 |
| Search duration: 1-5 months | 0.603 | 0.000 | 0.609 | 0.000 | 0.152 | 0.012 | 0.152 | 0.013 |
| Search duration: 6-11 months | 0.511 | 0.000 | 0.508 | 0.000 | 0.072 | 0.220 | 0.072 | 0.224 |
| Employed family members | 0.045 | 0.120 | 0.036 | 0.216 | -0.005 | 0.752 | -0.006 | 0.694 |
| Single living alone | | | | | -0.094 | 0.168 | -0.099 | 0.148 |
| Household head | | | | | -0.153 | 0.002 | -0.147 | 0.003 |
| Spouse | | | | | -0.302 | 0.000 | -0.299 | 0.000 |
| Student | | | | | -1.046 | 0.000 | -1.041 | 0.000 |
| Housewife | | | | | -1.269 | 0.000 | -1.266 | 0.000 |
| Other inactive condition | | | | | -0.975 | 0.000 | -0.994 | 0.000 |
| Number of non-working household members | | | | | 0.017 | 0.110 | 0.017 | 0.115 |
| Constant | -0.719 | 0.041 | -0.946 | 0.007 | -0.033 | 0.879 | -0.009 | 0.966 |
| Number of observations: | 46,131 | | 46,131 | | | | | |
| of which uncensored: | | | 5,731 | | | | | |

Source: author's elaboration on LFS data. Note: White-robust standard errors adjusted for clustering.

Table 9: **Marginal effects on hazard rates and search intensity (probit model)**

| | Hazards to employment: men | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (9.1) | | (9.2) | | (9.3) | | (9.4)(*) | | (9.5)(*) | | (9.6)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.001 | 0.318 | 0.001 | 0.282 | | | 0.008 | 0.057 | 0.009 | 0.038 |
| Large city dummy | 0.059 | 0.057 | | | | | 0.063 | 0.065 | | | | |
| Super-district dummy | 0.077 | 0.036 | 0.081 | 0.035 | | | 0.084 | 0.030 | 0.086 | 0.032 | | |
| Industrial district dummy | | | | | 0.043 | 0.107 | | | | | 0.051 | 0.075 |
| | Hazards to employment: women | | | | | | | | | | | |
| | (9.7) | | (9.8) | | (9.9) | | (9.10)(*) | | (9.11)(*) | | (9.12)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.002 | 0.189 | 0.002 | 0.226 | | | 0.010 | 0.002 | 0.010 | 0.002 |
| Large city dummy | 0.077 | 0.009 | | | | | 0.075 | 0.016 | | | | |
| Super-district dummy | 0.049 | 0.080 | 0.050 | 0.080 | | | 0.047 | 0.096 | 0.047 | 0.098 | | |
| Industrial district dummy | | | | | 0.006 | 0.789 | | | | | 0.007 | 0.737 |
| | Search intensity: men | | | | | | | | | | | |
| | (9.13) | | (9.14) | | (9.15) | | (9.16)(*) | | (9.17)(*) | | (9.18)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.001 | 0.141 | 0.001 | 0.181 | | | -0.002 | 0.272 | -0.002 | 0.240 |
| Large city dummy | -0.004 | 0.744 | | | | | -0.009 | 0.494 | | | | |
| Super-district dummy | 0.005 | 0.600 | 0.006 | 0.541 | | | 0.005 | 0.636 | 0.005 | 0.643 | | |
| Industrial district dummy | | | | | -0.007 | 0.424 | | | | | -0.009 | 0.255 |
| | Search intensity: women | | | | | | | | | | | |
| | (9.19) | | (9.20) | | (9.21) | | (9.22)(*) | | (9.23)(*) | | (9.24)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.000 | 0.134 | 0.000 | 0.196 | | | 0.000 | 0.985 | 0.000 | 0.917 |
| Large city dummy | -0.001 | 0.885 | | | | | -0.003 | 0.656 | | | | |
| Super-district dummy | 0.002 | 0.827 | 0.002 | 0.777 | | | 0.001 | 0.867 | 0.001 | 0.866 | | |
| Industrial district dummy | | | | | -0.006 | 0.239 | | | | | -0.007 | 0.155 |

Source: author's elaboration on LFS data. Note: White-robust standard errors adjusted for clustering.

(*) Computed on the sub-sample excluding the three largest LLMs (i.e., Rome, Milan, and Naples).

Table 10: **Hazard rates and search intensity (bivariate probit with sample selection)**

| | Hazards to employment: men | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (10.1) | | (10.2) | | (10.3) | | (10.4)(*) | | (10.5)(*) | | (10.6)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.046 | 0.258 | 0.048 | 0.235 | | | 0.181 | 0.135 | 0.193 | 0.105 |
| Large city dummy | 0.152 | 0.093 | | | | | 0.149 | 0.114 | | | | |
| Super-district dummy | 0.202 | 0.044 | 0.212 | 0.041 | | | 0.211 | 0.037 | 0.216 | 0.038 | | |
| Industrial district dummy | | | | | 0.104 | 0.168 | | | | | 0.116 | 0.129 |
| | Hazards to employment: women | | | | | | | | | | | |
| | (10.7) | | (10.8) | | (10.9) | | (10.10)(*) | | (10.11)(*) | | (10.12)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.072 | 0.161 | 0.068 | 0.198 | | | 0.308 | 0.002 | 0.308 | 0.002 |
| Large city dummy | 0.238 | 0.010 | | | | | 0.220 | 0.019 | | | | |
| Super-district dummy | 0.146 | 0.107 | 0.149 | 0.105 | | | 0.135 | 0.132 | 0.134 | 0.135 | | |
| Industrial district dummy | | | | | 0.010 | 0.889 | | | | | 0.011 | 0.875 |
| | Search intensity: men | | | | | | | | | | | |
| | (10.13) | | (10.14) | | (10.15) | | (10.16)(*) | | (10.17)(*) | | (10.18)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.042 | 0.153 | 0.039 | 0.193 | | | -0.095 | 0.279 | -0.100 | 0.246 |
| Large city dummy | -0.019 | 0.775 | | | | | -0.041 | 0.527 | | | | |
| Super-district dummy | 0.022 | 0.651 | 0.027 | 0.584 | | | 0.019 | 0.683 | 0.019 | 0.686 | | |
| Industrial district dummy | | | | | -0.034 | 0.416 | | | | | -0.047 | 0.255 |
| | Search intensity: women | | | | | | | | | | | |
| | (10.19) | | (10.20) | | (10.21) | | (10.22)(*) | | (10.23)(*) | | (10.24)(*) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.036 | 0.140 | 0.032 | 0.204 | | | -0.001 | 0.994 | -0.009 | 0.927 |
| Large city dummy | -0.010 | 0.861 | | | | | -0.027 | 0.629 | | | | |
| Super-district dummy | 0.011 | 0.848 | 0.015 | 0.797 | | | 0.008 | 0.895 | 0.008 | 0.892 | | |
| Industrial district dummy | | | | | -0.047 | 0.235 | | | | | -0.055 | 0.151 |

Source: author's elaboration on LFS data. Note: White-robust standard errors adjusted for clustering.

(*) Computed on the sub-sample excluding the three largest LLMs (i.e., Rome, Milan, and Naples).

Table 11: **Robustness checks (probit model, marginal effects)**

| | | | | | | |
|---|---|---|---|---|---|---|
| Hazards to employment: foreigners resident for more than 5 years | | | | | | |
| | (11.1) | | (11.2) | | (11.3) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | -0.002 | 0.412 | -0.004 | 0.112 |
| Large city dummy | 0.028 | 0.731 | | | | |
| Super-district dummy | 0.168 | 0.316 | 0.099 | 0.467 | | |
| Industrial district dummy | | | | | -0.036 | 0.106 |
| Search intensity: foreigners resident for more than 5 years | | | | | | |
| | (11.4) | | (11.5) | | (11.6) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.004 | 0.346 | 0.005 | 0.315 |
| Large city dummy | 0.005 | 0.955 | | | | |
| Super-district dummy | -0.047 | 0.506 | -0.033 | 0.680 | | |
| Industrial district dummy | | | | | -0.006 | 0.909 |
| Hazards to employment: stayers | | | | | | |
| | (11.7) | | (11.8) | | (11.9) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.002 | 0.138 | 0.002 | 0.141 |
| Large city dummy | 0.062 | 0.011 | | | | |
| Super-district dummy | 0.056 | 0.022 | 0.059 | 0.020 | | |
| Industrial district dummy | | | | | 0.024 | 0.191 |
| Search intensity: stayers | | | | | | |
| | (11.10) | | (11.11) | | (11.12) | |
| | Coeff. | P-val. | Coeff. | P-val. | Coeff. | P-val. |
| LLM's population | | | 0.001 | 0.068 | 0.001 | 0.112 |
| Large city dummy | -0.001 | 0.879 | | | | |
| Super-district dummy | 0.003 | 0.670 | 0.004 | 0.607 | | |
| Industrial district dummy | | | | | -0.007 | 0.166 |

Source: author's elaboration on LFS data. Note: White-robust standard errors adjusted for clustering.

## Appendix 1: Attrition analysis

I reconstructed the LFS longitudinal data with a deterministic method. The loss of observations implied by this method can be due to reporting errors in the household identifier or in the other individual variables (typically, the date of birth), but it can be also due to genuine "attrition": this is the loss of information deriving from the non-availability of some of the people to be re-interviewed at time $t+1$. In what follows I use the term "attrition" for both types of losses.

If the information loss was correlated to working condition changes, attrition would be a potential source of bias for the estimation of labor market dynamics. This typically occurs when people change residence because they find employment in a different location, in which case the exit from the LFS sample is determined by a movement towards employment.

In order to test for the effects of attrition in the estimation of labor market dynamics, I follow the approach proposed by Jiménez-Martín and Peracchi (2003), looking at individuals' survey participation at time $t$, $t+1$ and $t+4$ (i.e., respectively, one quarter and one year after the first LFS interview). As Jiménez-Martín and Peracchi (2003), I identify two sets of individuals: (1) those participating at all the three surveys (full-time respondents); and (2) those participating at time $t$ and $t+1$ but not at time $t+4$ (non full-time respondents). More formally, let $D$ be an indicator equal to 1 if the person is a full-time respondent and to 0 elsewhere, and consider a standard three-state labor market. Let $\pi_{ij}^{D}$ be the probability of moving from state $i = U, O$ at time $t$ to state $j = E, U, O$[72] at time $t+1$, for an individual whose sample participation is denoted by $D = 0, 1$. Attrition may bias transition probabilities if:

$$\pi_{ij}^{0} \neq \pi_{ij}^{1} \tag{11}$$

for $i = U, O$, $j = E, U, O$.

Consider the statistic $l_{ij} = \pi_{ij}^{0} - \pi_{ij}^{1}$. If attrition was not a source of bias for transition probabilities, under the null hypothesis $l_{ij}$ would be equal to zero. In other words, if full time respondents and people who are subject to attrition have the same probability to move towards all the other labor market states then I can assume that attrition does not affect transition probabilities.

Critical values for $l_{ij}$ can be easily derived. Because of the central limit theorem, $l_{ij}$ divided by its standard error has a $t$-Student's distribution. Rejection at 95 percent significance level, for instance, occurs for values of $l_{ij}$ greater than 2 in absolute value. Table 6 reports the test statistics by gender, age group (15-34 and 35+) and area of residence (North–West, North–East,

[72] E=Employed, U=Unemployed, O=Out of the labor force.

Center, South). As the table shows, the test results confirm the adequacy of the adopted matching procedure in my study of labor market movements, for all the socio-demographic groups considered.

Table A1. Testing for the effect of attrition

| | Men | | Women | |
|---|---|---|---|---|
| | Age 15–34 | Age 35–64 | Age 15–34 | Age 35–64 |
| | North West | | | |
| $l_{UE}$ | 0.33 | -0.10 | 0.35 | 0.13 |
| $l_{UU}$ | -0.44 | -0.25 | 0.31 | -0.27 |
| $l_{UO}$ | 0.07 | -0.05 | -0.20 | -0.28 |
| $l_{OE}$ | 0.09 | 0.02 | -0.04 | 0.00 |
| $l_{OU}$ | -0.06 | 0.00 | -0.03 | 0.01 |
| $l_{OO}$ | 0.03 | 0.02 | -0.06 | 0.35 |
| | North East | | | |
| $l_{UE}$ | -0.39 | 0.21 | -0.91 | -0.04 |
| $l_{UU}$ | -0.46 | -1.05 | 0.02 | -0.43 |
| $l_{UO}$ | 0.57 | 0.23 | 0.03 | -0.21 |
| $l_{OE}$ | 0.15 | 0.05 | -0.07 | -0.02 |
| $l_{OU}$ | -0.02 | -0.01 | -0.10 | -0.04 |
| $l_{OO}$ | -1.31 | 0.12 | -0.95 | -0.19 |
| | Centre | | | |
| $l_{UE}$ | 0.07 | 0.01 | 0.13 | 0.00 |
| $l_{UU}$ | -0.10 | -0.54 | -0.11 | 0.03 |
| $l_{UO}$ | -0.10 | 0.65 | 0.12 | -0.44 |
| $l_{OE}$ | -0.01 | -0.05 | 0.02 | 0.04 |
| $l_{OU}$ | 0.03 | 0.01 | 0.01 | 0.02 |
| $l_{OO}$ | -0.73 | 0.34 | -1.00 | -0.66 |
| | South | | | |
| $l_{UE}$ | -0.06 | 0.22 | -0.09 | -0.06 |
| $l_{UU}$ | -0.91 | -2.03 | -1.14 | -0.34 |
| $l_{UO}$ | -0.18 | 0.00 | -0.11 | -0.24 |
| $l_{OE}$ | -0.10 | -0.01 | -0.01 | 0.02 |
| $l_{OU}$ | -0.29 | 0.01 | -0.20 | 0.00 |
| $l_{OO}$ | -0.80 | 0.08 | -1.55 | -1.38 |

Source: elaboration on LFS data.

## References


Andersson, Fredrik, Simon Burgess, and Julia Lane. 2004. "Cities, Matching and the Productivity Gains of Agglomeration," Discussion Paper No. 4598, CEPR.

Arzaghi, Mohammad, and J. Vernon Henderson. 2005. "Networking off Madison Square," Working Papers, US Census Bureau, Center for Economic Studies.

Au, Chun-Chung, and J. Vernon Henderson. 2006. "How Migration Restrictions Limit Agglomeration and Productivity in China," *Journal of Economic Development*.

Berliant, Marcus, Reed III, Robert R., and Ping Wang. 2006. "Knowledge Exchange, Matching, and Agglomeration," *Journal of Urban Economics*, 60, 69–95.

Black, Dan A. 1995. "Discrimination in an Equilibrium Search Model," *Journal of Labor Economics*, 13, 309–334.

Brandolini, Andrea, Piero Cipollone, and Eliana Viviano. 2004. "Does the ILO Definition Capture all Unemployment?," Temi di Discussione No. 529, Bank of Italy.

Broersma, L., and J.C. van Ours. 1999. "Job Searchers, Job Matches and the Elasticity of Matching," *Labor Economics*, 6, 77–93.

Burgess, Simon. 1993. "A Model of Competition between Unemployed and Employed Job-Searchers: An Application to the Unemployment Outflow Rate in Britain," *The Economic Journal*, 103, 1190–1204.

Cannari, Luigi, Francesco Nucci, and Paolo Sestito. 2000. "Geographic Labor Mobility and the Cost of Housing: Evidence from Italy," *Applied Economics*, 32, 1899–1906.

Cannari, Luigi, and Luigi F. Signorini. 2000. "Nuovi strumenti per la classificazione del sistemi industriali," in L.F. Signorini (ed.), *Lo sviluppo locale*, Roma: Donzelli, pp. 123–151.

Ciccone, Antonio. 2002. "Agglomeration Effects in Europe," *European Economic Review*, 46, 213–227.

Ciccone, Antonio, and Robert E. Hall. 1996. "Productivity and the Density of Economic Activity," *American Economic Review*, 86, 54–70.

Coles, M.G., and E. Smith. 1996. "Cross Section Estimation of a Matching Function: Evidence from England and Wales," *Economica*, 63, 589–598.

Coulson, N. Edward, Laing, Derek and Ping Wang. 2001. “Spatial Mismatch in Search Equilibrium,” *Journal of Labor Economics*, 19, 949–972.

de Blasio, Guido, and Sabrina Di Addario. 2005. ”Do Workers Benefit from Industrial Agglomeration?,” *Journal of Regional Science*, 45, 797–827.

Del Boca, Daniela. 2001. “L’offerta di lavoro,” in Brucchi Luchino (ed.), *Manuale di economia del lavoro*, Bologna: Il Mulino.

Devine, T.J., and N.M. Kiefer. 1991. *Empirical Economics*, Oxford University Press, New York and Oxford.

Di Addario, Sabrina. 2002. “Italian Household Tenure Choices and Housing Demand,” Lavoro preparatorio alla Relazione sul 2001, Bank of Italy.

Di Addario, Sabrina, and Eleonora Patacchini. 2006. “Is There an Urban Wage Premium in Italy?,” Temi di Discussione No. 570, Bank of Italy.

Dohmen, Thomas J. 2005. “Housing, Mobility and Unemployment,” *Regional Science and Urban Economics*, 35, 305–325.

Duranton, Gilles, and Diego Puga. 2004. “Micro-Foundations of Urban Agglomeration Economies,” in J.V. Henderson and J.F. Thisse (eds.), *Handbook of Urban and Regional Economics*, Amsterdam and New York: North Holland.

Duranton, Gilles, and Vassilis Monastiriotis. 2002. ”Mind the Gaps: The Evolution of Regional Earnings Inequalities in the U.K., 1982-1997,” *Journal of Regional Science*, 42, 219-256.

Flabbi, Luca. 2001. “La discriminazione: evidenza empirica e teoria economica,” in Brucchi Luchino (ed.), *Manuale di economia del lavoro*, Bologna: Il Mulino.

Gautier, Pieter A., and Coen N. Teulings. 2003a. “An Empirical Index for Labor Market Density,” *The Review of Economics and Statistics*, 85, 901–908.

Glaeser, Edward L., and David C. Mare’. 2001. “Cities and Skills,” *Journal of Labor Economics*, 19, 316–342.

Helsley, Robert W., and William C. Strange. 1990. “Matching and Agglomeration Economies in a System of Cities,” *Regional Science and Urban Economics*, 20, 189–212.

Henley, A. 1998. "Residential Mobility, Housing Equity and the Labor Market," *Economic Journal*, 108, 414–428.

Ioannides, Yannis M. and Linda D. Loury. 2004. "Job Information Networks, Neighborhood Effects, and Inequality," *Journal of Economic Literature*, 42, 1056–1093.

Istat. 1997. *I sistemi locali del lavoro 1991.* Rome, Istat.

Jiménez-Martín, S., and Franco Peracchi. 2003. "Sample Attrition and Labor Force Dynamics: Evidence from the Spanish Labor Force Survey," *Spanish Economic Review*, 4, 79–102.

Lancaster, T. 1979. "Econometric Methods for the Duration of Unemployment," *Econometrica*, 47, 939–956.

OECD. 2002. *Redefining Territories. The Functional Regions*, OECD: Paris.

Paggiaro, Adriano, and Nicola Torelli. 1999. "Una procedura per l'abbinamento di record nella rilevazione trimestrale delle Forze di Lavoro," presented at the Conference "Metodi e studi di valutazione degli effetti di politiche del lavoro, di aiuto alle imprese e di welfare", 2003.

Patacchini, Eleonora, and Yves Zenou. 2005. "Search Intensity, Cost of Living and Local Labor Markets," *Regional Science and Urban Economics*, 36, 227–248.

Peracchi, Franco, and Eliana Viviano. 2004."An Empirical Micro Matching Model with an Application to Italy and Spain," Temi di Discussione No. 538, Bank of Italy.

Petrongolo, Barbara. 2001. "Reemployment Probabilities and Returns to Matching," *Journal of Labor Economics*, 19, 716–741.

Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39, 390–431.

Petrongolo, Barbara, and Christopher A. Pissarides. 2006. "Scale Effects in Markets with Search," *The Economic Journal*, 116, 21–44.

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory*, Cambridge: The MIT Press.

Putnam, Robert D. 1993. *Making Democracy Work*, Princeton: Princeton University Press.

Rosenthal, Stuart S., and William C. Strange. 2002. "The Urban Rat Race," presented at the CEPR meeting "The Economics of Cities," London, 6-8 June 2003, mimeo.

Rosenthal, Stuart S., and William C. Strange. 2004. "Evidence on the Nature and Sources of Agglomeration Economies," in J.V. Henderson and J.F. Thisse (eds.), *Handbook of Urban and Regional Economics*, Amsterdam and New York: North Holland.

Smith, Tony E., and Yves Zenou. 2003. "Spatial Mismatch, Search Effort, and Urban Spatial Structure," *Journal of Urban Economics*, 54, 129–156.

van de Ven, Wynand, and Bernard van Praag. 1981. "The Demand for Deductibles in Private Health Insurance. A Probit Model with Sample Selection," *Journal of Econometrics*, 17, 229–252.

Venables, Anthony J. 2002. "Labour Sorting by Cities: Partnership, Self-selection, and Agglomeration," mimeo.

Wahba, J., and Yves Zenou. 2005. "Density, Social Networks and Job Search Methods: Theory and Application to Egypt," *Journal of Development Economics*, 78, 443–473.

Wasmer, Etienne, and Yves Zenou. 2002. "Does City Structure Affect Job Search and Welfare?," *Journal of Urban Economics*, 51, 515–541.

Wheeler, Christopher H. 2001. "Search, Sorting, and Urban Agglomeration," *Journal of Labor Economics*, 19, 879–899.