

Marshallian Labor Market Pooling: Evidence from Italy

Monica Andini*[†]

Bank of Italy

Guido de Blasio*[‡]

Bank of Italy

Gilles Duranton*[‡]

University of Toronto

William C. Strange*[§]

University of Toronto

First draft: 6 February 2012

ABSTRACT: This paper employs a unique Italian data source to take a comprehensive approach to labor market pooling. It jointly considers many different aspects of the agglomeration - labor market relationship, including turnover, learning, matching, and hold up. It also considers labor market pooling from the perspective of both workers and firms and across a range of industries. The paper reports a general positive relationship of turnover to local population density, which is consistent with theories of agglomeration and uncertainty. The paper also finds evidence of on-the-job learning that is consistent with theories of labor pooling, labor poaching, and hold up. In addition, the paper provides evidence consistent with agglomeration improving job matches. However, the labor market pooling gains that we measure are unlikely to account for a significant share of the agglomeration benefits accruing to worker and firms.

Key words: local labor markets, matching, turnover, learning, hold-up, agglomeration

JEL classification: R23, J60

*We are deeply indebted to people in the Statistics Department of the Bank of Italy, who helped us to phrase the questions included in the SHIW and SISF special sections on local labor markets. Our sincere thanks go to Giovanni D'Alessio, Leandro D'Aurizio, Ivan Faiella, Stefano Iezzi, Andrea Neri, and Federico Signorini. We are also grateful to Sabrina Di Addario, Massimo Omiccioli, Marcello Pagnini, Giacomo Calzolari and the participants to the *Seminario di analisi economica territoriale* (Bank of Italy, December 2011) and to the conference *Trasformazioni dei sistemi produttivi locali* (Bank of Italy and University of Bologna, February 2012) for insightful comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily correspond to those of the Institutions they are affiliated with.

[†]Bank of Italy, Branch of Napoli, Via Miguel Cervantes, 71, 80133 Napoli, Italy (e-mail: monica.andini@bancaditalia.it.).

[‡]Bank of Italy, Structural economic analysis department, Via Nazionale 91, 00184 Roma, Italy (e-mail: guido.deblasio@bancaditalia.it; website: <http://sites.google.com/site/guidodeblasiohomepage/Home>).

[‡]Department of Economics, University of Toronto, 150 Saint George Street, Toronto, Ontario M5S 3G7, Canada (e-mail: gilles.duranton@utoronto.ca; website: <http://individual.utoronto.ca/gilles/default.html>). Also affiliated with the Centre for Economic Policy Research, the Rimini Centre for Economic Analysis, the Centre for Economic Performance and the Spatial Economic Research Centre at the London School of Economics.

[§]Rotman School of Management, 105 Saint George Street, Toronto, Ontario M5S 3E6, Canada (e-mail: wstrange@rotman.utoronto.ca; website: <http://www.rotman.utoronto.ca/~wstrange/>).

1. Introduction

As with most economic research on urban labor markets, this paper begins with Marshall (1890). His well-known taxonomy of the sources of external economies of scale includes knowledge spillovers, input sharing, and - most importantly for our purposes - labor market pooling. The latter refers to the advantages for workers and firms deriving from sharing a labor market that is territorially limited to a small area: the local labor market. For instance, in a thicker local labor market workers might be able find a job faster. Similarly, firms might also fill their vacancies faster. In addition, firms and workers are more likely to find better matches for each other in terms of skills and experience. Moreover, workers might acquire more knowledge through learning spillovers. At the same time, job opportunities in competing firms might discourage firms to invest in their workers' training.

This paper employs a unique Italian data source to take a comprehensive approach to labor market pooling. The paper looks across all industries from the perspectives of both workers and firms and it considers many different aspects of labor market pooling, including turnover, matching, hold up and learning. To our knowledge, this is the first time such variables are used in a study of the economic effects of agglomeration. Our main data sources are the 2006 Survey of Household Income and Wealth (SHIW) and the 2007 Survey on Industrial and Service Firms (SISF). These Bank of Italy Surveys are described in greater detail below. They are valuable for our purposes because they provide information on aspects of labor market pooling such as turnover, the suitability of a worker for his or her job, on-the-job learning, training, and so on. This type of information is not available from the standard administrative sources used by previous research on the subject. We match this data with data from the Italian National Institute of Statistics to assess the thickness of the labor market in which firms and workers operate and to control for other aspects of these locations.

In order to establish a context for our investigation of labor market pooling, we begin by estimating models of the urban wage premium and of the relationship between agglomeration and firm output per worker. Our results here are consistent with the pattern of results from other empirical work on agglomeration. There is consistent evidence of an urban wage premium. In addition, firm output per worker is positively related to population density.

The labor market pooling results that we find are, when taken as a whole, largely consistent with the theoretical literature. There is a general positive relationship of turnover to density, which is consistent with theories of agglomeration and uncertainty. The paper also finds evidence of on-the-job learning that is consistent with theories of labor pooling, labor poaching, and hold up. In addition, the paper provides evidence consistent with agglomeration improving job matches. Overall, we find evidence of a variety of channels for labor market pooling.

We also note that these effects are rather modest, which can be interpreted in a number of ways. One possibility is that greater urban density improves the workings of local labour markets but only modestly so. Another is that the weak relationship may, in some cases, reflect a complicated equilibrium relationship between labor pooling and density. For instance, we find a relatively weak relationship between a worker's self-reported appropriate experience for a job and density.

This should arguably reflect the combination of two different effects: the influence of a thick market on the worker-job match (which would tend to find better fit with higher density) and the tendency of jobs requiring specialized skills to locate in thick markets (which would tend to have the opposite effect). Another possible interpretation of the modest coefficients is that labor market pooling operates differently across different industries. For instance, it is common to consider the relationship between agglomeration and turnover for high-technology industries. If the relationship is strong in this sector but not in others, then estimating over all industries will produce aggregate coefficients that fail to capture the relationships at work in individual sectors. More generally, if agglomeration effects are particular to sectors or industries, imposing the specification that effects are the same across sectors can fail to uncover agglomeration effects. Unfortunately, our data do not allow us to say more about the sources of the small coefficients. We hope further research will be able to shed more light on this issue. For now we offer the following conclusion. We find evidence consistent with a variety of local labour market pooling mechanisms. However, the effects we evidence are small and appear to account for only a small fraction of agglomeration economies.

The remainder of the paper is organized as follows. Section 2 provides the literature context. Section 3 discusses theories of labor market pooling. Section 4 presents the details of the paper's data sources. Section 5 includes the results of the estimates of the agglomeration-wage and agglomeration-productivity relationship. Section 6 contains the estimates of the relationship between agglomeration and turnover, learning, matching, and other aspects of labor market pooling. Section 7 assesses the importance of our measures of labor market pooling in the agglomeration-wage and agglomeration-productivity relationship. Section 8 concludes.

2. Related literature

Marshall's insights have motivated a long line of research on labor market pooling as a micro-foundation for agglomeration economies. This section provides a snapshot on the theoretical and empirical contributions of the literature.

Theoretical research on labor market pooling formalizes the elements of Marshall's analysis and also extends them in various directions. Helsley and Strange (1990) show how the matching of workers who are heterogeneous in their skills and firms who are heterogeneous in their labor demands can generate an agglomeration economy. Krugman (1991) models the effects of shocks on workers and firms. Overman and Puga (2010) extend this approach to derive the specific prediction that industries facing stronger idiosyncratic shocks will exhibit a greater tendency to agglomerate and that agglomeration will be associated with worker turnover. Matouschek and Robert-Nicoud (2005), Combes and Duranton (2006), and Almazan, de Motta, and Titman (2007) all consider the tension between the beneficial turnover considered by Marshall and the risks that firms and workers face that others – either their opposites or their rivals – will expropriate value created by specific investments. In particular, a firm may be reluctant to train its workers if this training would provoke either opportunism by its employees or poaching by its rivals. More recent theoretical papers on labor pooling include Gerlach, Rønde, and Stahl (2009), who consider the

interaction between labor pooling and innovation, and Picard and Wildasin (2011), who consider the interaction with input sharing. A survey of the larger microfoundations literature, including labor market pooling, can be found in Duranton and Puga (2004).

The empirical literature on labor market pooling is a part of the very large literature that considers agglomeration economies more generally. This literature has established a robust relationship between various sorts of agglomeration and productivity. Although much of this literature has focused on manufacturing industries, the relationship is also present in service sectors. Theories of agglomeration economies capturing all three of Marshall's microfoundations all predict this agglomeration-productivity relationship. As a result of this "Marshallian equivalence" (see Duranton and Puga, 2004), there remains a lot of uncertainty about the relative strengths of the various agglomeration forces. Looking at coagglomeration patterns across a range of industries, Ellison, Glaeser, and Kerr (2010) find that firms drawing from the same sorts of labor pool tend to coagglomerate. Jofre-Montseny, Marín-López, and Viladecans-Marsal (2011) carry out a similar exercise and also find evidence consistent with labor market pooling.¹

There is also a smaller but growing empirical literature that has looked specifically at labor market pooling. Papers in this literature have uncovered a number of instances where Marshallian labor market pooling seems to be at work. Fallick, Fleischman, and Rebitzer (2006), for instance, show that mobility rates in California's computer clusters, including the Silicon Valley, are high. Freedman (2008) finds that agglomeration in the software publishing industry to be associated with more turnover in the sense that job durations are shorter and mobility is greater. Wheeler (2008) finds the agglomeration-turnover relationship to be strongest for young workers. With regard to matching, Andersson, Burgess, and Lane (2007) find evidence of stronger positive assortative matching in larger markets, while Di Addario (2011), using Italian data, finds a greater rate of transitions from unemployment to employment. Glaeser and Maré (2001), Wheeler (2006), and De La Roca and Puga (2012) all provide evidence that the urban wage premium rises with a workers tenure in a city, a finding consistent with learning. Bacolod, Blum, and Strange (2010) provide some direct evidence of skill acquisition in cities. Finally, Overman and Puga (2010) show that industries more subject to shocks are more likely to cluster, a result consistent with the labor market pooling reducing risk.

3. Local labor markets: theories

This section does not break new theoretical ground. Its purpose is to outline a heuristic theory of labor market pooling that can be used to understand the patterns in the data. Suppose that there exist two classes of agent: workers and entrepreneurs. The latter are the creators and managers of the firms that employ the workers.

Workers choose locations to maximize expected lifetime utility. There are many components to the worker's objective: the worker's current wage and cost of living, the learning that takes place, the impacts of learning on future wages. Workers have heterogeneous skills and experience, which

¹For further references, see the surveys by Rosenthal and Strange (2004), Glaeser and Gottlieb (2009), and Puga (2010).

impacts the productivity of their match with an employer. This differentiation is both vertical (in the sense that some workers are better matches with all potential employers) and horizontal (in the sense that some workers are better matches with particular employers). The key labor market equilibrium condition is that no worker of a given type can obtain a higher level of expected utility in another city.

Entrepreneurs choose whether or not to become active in a given location.² This depends on the expected profit of the created enterprise. There are many risks that an entrepreneur faces. The latent profitability of the enterprise and entrepreneur are uncertain at the time of entry. The challenges and opportunities that will confront the firm are also uncertain. And there is an uncertain element to the profit contribution from the match with a particular worker, which may depend on the training that the firm gives to the worker. Ex post, some firms will survive, while others will fail depending on how these uncertainties are resolved. The key equilibrium condition for entrepreneurs is that entry drives profit to zero.

Cities have many impacts on workers, firms, and the labor markets that connect them. Workers may learn more in cities, developing skills that are valuable to their employers. Workers and firms may find closer matches in cities, increasing productivity. Workers in cities may be more mobile between firms, allowing firms to be more flexible and to grow more quickly in response to positive shocks. All of these are instances of agglomeration economies.

In addition to agglomeration economies, in this situation we also have sorting of both workers and firms, with workers migrating between cities and entrepreneurs choosing whether or not to become active according to city characteristics. We also have selection, where some matches of workers and firms survive.

This heuristic analysis has a number of predictions that can be taken to data. First, wages and other productivity measures will be positively associated with agglomeration. An observed relationship between agglomeration and wages or productivity could, of course, reflect any combination of agglomeration spillovers or sorting or selection. Second, turnover will tend to be greater in thicker markets. Third, agglomeration will have an ambiguous effect on learning, with knowledge spillovers encouraging learning and holdup discouraging learning. Fourth, agglomeration will have an ambiguous effect on measures of the match quality between worker and firm, since any better matching that cities provide will attract jobs where matching is more difficult. We now turn to the data that we use to evaluate these predictions.

4. Data

We turn attention to our data. Our two main sources of data are the 2006 Survey of Household Income and Wealth (SHIW) and the 2007 Survey on Industrial and Service Firms (SISF). Both surveys are conducted by the Bank of Italy. Appendix A provides further details about them.

²An alternative approach would be to suppose that entrepreneurs are mobile or “footloose.” This will give a different allocation of firms across space, but there would still be an equal profit condition, and this would mean that the analysis would be essentially similar. Compare for instance Venables (2011) (with mobile workers) with Nocke (2006) or Baldwin and Okubo (2006) (with mobile entrepreneurs/firms).

These surveys regularly collect standard information about households and firms in the manner of, for instance, the US Current Population Survey for households. These two surveys are also supplemented by special sections. The 2006 SHIW and the 2007 SISF each contain a section of questions about local labor markets. These questions were designed jointly between us and the survey administrators at the Statistics Department of the Bank of Italy to investigate the functioning of local labor markets. The household survey contains 12 questions about the working of local labor markets and the firm survey contains another 5.³ A full list of these questions is reported in Appendix B. Using the two surveys together allows us to consider jointly the worker and firm sides of local labor markets while matching these outcomes with information about local labor markets.

Our sample includes 4,367 workers (excluding government employees) and 3,660 firms. Appendix C documents a number of data issues. In particular, the questions pertaining to labor market issues were often asked to only a subsample of firms or workers. To link workers and firms we make use of confidential information about the municipality of residence for workers and location for firms.

Workers and firms are distributed over 226 and 439 local labor markets which we refer to as ‘cities’. Cities are functional areas based on the self-containment of commuting flows. They are defined by the Italian National Institute of Statistics (ISTAT) on Census commuting data at the municipality level. Appendix D provides additional details. The agglomeration measure we use is the log of 2001 city population density, provided by ISTAT. For instrumental variable (IV) estimations we use long lagged values of density from 1871 and 1921. In this we follow the literature and use the fact that local employment is to a large extent historically pre-determined while local productivity is likely to have changed a lot over time.

We also make use of measures of industrial agglomeration. Some cities are identified as industrial districts (IDs), based on ISTAT’s Cluster Mapping Project (ICMP). Details are provided in Appendix D.

Table 1 reports some summary statistics for the answers to the questions from the two surveys and for other variables we use. Throughout the paper, household variables (collected through the SHIW) are labelled with H and firm variables (collected through the SISF) with F. The questions pertaining to the functioning of local labor markets are organized around three themes:

Turnover and flexibility. To assess whether denser markets are associated with greater labor market flexibility, workers are asked if they changed employer *or* type of work in the recent past (H2). As discussed above, within denser areas job changes are more likely within occupations, a question allows us to disentangle changes in employer that do not carry with them changes in the type of job (H3). On the firm side, the sort of turnover associated with labor market pooling is voluntary. We therefore measure turnover by the percentage of terminations due to voluntary

³The original design of the local labor market sections of these two surveys contained more questions. For instance, questions on labor market prospects for workers or questions about the poaching of employees by competition were proposed. Unfortunately, these questions could not be included because, as the pilot showed, they were extremely difficult to answer. Cost-effectiveness considerations required to cut them off. In addition, we also discarded a small number of questions included in the surveys because of a small number of observations.

Table 1: Summary statistics for our main variables

Question	Variable	Obs.	Mean	Std. Dev.	Min.	Max.
Panel A. Household survey (SHIW)						
H1	log wage	4367	2.01	0.42	0.16	3.87
H2	Change of employer or type of work	1287	0.09	0.28	0	1
H3	Change of employer but not of type of work	117	0.47	0.50	0	1
H4	Workplace learning	1287	0.02	0.14	0	1
H5	Useful past experience	945	0.48	0.49	0	1
H6	Training by firm	1287	0.28	0.45	0	1
H7	Skill transferability	1287	0.84	0.35	0	1
H8	Difficulty of finding a replacement by employer	1287	0.10	0.31	0	1
H9	Difficulty of finding an equivalent job	1287	0.15	0.36	0	1
H10	Worker specialization	1606	0.07	0.26	0	1
H11	Appropriate experience	1606	0.83	0.37	0	1
H12	Appropriate education	1606	0.73	0.43	0	1
	log density	4367	5.71	1.08	3.13	8.28
	Industrial district	4367	0.26	0.43	0	1
	South	4367	0.25	0.43	0	1
	Male	4367	0.64	0.47	0	1
	Education (years)	4367	10.77	3.44	0	20
	Experience (years)	4367	21.05	11.02	1	49
Panel B. Firm survey (SISF)						
F1	log output per worker	3660	5.34	0.78	3.30	7.37
F2	Share of terminations voluntary	2750	0.67	0.39	0	1
F3	Share of vacancies filled from same sector	2452	0.46	0.40	0	1
F4	Number of days to train key workers	2553	17.18	23.71	0	160
F5	Appropriate experience and education of new key workers	2946	0.72	0.45	0	1
	log density	3660	5.68	1.06	2.53	8.28
	Industrial district	3660	0.28	0.45	0	1
	South	3660	0.31	0.46	0	1
	Age	3660	32.51	24.41	1	272
	Status: limited liability	3660	0.95	0.22	0	1
	Part of a group	3649	0.42	0.49	0	1

Sources: Bank of Italy, Survey on Italian Households' Income and Wealth 2006 and Survey of Industrial and Service Firms 2007.

resignations (F2), as reported by the owner or the manager of the firm. Relatively few workers report changing jobs (9%), so the overall magnitude of turnover is moderate.⁴ The worker survey reports that roughly half of the job changes entail a different employer but not a different type of work. The firm survey shows that 67% of terminations are voluntary. All this paints a picture with relatively modest labor market fluctuations. We return to this below.

⁴See Shimer (2005) for US figures.

Learning and holdup. To gauge the importance of density for learning, workers are asked whether they acquired their skills informally from colleagues inside or outside the firm (H4) and whether they find it useful for their current job any previous experience gained in the same field (H5). The possibility of hold-up problems is investigated by looking at the training provided by firms (H6) and skill transferability (H7). By the same token, firms are requested to report the percentage of vacancies filled in by workers with previous experience in the same sector (F3) and the amount of formal training they provide to their workers (F4). The responses here show training taking place within the employee-employer relationship. They also show that past worker experience is relevant to the worker's current job. It is striking, however, that only 2% of workers say that they have learned from informal contacts within the firm. We return to this issue below as well.

Matching. The theories discussed above also highlight that workers and firms will both face less risk in finding a job and filling a vacancy in a denser labor market. We ask workers to assess the ease of replacement faced by their employer, should the worker quit (H8). We also ask workers how easy would be for them to find another job similar in terms of salary or overall quality, should they lose their current job (H9). Relatively few workers –10% and 15% respectively – report substantial difficulty in these employment transitions. The degree of job specialization is measured by the answers from a question that requires workers to compare their level of specialization with that of other people in Italy who perform the same job (H10). Even fewer workers (7%) report having highly specialized jobs. The quality of matches is captured by two questions on the appropriateness for the job of, respectively, work experience (H11) and educational qualification (H12). A large majority of workers consider themselves to be well-matched according to these two measures (83% and 73% respectively). Similarly, firms are requested to assess the suitability of their workers in terms of experience and education (F5). The answers here are quantitatively consistent with those for the parallel questions in the worker survey. Overall, the survey responses seem to show labor to be relatively unspecialized, with good matches of workers to jobs.

We also use a number of control variables in our regressions to alleviate possible concerns about selection in our samples of workers and firms.⁵ At worker level, our controls include gender, education, experience, and its square. In a robustness check, we make use of the (confidential) data on the birthplace, which allows us to identify the movers (i.e., those who moved away from their birthplace). At the firm level, controls include: age of the firm, legal status (limited or unlimited liability) and being part of a broader corporate structure. In the robustness checks, we also look more specifically at manufacturing firms and small and medium enterprises (SMEs).

5. Agglomeration

Our interest in labor market pooling arises from its role in the generation of agglomeration economies. Before turning to labor market pooling in the next section, this section considers agglomeration economies. The specific focus is on the relationship between agglomeration and outcome measures such as wages and output to workers.

⁵Sample selection remains a concern if it is driven by an unobservable characteristic that is correlated with an observable of interest. It is econometrically equivalent to the issues of simultaneity and missing variables that we discuss below.

Our main estimating equation for worker i is

$$Y_i = \beta_0 + B_{c(i)} \beta_1 + X_i \beta_2 + \epsilon_i \quad (1)$$

where our dependent variable Y_i is here the log of the hourly wage of worker i , $B_{c(i)}$ is a vector of characteristics for city c where worker i works, and X_i is a vector of individual characteristics.⁶ Finally, ϵ_i is an error term that needs to be clustered by city given that city level explanatory variables apply to all workers within a city (Moulton, 1990).⁷

Table 2 reports results for eight wage regressions. In column 1, we regress log wages on the log of city density alone.⁸ In column 2 we add dummy variables for being classified as an industrial district and being located in the South of Italy. These are standard controls for Italian data. In column 3, we also add four standard individual controls: a male dummy, years of education, labor market experience, and its square. In columns 4 to 7, we retain the same specification as in column 3 but consider different subsamples of observations; males in column 4, workers above the age of 40 in column 5, workers with more than 13 years of education (which corresponds to higher education) in column 6, and workers who do not live in their place of birth in column 7 (movers). Finally, in column 8 we recognize that density and wages might be simultaneous determined since we expect places that pay higher wages to be more densely populated. Following earlier literature (e.g. Ciccone and Hall, 1996, Combes, Duranton, Gobillon, and Roux, 2010), we instrument contemporaneous density with long population density lags, namely log 1871 and log 1921 population density.⁹

In column 1, the estimate for the elasticity of wages with respect to density is 0.040. Controlling for other local characteristics and for individual characteristics lowers this estimate to 0.027. Columns 4 to 7 show that agglomeration effects appear stronger for old workers and university educated workers. However for none of these subsamples is the difference relative to whole sample statistically significant. Finally, in column 8, instrumenting for density yields

⁶Location controls include population density (our variable of interest), being part of an industrial district, and a South of Italy dummy. Individual controls include gender, education, experience, and its square.

⁷Furthermore, we did not make use of sample weights in the estimations reported in the paper. We duplicated all our estimations using sample weights. We also tried to weight our estimations so that the SHIW and SIF samples would match the distribution of area employment in Italy. Finally, we also replicated all our estimation excluding all observation from areas with less than 25,000 inhabitants. For these three robustness checks, the results are essentially the same as those reported here.

⁸We use density rather than total population because this variable more robust to the way boundaries are drawn. In particular, municipalities are part of the same local labor market only when both the share of working residents working locally and the share of employees residing locally is at least 75%. This is a restrictive definition relative to other countries as some ‘suburban’ municipalities may form a separate local labour market even though they belong to the same metropolitan areas in many other dimension. This should also ease comparisons since density is used more often than population in the recent literature (see Combes, Duranton, and Gobillon, 2011, for a longer discussion of these issues). In particular density is less sensitive to the modifiable areal unit problem (MAUP) than population.

⁹In our context, these density lags provide strong predictors for contemporaneous density. To be valid these instruments need to be correlated with wages only through contemporaneous density. The argument that ‘residual productivity’ (i.e., the error term in the wage regression) should be uncorrelated with our instruments rely on the fundamental changes that have affected the Italian economy since 1871 and 1921. Other research has also shown that these historical instruments yield results similar to alternative instruments based on geology (Combes *et al.*, 2010). Of course, the validity of an instrument is always potentially problematic. This and other issues in the estimation of (1) are discussed in Combes *et al.* (2011).

Table 2: Wage equations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Raw	Geog	Geog+Indiv	Males	Old	HighEd	Movers	IV
log density	0.040 ^a (0.015)	0.035 ^a (0.0081)	0.027 ^a (0.0066)	0.026 ^a (0.0078)	0.035 ^a (0.0080)	0.049 ^b (0.023)	0.027 ^b (0.012)	0.022 ^a (0.0085)
Indus. district		-0.028 (0.021)	-0.0016 (0.018)	-0.0038 (0.022)	-0.0048 (0.024)	0.067 (0.053)	0.0096 (0.026)	0.0024 (0.018)
South		-0.18 ^a (0.026)	-0.14 ^a (0.024)	-0.14 ^a (0.025)	-0.13 ^a (0.027)	-0.16 ^b (0.071)	-0.15 ^a (0.046)	-0.13 ^a (0.024)
Male			0.15 ^a (0.012)		0.18 ^a (0.019)	0.14 ^a (0.045)	0.19 ^a (0.018)	0.15 ^a (0.012)
Education			0.041 ^a (0.0024)	0.043 ^a (0.0030)	0.044 ^a (0.0035)	0.059 (0.036)	0.042 ^a (0.0038)	0.041 ^a (0.0025)
Experience			0.026 ^a (0.0024)	0.027 ^a (0.0032)	0.017 ^b (0.0076)	0.045 ^a (0.0090)	0.022 ^a (0.0040)	0.026 ^a (0.0025)
Experience2 (X100)			-0.038 ^a (0.0055)	-0.037 ^a (0.0069)	-0.024 ^c (0.0013)	-0.073 ^a (0.027)	-0.033 ^a (0.0091)	-0.037 ^a (0.0056)
Observations	4,367	4,367	4,367	2,804	2,293	398	1,162	4,252
R ²	0.01	0.04	0.19	0.18	0.19	0.28	0.21	-

Notes: The dependent variable is log wage (question H1). All regressions include a constant. Robust standard errors clustered by city in parentheses. *a, b, c*: significant at 1%, 5%, 10%. Columns 1 to 7 are estimated with OLS. Column 8 is estimated by TSLS using 1871 and 1921 populations as instruments. The first-stage statistic for these two instruments is 136, and the p-value of the overidentification test is 0.27.

a slightly lower point estimate of 0.022. This estimate implies that a one standard deviation increase in log density (i.e., +1.09) leads to a relatively modest 2.4% increase in wages. Turning to the other coefficient, we find a negative wage penalty for the South of Italy and an absence of significant results for being part of an industrial district. For individual characteristics, we find – unsurprisingly – higher wages for male workers, more educated workers, and more experienced workers.¹⁰

Next, we conduct a similar exercise using firm level data and estimate regressions corresponding to (1) on the firm side. For firm j , our estimating equation is

$$Y_j = \beta_0 + B_{c(j)} \beta_1 + X_j \beta_2 + \epsilon_j, \quad (2)$$

and our dependent variable Y_i is now the log of output per worker i , $B_{c(j)}$ is still a vector of characteristics for city c where firm j is located, and X_j is a vector of characteristics of firm j .

Table 3 report results for the estimation of equation (2). The structure of this table mirrors that of table 2. Column 1 uses only log density as explanatory variable. Column 2 adds again dummy variables for the South of Italy and industrial district. In column 3, the firm characteristics we

¹⁰The measured returns to education in table 2 are on the low side relative to estimates for other countries. This finding is not unique to our work (see for instance Di Addario and Patacchini, 2008, Dalmazzo and de Blasio, 2011, for work with a regional focus) and may be related, among other things to the fact that the skills provided by the public education system are different from those demanded by the firms (see for instance Tarantola, 2011).

Table 3: Output per worker

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Raw	Geog	Geog+Firm	Old	Group	SME	Manuf	IV
log density	0.062 ^c (0.033)	0.065 ^a (0.022)	0.036 ^c (0.019)	0.052 ^b (0.025)	0.030 (0.028)	0.057 ^a (0.017)	0.053 ^a (0.013)	0.033 ^c (0.019)
Indus. district		0.024 (0.039)	0.034 (0.037)	0.012 (0.063)	0.057 (0.050)	0.0012 (0.042)	-0.020 (0.036)	0.038 (0.039)
South		-0.26 ^a (0.044)	-0.17 ^a (0.041)	-0.098 (0.081)	-0.17 ^a (0.065)	-0.20 ^a (0.042)	-0.24 ^a (0.035)	-0.17 ^a (0.042)
Age			0.0016 ^a (0.00057)	6.0e-07 (0.00096)	0.0012 ^c (0.00066)	0.0015 ^b (0.00071)	0.0015 ^a (0.00050)	0.0014 ^b (0.00057)
Limited liability			0.35 ^a (0.099)	0.039 (0.14)	0.48 ^a (0.15)	0.15 (0.10)	-0.18 ^c (0.094)	0.38 ^a (0.099)
Group			0.29 ^a (0.026)	0.31 ^a (0.058)		0.30 ^a (0.031)	0.37 ^a (0.030)	0.30 ^a (0.027)
Observations	3,660	3,660	3,649	908	1,518	2,777	2,645	3,551
R ²	0.01	0.03	0.08	0.06	0.03	0.07	0.13	-

Notes: The dependent variable is log output per worker (question F1). All regressions include a constant. Robust standard errors clustered by city in parentheses. *a, b, c*: significant at 1%, 5%, 10%. Columns 1 to 7 are estimated with OLS. Column 8 is estimated by TSLS using 1871 and 1921 populations as instruments. The first-stage statistics for these two instruments is 196 and the p-value of the overidentification test is 0.095.

consider are: age, a dummy variable for limited liability status, and a dummy variable for being part of a group.¹¹ In columns 4 to 7, we consider different subsamples; firms older than 40 years, firms that are part of a larger group of firms, small and medium enterprises with employment below 100, and manufacturing firms. Finally, in column 8 we instrument log density as previously.

In column 1, the estimate for the elasticity of output per worker with respect to density is 0.062. Controlling for other local characteristics raises this estimate marginally while the introduction of establishment level characteristics lowers it to 0.036. Columns 4 to 7 show that agglomeration effects appear stronger for old establishments, SMEs, and manufacturing establishments. The estimates for establishments that are part of a group is insignificant. Dealing with the endogeneity of log density in column 8 again lowers the estimate to 0.033. Consistent with table 2, productivity per worker is much lower in the South of Italy whereas being part of an industrial district makes no significant difference. We also find that establishments that are part of a group and old establishments are more productive. For the limited liability status dummy, the picture is more mixed.

Overall these results are consistent with previous findings of the literature. Relative to results for France (e.g., Combes *et al.*, 2010) and the US (e.g., Ciccone and Hall, 1996, Glaeser and Resseger, 2010), the estimated elasticity of wages with respect to density is slightly lower. It is about one to

¹¹We do not consider firm size since it could be a consequence of being more productive.

two percentage points lower than in these other countries. Relative to existing agglomeration findings on Italian data (de Blasio and Di Addario, 2005, Di Addario and Patacchini, 2008, Mion and Naticchioni, 2009), we find slightly higher coefficients, by about one percentage point.¹² Overall we take these magnitudes as very close given differences in the data being used and differences in the estimation. Many detailed aspect of the wage findings are also in line with existing results.¹³ The results from table 3 are more difficult to compare since output per worker is seldom used in the recent literature. The estimates for the density elasticity of output per worker are nonetheless close to the estimates of the density elasticity of TFP in Combes *et al.* (2010).¹⁴

Compared to other countries, agglomeration effects are thus slightly lower when measured on wages, while they seem to be about the same when measured on output per worker. A possible explanation for this refers to the centralized bargaining system, which prevents wages from reacting in full to local labor market conditions.¹⁵

Simple theoretical considerations suggest that when agglomeration increases productivity and when local wages and local prices are set competitively, denser areas should exhibit higher wages and higher prices for nontradables. Imposing uniform wages across locations can lead to a higher pressure on local prices for nontradables, as these goods capitalize all the agglomeration gains (Dalmazzo and de Blasio, 2011). Uniform wages can also reduce mobility across cities. This is consistent with the fact that labour mobility is indeed low in Italy (Faini, Galli, Gennari, and Rossi, 1997) and that economic activity in Italy is less concentrated than in some other countries. For instance, the four most populated Italian cities host less than 20% of the national population whereas in the UK or France, the same proportion is attained by the largest city only.¹⁶

Finally, we note that the regressions reported in tables 2 and 3 follow the current practice in agglomeration work and focus mostly on gains from agglomerations that take place across sectors (e.g., Ciccone and Hall, 1996, Combes *et al.*, 2010). The generally insignificant coefficient on industrial districts and the results in column 7 of table 3 (where we restrict our sample of firms to manufacturing) relative to the other columns of the same table are not supportive of major agglomeration effects within sectors. In regressions not reported here, we experimented extensively with agglomeration effects at the sector level using the share of workers employed in

¹²If we consider only full-time workers, the coefficient on density decreases marginally. This is unsurprising. labour flexibility is arguably part of the benefits from agglomeration. Excluding flexible contractual arrangements may lead to the underestimation of the benefits from agglomeration.

¹³Among the features already documented in the references cited above and others: the halving of the density elasticity when individual controls are induced, the small further decline of this elasticity when instrumenting by long historical lags, the higher estimates for movers and more highly skilled workers, etc.

¹⁴In addition, the stronger agglomeration effects for firms that are smaller and not part of a group is also reminiscent of results by Henderson (2003) and Rosenthal and Strange (2010).

¹⁵For instance, the IMF recently mentioned that: "Italy's wage setting system is an obstacle for growth" (IMF, 2011).

¹⁶Clearly, the wage setting scheme might not be the only reason behind reduced mobility. As highlighted by Saraceno (1994) and Alesina and Ichino (2009), a first reason is the "familism" of the provision of welfare-related services. The family network in Italy offers most of the support for higher education, unemployment, and childcare and elderly-care. Thus, the opportunity cost of migration are high in Italy relative to the US (where the family network is much weaker) or other European countries (with a stronger welfare state not based on family ties). There is also evidence that labor mobility has been discouraged by the generous provision of public funds (government social transfers, which include unemployment benefits, social assistance, regular and invalidity pensions and health payments; Brunello, Lupi, and Ordine, 2001) and public sector jobs (Alesina, Danninger, and Rostagno, 2001) deployed in lagging (Southern) regions in the last decades.

the same industry to capture them.¹⁷ We failed to find robust results supportive of agglomeration effects taking mainly within broad sectors. This lack of result is consistent with recent literature which finds evidence for agglomeration effects within sectors but finds that those between sectors matter more (see, for instance, Combes, Duranton, and Gobillon, 2008, for a detailed comparison between these two types of agglomeration effects and the robustness of urbanisation effects to detailed sectoral controls).

6. Labor market pooling

Overview

In this section, we take advantage of the richness of the two surveys to explore the manifestation of labor market pooling in Italy. As noted previously, our approach departs from prior work by taking a comprehensive approach. We look across all industries, rather than focusing on a few. We look at both workers and firms, rather than looking at only one side of the market. Most importantly, we examine a number of possible ways that labor market pooling may manifest itself.

The approach is parallel to the previous section’s analysis of agglomeration. In the worker sample, we estimate regressions of the form

$$Z_i = \beta_0 + B_{c(i)} \beta_1 + X_i \beta_2 + \epsilon_i. \quad (3)$$

This specification mirrors regression (1) but considers, as dependent variable Z_i , a measure of labor market pooling for worker i . We consider a number of aspects of labor market pooling. Z_i thus includes dummy variables such as the change of employer or type of work or both (H2 and H3), workplace learning (H4), past experience (H5), training by the firm (H6), skill transferability (H7), difficulty of replacing the worker or finding another job (H8 and H9), and measures of specialization and the appropriateness of experience and education (H10-H12).

We are concerned primarily with the relationship of the labor market pooling variables with agglomeration but, in estimating equation (3), we include the other controls from the wage models reported in Table 2 as well. For questions from the worker sample, the tables in this section are organized in a way that parallels Table 2. In column 1, we regress the labor market pooling measure on city density. Column 2 adds the controls for a worker being located in an industrial district or in the South. Column 3 again adds the individual controls. Columns 4-7 are estimates of the specification in column 3 over subsamples of males, older workers, more educated workers, and workers who have moved. Finally, column 8 is the long lagged instruments model. Estimation is carried out by Probit and iv Probit for which we report marginal effects.

In the firm sample, we estimate equations of a similar form:

$$Z_j = \beta_0 + B_{c(j)} \beta_1 + X_j \beta_2 + \epsilon_j, \quad (4)$$

This specification mirrors regression (2) but considers, as dependent variable Z_j , a measure of labor market pooling for firm j . These include the share of terminations that are voluntary (F2), the share

¹⁷For instance, including a dummy variable for manufacturing does not change anything the our results.

of vacancies filled from workers previously employed in the same sector (F_3), and the numbers of days to train and find key workers (F_4 and F_5).

In estimating (4), we include the other controls from Table 3. This produces results for firms that are organized in a way that parallels table 3. Column 1 presents a simple regression using density alone. Column 2 augments it with geographic controls. Column 3 also controls for firm age and dummy variables for limited liability status and for belonging to a group. The rest of the columns include regressions for the subsamples of older firms (more than 40 years), small and medium sized enterprises (employment below 100), and manufacturing firms. The final column presents the iv results using lagged densities, as above.

Turnover

Table 4 presents results that relate to the turnover element of labor market pooling. In this and in other tables, because our focus is on agglomeration, we report only the density coefficients. The only other agglomeration variable in the specifications is for a worker's location in a designated industrial district. This variable is in nearly every instance insignificant, so these coefficients are not reported. The evidence of our estimation does not show the industrial districts to be related to labor market pooling, either positively or negatively.¹⁸

Panel A presents models of the relationship between overall worker turnover and density. The results are quite consistent across specifications. There is not a significant relationship between worker job change and density. Looking across all industries in an Italian setting, we do not find more job turnover in dense markets. In light of prior literature (e.g., Wheeler, 2008), this could be due to the fact that workers experience greater turnover in denser cities early on in their career but eventually find better job matches and experience less turnover.

The sort of turnover that is predicted by Marshallian theories is of a particular sort, with workers moving jobs without changing houses. Panel B presents results on worker change of employer without change of type of employment. In contrast to overall turnover, this sort of turnover is positively and significantly related to density. Similarly, Panel C considers the firm-reported share of terminations that were voluntary. This is also positively and significantly related to density. Thus, although raw turnover does not show a strong relationship with agglomeration, the more refined measures in Panels B and C do show a consistent relationship. This evidence is qualitatively consistent with labor market pooling in the spirit of Marshall.

It is important, however, not to forget how little worker turnover there is in the Italian sample. As noted earlier, only 9% of workers reported changing jobs or employers in the previous two years. Of these, relatively few were turnovers that involved change of employer but not type of work (only 117 instances, amounting to 47% of the turnovers). Likewise, for the firms, only 67% of terminations were voluntary. Together, these results mean that although Marshallian turnover

¹⁸This result is consistent with the wage results of de Blasio and Di Addario (2005), who find no wage premium associated with being located in an industrial district. It casts a shadow on the empirical relevance of labor market advantages, which are regularly mentioned in the qualitative literature on Italian industrial districts originated from Becattini (1978, 1979); see also Brusco and Paba (1997).

Table 4: Turnover and flexibility

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Raw	Geog	Geog+Indiv	Males	Old	HighEd	Movers	IV
Panel A. Dependent variable: Worker job change (Question H2), probit and IV probit								
log density	-0.0048 (0.0088)	-0.0035 (0.0091)	-0.0027 (0.0092)	-0.0087 (0.011)	-0.0042 (0.0083)	0.025 (0.023)	-0.019 (0.016)	0.0077 (0.011)
Observations	1,287	1,287	1,287	806	721	100	383	1,251
pseudo-R ²	0.00	0.00	0.01	0.01	0.02	0.12	0.03	-
Panel B. Dependent variable: Worker change of employer but no change of type of work (Question H3), probit and IV probit								
log density	0.11 ^b (0.058)	0.12 ^b (0.057)	0.13 ^b (0.060)	0.10 (0.070)	0.25 ^b (0.11)	-0.00062 (0.0020)	0.043 (0.076)	0.14 ^b (0.069)
Observations	117	117	117	79	46	11	43	110
pseudo-R ²	0.03	0.04	0.06	0.10	0.20	1.00	0.09	-
Panel C. Dependent variable: Firm share of terminations voluntary (Question F2), OLS and TSLS								
	Raw	Geog	Geog+Firm	Old	Group	SME	Manuf	IV
log density	0.030 ^a (0.01)	0.034 ^a (0.01)	0.035 ^a (0.01)	0.042 ^a (0.01)	0.027 ^a (0.01)	0.041 ^a (0.01)	0.023 ^a (0.01)	0.050 ^a (0.01)
Observations	2,750	2,750	2,743	691	1,165	2,066	1,987	2,666
R ²	0.01	0.02	0.03	0.05	0.02	0.05	0.03	-

Notes: All regressions include a constant and follow the specifications of tables 2 and 3. Robust standard errors clustered by city in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. In panels A and B, columns 1 to 7 are estimated with probit. Column 8 is estimated by IV probit using 1871 and 1921 populations as instruments. Marginal effects reported for all probit regressions. In panel C, columns 1 to 7 are estimated with OLS. Column 8 is estimated by TSLS using 1871 and 1921 populations as instruments.

does increase with density, the magnitude of this turnover is modest. When this is combined with prior persuasive evidence of job-hopping in certain industries and certain places (e.g., Fallick *et al.*, 2006, Freedman, 2008), this suggests that the agglomeration-turnover relationship often highlighted in the literature is a particular one. It does not seem to apply in all situations.

Learning and holdup

Table 5 presents results that relate to the learning element of labor market pooling. Panel A presents estimates of models where the dependent variable is workplace learning. The results here are only weakly consistent with Marshall's insights on knowledge spillover in that learning increases with density in all specifications except for the older worker sample (age over 40). That there is less workplace learning by older workers is sensible and not really inconsistent with the idea of learning in cities. However, the estimates are noisy and mostly insignificant, including the preferred specification in column 3. In addition, very few workers report this sort of informal learning, only 2% as noted in Table 2. So while the positive relationship of workplace learning

to density is Marshallian, the lack of precision and the small number of workers impacted do not provide strong support for the knowledge spillovers of this sort as an agglomeration economy that operates across industries.

Panel B reports results for models of the importance of past experience in the same field. Across most of the models, density is significant and positive. This can be interpreted as evidence of prior learning in cities. However, the weak results in Panel A and the low rate of learning from other workers suggest that the learning is not very Marshallian. It is worth pointing out that the density coefficients are largest for the samples of educated workers and movers in columns 7 and 8 of the table. While this coefficient of 0.095 is not statistically different from the full sample coefficient in column 3 of 0.079, the larger coefficient for movers is at least somewhat suggestive of a role for sorting of high skill worker into larger cities.

Panels C and D of Table 5 address holdup. Panel C presents models of worker training by firms, while Panel D presents models of worker skill transferability. As noted in Combes and Duranton (2006), there is tension between a firm's desire to draw from a large labor pool and its aversion to competition with other firms for skilled workers. To the extent that a firm has trained an employee and the employee has thus acquired skills that are potentially transferable to other local employers, then the firm risks what Combes and Duranton (2006) call "labor poaching." This will discourage firms from training workers in ways that develop transferable skills. Panel C's results on training are consistent with firms wanting to avoid this sort of hold up. Worker training by the firm is consistently negatively and significantly related to density. Which means that while workers seem to have obtained useful past experience (Panel B), they have obtained the experience neither from other workers (Panel A) nor from training provided by employers (Panel C). In Panel D we see a relationship between density and worker skill transferability that is insignificant in all of the models.

The firm results presented in the last two panels of the table are noisier, but they ultimately tell a rather similar story. In Panel E, we observe a positive relationship between density and the firm's share of vacancies filled with workers with same sector experience. The coefficients are all insignificant, with the notable exception of the sample of firms that belong to a group (which likely reflects within-group labor market practices). Likewise, Panel F shows an insignificant positive relationship between the firm training days for a new key worker and density. The difference between this and the clear negative worker training results from Panel C might possibly hinge on the difference between "key" workers and the rest of the workforce.

Matching

We now discuss the matching aspect of labor market pooling. The main idea is that agglomeration allows a better match between firms and workers, to the benefit of both. As noted in Section 3, while one expects matches to be better in thick markets, one also expects jobs where matching is difficult to be found in cities. This means that the theoretical relationship between the surveys' measures of matching and density are theoretically ambiguous.

Table 5: Learning and holdup

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Raw	Geog	Geog+Indiv	Males	Old	HighEd	Movers	IV
Panel A. Dependent variable: Worker workplace learning (Question H4), probit and IV probit								
log density	0.0081 ^c (0.0041)	0.0052 (0.0032)	0.0047 (0.0030)	0.0072 ^a (0.0027)	0.00063 (0.0048)	0 (0)	0.000088 (0.0045)	0.0071 ^c (0.0041)
Observations	1,287	1,287	1,287	806	543	55	214	1,251
pseudo-R ²	0.01	0.06	0.08	0.09	0.01	0.90	0.07	-
Panel B. Dependent variable: Useful past experience (Question H5), probit and IV probit								
log density	0.076 ^a (0.018)	0.076 ^a (0.018)	0.079 ^a (0.018)	0.073 ^a (0.021)	0.061 ^a (0.018)	0.096 ^c (0.055)	0.095 ^b (0.033)	0.079 ^a (0.021)
Observations	945	945	945	594	553	80	304	913
pseudo-R ²	0.01	0.01	0.02	0.01	0.01	0.11	0.02	-
Panel C. Dependent variable: Worker training by firm (Question H6), probit and IV probit								
log density	-0.022 (0.020)	-0.028 ^c (0.017)	-0.031 ^c (0.017)	-0.033 ^b (0.015)	-0.025 (0.022)	-0.081 (0.069)	-0.033 (0.028)	-0.034 (0.021)
Observations	1,287	1,287	1,287	806	721	100	383	1,251
pseudo-R ²	0.00	0.01	0.04	0.05	0.06	0.07	0.03	-
Panel D. Dependent variable: Worker skill transferability (Question H7), probit and IV probit								
log density	0.012 (0.015)	0.010 (0.015)	0.011 (0.014)	0.0077 (0.013)	0.012 (0.014)	-0.00040 (0.030)	-0.0089 (0.020)	0.020 (0.013)
Observations	1,287	1,287	1,287	806	721	100	383	1,251
pseudo-R ²	0.00	0.02	0.05	0.04	0.04	0.03	0.07	-
Panel E. Dependent variable: Firm share of vacancies filled by workers with same sector experience (Question F3), OLS and TSLS								
	Raw	Geog	Geog+Firm	Old	Group	SME	Manuf	IV
log density	0.010 (0.01)	0.0081 (0.01)	0.011 (0.01)	0.0053 (0.02)	0.026 ^b (0.01)	0.0074 (0.01)	0.018 (0.02)	0.0060 (0.01)
Observations	2,452	2,452	2,443	604	1,014	1,878	1,767	2,376
R ²	0.00	0.01	0.01	0.02	0.01	0.01	0.02	-
Panel F. Dependent variable: Firm training days for a new key worker (Question F4), OLS and TSLS								
log density	0.26 (0.48)	0.32 (0.49)	0.36 (0.49)	0.97 (1.01)	0.38 (0.56)	0.81 (0.53)	-0.14 (0.56)	0.45 (0.64)
Observations	2,553	2,553	2,548	614	1,060	1,955	1,839	2,484
R ²	0.00	0.00	0.00	0.01	0.00	0.00	0.01	-

Notes: All regressions include a constant and follow the specifications of table 2 and 3. Robust standard errors clustered by city in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. In panels A to D, columns 1 to 7 are estimated with probit. Column 8 is estimated by IV probit using 1871 and 1921 populations as instruments. In panels E and F, columns 1 to 7 are estimated with OLS. Column 8 is estimated by TSLS using 1871 and 1921 populations as instruments. 16

Some of the previous results can also be interpreted as bearing on matching. The results on the relationship of agglomeration to useful past experience in Panel B of the previously discussed Table 5 clearly bear on matching as well as the learning that we have previously discussed. The results show a consistent positive relationship between a worker's useful experience and density. This is consistent with the idea that agglomeration improves matches.

Table 6 presents further results on the matching of workers to firms. Panel A presents results on the difficulty an employer is expected by a worker to encounter in finding a replacement. This is negatively related to density in all the models but one. The relationship is significant in all but the samples of highly educated workers and movers.

The rest of the worker estimates typically feature signs that are consistent with matching but with small coefficients. Panel B presents models of a worker's difficulty of finding an equivalent job. Although the coefficients here are all negative, none are significant. Panel C presents results of models of the relationship of a worker's specialization to density. Again, although the results are all negative, they are also all insignificant. In a similar spirit, the results in Panel D show a positive but insignificant relationship between a worker's appropriate experience and agglomeration. The only exception to this pattern of predicted signs but noisy measurement is in Panel E. Here, a worker's appropriate education is negatively but insignificantly associated with density.

The firm results in Panel F also address a worker's possession of appropriate education and experience for the job. This is positively related to density, but the coefficients are small and insignificant.

Taken as a group, these results offer some support for the idea that matching is improved by agglomeration, but like the results on knowledge spillovers, the support is weak. For two of the relevant survey questions, the results are significant and consistent with prediction. For five questions, the results are consistent with prediction, but insignificant. In only one case, results are inconsistent with prediction and insignificant. In no case do we have significant coefficients that are at odds with the predictions of matching models.

7. Labor market pooling and agglomeration

In this section we return to the agglomeration-wage relationship and to the agglomeration-productivity explored in section 5 but consider density and our measures of labor market pooling jointly as explanatory variables. This allows us to assess the association between specific dimensions of labour market pooling and wages or output per worker. We estimate regressions of the following form:

$$Y_i = \beta_0 + B_{c(i)} \beta_1 + X_i \beta_2 + Z_i \beta_3 + \epsilon_i, \quad (5)$$

where Y_i is the wage of worker i , $B_{c(i)}$ is a set of characteristics of city c where worker i works, X_i is a set of individual characteristics, and Z_i is a labor market pooling variable measured for worker i . We also estimate the corresponding regressions for output per worker

$$Y_j = \beta_0 + B_{c(j)} \beta_1 + X_j \beta_2 + Z_j \beta_3 + \epsilon_j, \quad (6)$$

Table 6: Matching

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Raw	Geog	Geog+Indiv	Males	Old	HighEd	Movers	IV
Panel A. Dependent variable: Difficulty of worker replacement for employer (Question H8), probit and IV probit								
log density	-0.019 ^c (0.010)	-0.017 ^c (0.0093)	-0.018 ^b (0.0091)	-0.025 ^b (0.010)	-0.012 (0.0099)	0.020 (0.016)	-0.029 (0.019)	-0.028 ^b (0.012)
Observations	1,287	1,287	1,287	806	721	100	383	1,251
pseudo-R ²	0.00	0.01	0.02	0.03	0.04	0.12	0.03	-
Panel B. Dependent variable: Worker difficulty of finding equivalent job (Question H9), probit and IV probit								
log density	-0.012 (0.012)	-0.0093 (0.012)	-0.013 (0.012)	-0.017 (0.013)	-0.0095 (0.015)	-0.010 (0.034)	-0.032 (0.020)	-0.022 (0.016)
Observations	1,287	1,287	1,287	806	721	100	383	1,251
pseudo-R ²	0.00	0.00	0.03	0.03	0.02	0.14	0.09	-
Panel C. Dependent variable: Worker specialisation (Question H10), probit and IV probit								
log density	-0.0020 (0.0076)	-0.0049 (0.0076)	-0.0066 (0.0074)	-0.0062 (0.0085)	-0.0012 (0.0073)	0.0055 (0.031)	-0.0041 (0.015)	-0.0038 (0.0097)
Observations	1,606	1,606	1,606	1,015	914	140	458	1,558
pseudo-R ²	0.00	0.01	0.00	0.03	0.06	0.05	0.02	-
Panel D. Dependent variable: Worker appropriate experience (Question H11), probit and IV probit								
log density	0.0079 (0.012)	0.0086 (0.012)	0.0087 (0.012)	0.0092 (0.011)	0.0011 (0.015)	-0.010 (0.026)	0.023 (0.023)	0.0033 (0.014)
Observations	1,606	1,606	1,606	1,015	914	140	458	1,558
pseudo-R ²	0.00	0.00	0.00	0.01	0.01	0.05	0.03	-
Panel E. Dependent variable: Worker appropriate education (Question H12), probit and IV probit								
log density	-0.014 (0.015)	-0.0085 (0.015)	-0.0088 (0.015)	-0.0083 (0.014)	-0.0099 (0.019)	0.029 (0.033)	0.0073 (0.026)	-0.011 (0.016)
Observations	1,606	1,606	1,606	1,015	914	140	458	1,558
pseudo-R ²	0.00	0.00	0.00	0.00	0.00	0.05	0.01	-
Panel F. Dependent variable: Appropriate education and experience of new key workers (Question F5), probit and IV probit								
	Raw	Geog	Geog+Firm	Old	Group	SME	Manuf	IV
log density	0.0094 (0.019)	0.012 (0.015)	0.0086 (0.015)	0.010 (0.020)	0.0070 (0.021)	-0.0014 (0.015)	0.017 (0.019)	-0.00052 (0.017)
Observations	2,946	2,946	2,939	731	1,222	2,252	2,143	2,862
pseudo-R ²	0.00	0.00	0.01	0.02	0.02	0.01	0.00	-

Notes: All regressions include a constant and follow the specifications of table 2 and 3. Robust standard errors clustered by city in parentheses. *a*, *b*, *c*: significant at 1%, 5%, 10%. In all panels A to E, columns 1 to 7 are estimated with probit. Column 8 is estimated by IV probit using 1871 and 1921 populations as instruments. In panel F, columns 1 to 7 are estimated with OLS. Column 8 is estimated by TSLS using 1871 and 1921 populations as instruments.

where the our dependent variable Y_i is now the log of output per worker i , $B_{c(j)}$ is still a vector of characteristics for city c where firm j is located, X_j is a vector of characteristics of firm j , and Z_i is a labor market pooling variable now measured for firm j . These two specifications basically augment specifications (1) and (2) with a labour market pooling variable.

In the different specifications reported in panels A and B of table 7, we consider all our measures of labor pooling at the worker level in turn. In the specifications reported in panel C of table 7, we also consider all our measures of labor pooling at the worker level. As can be seen from the results, only about half the labor pooling variables are significant in our estimation of equation (5) for workers. On the other hand, all the labor pooling variables are significant in our estimation of equation (6) for firms. Turning to the coefficient on density, it is significant in all cases but one. This coefficient is either slightly lower or of the same magnitude the same coefficient in the corresponding regression without any labour market pooling variable in column 3 of table 2 for the worker regressions of panels A and B. The standard errors are nonetheless generally higher due to the smaller samples of workers which were surveyed with the labour pooling questions. In panel C for the firm level regressions, the coefficient on density is insignificant and about one percentage point lower than in the corresponding regression in column 3 of table 3. Like with workers, the higher standard errors are caused in part by the fact that only a subsample of firms was surveyed on labour pooling issues. Some caution is obviously needed when interpreting these regressions since labour pooling variables measuring for instance a change of employer is likely to be determined simultaneously with the wage.

By estimating the relationship between our labour market variables and wages and knowing the relationship between density and our labour market variables from the tables above, we can assess how much of the relationship between density and wages (output per worker) is accounted for by these labour market variables. For instance, we know from column 3 of panel C of table 4 that an increase in log density by one point is associated with an increase of 0.035 in the share of voluntary turnover for firms. In column 1 of panel C of table 7, we report that the share of voluntary turnover for firms is positively associated with log output per worker with a coefficient of 0.13 for the former variable. Hence, an increase in log density by one point is associated with an increase in log output per worker of $0.035 \times 0.13 = 0.0046$. This represents about 13% of the total effect of log density reported in column 3 of table 3. The same calculation can be repeated to assess the role of the other labor market variables of 7. By doing that, we find that overall the labor market variables explain only a limited share of the urban wage/productivity premia.¹⁹

8. Conclusion

This paper looks at several different aspects of labor market pooling across a range of industries and from the perspectives of both firms and workers. The focus is on the microfoundations of agglomeration economies.

¹⁹We also performed a principal component analysis to obtain a synthetic indicator of labor market pooling channels and use it as alternative explanatory variable in the estimation of equations (5) and (6). The results confirm those presented in this section.

Table 7: Labour market pooling and wages / output per worker

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Dependent variable: log wage						
Labor market pooling variable:	job change	emp. change	learning	useful exp.	training	skill trans.
	(H2)	(H3)	(H4)	(H5)	(H6)	(H7)
Labor market pooling variable	-0.13 ^a	0.15 ^c	0.014	0.0069	0.072 ^a	0.025
	(0.043)	(0.079)	(0.060)	(0.031)	(0.020)	(0.026)
log density	0.016 ^c	0.0081	0.016 ^c	0.034 ^a	0.018 ^b	0.016 ^c
	(0.0082)	(0.044)	(0.0084)	(0.012)	(0.0085)	(0.0084)
Observations	1,287	117	1,287	945	1,287	1,287
R ²	0.24	0.13	0.23	0.16	0.24	0.23
Panel B. Dependent variable: log wage						
Labor market pooling variable:	difficu.	find eq.	spec.	approp.	approp.	
	replac.	job		exp.	edu.	
	(H8)	(H9)	(H10)	(H11)	(H12)	
Labor market pooling variable	-0.027	-0.023	0.020 ^a	0.017	0.053 ^a	
	(0.036)	(0.029)	(0.0054)	(0.024)	(0.020)	
log density	0.016 ^c	0.016 ^c	0.026 ^b	0.025 ^b	0.026 ^b	
	(0.0085)	(0.0084)	(0.011)	(0.010)	(0.011)	
Observations	1,287	1,287	1,606	1,606	1,606	
R ²	0.23	0.23	0.20	0.19	0.20	
Panel C. Dependent variable: log output per worker						
Labor market pooling variable:	volunt.	same	training	approp.		
	turnover	exp.	days	exp./edu.		
	(F2)	(F3)	(F4)	(F5)		
Labor market pooling variable	0.13 ^a	-0.15 ^a	0.0018 ^a	0.063 ^c		
	(0.039)	(0.039)	(0.00058)	(0.033)		
log density	0.018	0.025	0.027	0.029 ^c		
	(0.020)	(0.023)	(0.020)	(0.017)		
Observations	2,743	2,443	2,548	2,939		
R ²	0.09	0.08	0.07	0.07		

Notes: All regressions are estimated with OLS and include a constant. Robust standard errors clustered by city in parentheses. *a, b, c*: significant at 1%, 5%, 10%. In panels A to B, the specification is the same as that of column 3 of table 2 with one additional labour market pooling variable. In panel C, the specification is the same as that of column 3 of table 3 with one additional labour market pooling variable.

The paper's findings are broadly consistent with the many different dimensions of labor market pooling that have been discussed in the theoretical literature. The paper demonstrates a general positive relationship of turnover to density. It also offers evidence of on-the-job learning that is consistent with theories of labor pooling, labor poaching, and hold up. In addition, the paper provides evidence consistent with agglomeration improving job matches.

The results, however, are relatively modest. The paper shows that labor market pooling gains are unlikely to account for a significant share of the agglomeration benefits accruing to worker and firms. These results have several possible explanations. As noted above, this pattern may reflect, at least in part, the complex equilibrium relationships associated with agglomeration. It is also possible that labor market pooling is, at least in the Italian markets that we examine, not an important source of agglomeration economies. Or that there are different sources of agglomeration economies in different industries, making it difficult to identify a clear pattern of labor market pooling across all industries. The data do not allow us to determine which of these possible explanations are correct.

There is one strong suggestion that comes from the weak results, and that is that economists should attend to the specifics of industries in looking for evidence of the microfoundations of agglomeration economies. The various microfoundations proposed by Marshall and his successors may all be valid in certain situations but not in others. This means both that approaches that focus on particular and narrowly defined industries make a lot of sense and that one should be cautious in generalizing the results of these approaches. Similarly, policymakers should probably also be careful not to draw overly general lessons from the agglomeration successes of particular industries.

References

- Alesina, Alberto, Stephan Danninger, and Massimo Rostagno. 2001. Redistribution through public employment: The case of Italy. *IMF Staff Papers* 48(3):447–473.
- Alesina, Alberto and Andrea Ichino. 2009. *L'Italia fatta in casa. Indagine sulla vera ricchezza degli italiani*. Milano: Mondadori editore.
- Almazan, Andres, Adolfo de Motta, and Sheridan Titman. 2007. Firm location and the creation and utilization of human capital. *Review of Economic Studies* 74(4):1305–1327.
- Andersson, Fredrik, Simon Burgess, and Julia I. Lane. 2007. Cities, matching and the productivity gains of agglomeration. *Journal of Urban Economics* 61(1):112–128.
- Bacolod, Marigee, Bernardo S. Blum, and William C. Strange. 2010. Elements of skill: Traits, intelligences, education, and agglomeration. *Journal of Regional Science* 50(1):245–280.
- Baldwin, Richard E. and Toshihiro Okubo. 2006. Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. *Journal of Economic Geography* 6(3):323–346.
- Becattini, Giacomo. 1978. The development of light industry in Tuscany: An interpretation. *Economic Notes* 2-3:107–123.
- Becattini, Giacomo. 1979. Dal settore industriale al distretto industriale: Alcune considerazioni sull'unità d'indagine dell'economia industriale. *Rivista di Economia e Politica Industriale* 0(1):7–21.
- Brunello, Giorgio, Claudio Lupi, and Patrizia Ordine. 2001. Widening differences in Italian regional unemployment. *Labour Economics* 8(1):103–129.
- Brusco, Sebastiano and Sergio Paba. 1997. Per una storia dei distretti industriali italiani dal secondo dopoguerra agli anni novanta. In Francesco Barca (ed.) *Storia del capitalismo italiano dal dopoguerra a oggi*. Rome: Donzelli editore, 265–334.
- Bugamelli, Matteo, Riccardo Cristadoro, and Giordano Zevi. 2009. La crisi internazionale e il sistema produttivo italiano: un'analisi su dati a livello d'impresa. Occasional papers 58, Bank of Italy.
- Ciccone, Antonio and Robert E. Hall. 1996. Productivity and the density of economic activity. *American Economic Review* 86(1):54–70.
- Combes, Pierre-Philippe and Gilles Duranton. 2006. Labour pooling, labour poaching, and spatial clustering. *Regional Science and Urban Economics* 36(1):1–28.
- Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. Spatial wage disparities: Sorting matters! *Journal of Urban Economics* 63(2):723–742.
- Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2011. The identification of agglomeration economies. *Journal of Economic Geography* 11(2):253–266.
- Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2010. Estimating agglomeration economies with history, geology, and worker effects. In Edward L. Glaeser (ed.) *The Economics of Agglomeration*. Cambridge (Mass.): National Bureau of Economic Research, 15–65.
- Dalmazzo, Alberto and Guido de Blasio. 2011. Amenities and skill-biased agglomeration effects: Some results on Italian cities. *Papers in Regional Science* 90. Forthcoming.

- de Blasio, Guido and Sabrina Di Addario. 2005. Do workers benefit from industrial agglomeration? *Journal of Regional Science* 45(4):797–827.
- De La Roca, Jorge and Diego Puga. 2012. The dynamic earnings premium of dense cities. Processed, CEMFI and IMDEA Sociale Sciences.
- Di Addario, Sabrina. 2011. Job search in thick markets. *Journal of Urban Economics* 69(3):303–318.
- Di Addario, Sabrina and Eleonora Patacchini. 2008. Wages and the city. Evidence from Italy. *Labour* 15(5):1040–1061.
- Duranton, Gilles and Diego Puga. 2004. Micro-foundations of urban agglomeration economies. In J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2063–2117.
- Ellison, Glenn, Edward L. Glaeser, and William R. Kerr. 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review* 100(3):1195–1213.
- Faini, Riccardo, Gianpaolo Galli, Pietro Gennari, and Fulvio Rossi. 1997. An empirical puzzle: Falling migration and growing unemployment differentials among Italian regions. *European Economic Review* 41(3-5):571–579.
- Fallick, Bruce, Charles Fleischman, and James B. Rebitzer. 2006. Job-hopping in Silicon Valley: Some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics* 88(3):472–481.
- Freedman, Matthew. 2008. Job hopping, earnings dynamics, and industrial agglomeration in the software publishing industry. *Journal of Urban Economics* 64(3):590–600.
- Gerlach, Heiko A., Thomas Rønde, and Konrad O. Stahl. 2009. Labor pooling in R&D intensive industries. *Journal of Urban Economics* 65(1):99–111.
- Glaeser, Edward L. and Joshua D. Gottlieb. 2009. The wealth of cities: Agglomeration economies and spatial equilibrium in the United States. *Journal of Economic Literature* 47(4):983–1028.
- Glaeser, Edward L. and David C. Maré. 2001. Cities and skills. *Journal of Labor Economics* 19(2):316–342.
- Glaeser, Edward L. and Matthew R. Resseger. 2010. The complementarity between cities and skills. *Journal of Regional Science* 50(1):221–244.
- Helsley, Robert W. and William C. Strange. 1990. Matching and agglomeration economies in a system of cities. *Regional Science and Urban Economics* 20(2):189–212.
- Henderson, J. Vernon. 2003. Marshall's economies. *Journal of Urban Economics* 53(1):1–28.
- International Monetary Fund. 2011. Italy: 2011 Article IV Consultation Concluding Statement of the Mission, May 2011. Processed.
- Istituto Nazionale di Statistica (Italian National Institute of Statistics). 1997. *I sistemi locali del lavoro 1991*. Rome: ISTAT.
- Jofre-Montseny, Jordi, Raquel Marín-López, and Elisabet Viladecans-Marsal. 2011. The mechanisms of agglomeration. *Journal of Urban Economics* (forthcoming).
- Krugman, Paul R. 1991. *Geography and Trade*. Cambridge, MA: MIT Press.

- Marshall, Alfred. 1890. *Principles of Economics*. London: Macmillan.
- Matouschek, Niko and Frédéric Robert-Nicoud. 2005. The role of human capital investments in the location decision of firms. *Regional Science and Urban Economics* 35(5):570–583.
- Mion, Giordano and Paolo Naticchioni. 2009. The spatial sorting and matching of skills and firms. *Canadian Journal of Economics* 42(1):683–701.
- Moulton, Brent R. 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *Review of Economics and Statistics* 72(2):334–338.
- Nocke, Volker. 2006. A gap for me: Entrepreneurs and entry. *Journal of the European Economic Association* (forthcoming).
- Organisation for Economic Co-operation and Development . 2002. *Redefining territories. The functional regions*. Paris: OECD.
- Overman, Henry G. and Diego Puga. 2010. Labor pooling as a source of agglomeration: An empirical investigation. In Edward L. Glaeser (ed.) *The Economics of Agglomeration*. Cambridge (MASS.): National Bureau of Economic Research, 133–150.
- Picard, Pierre M. and David E. Wildasin. 2011. Outsourcing, labor market pooling, and labor contracts. *Journal of Urban Economics* 70(3):47–60.
- Puga, Diego. 2010. The magnitude and causes of agglomeration economies. *Journal of Regional Science* 50(1):203–219.
- Rosenthal, Stuart S. and William C. Strange. 2004. Evidence on the nature and sources of agglomeration economies. In J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics*, volume 4. Amsterdam: North-Holland, 2119–2171.
- Rosenthal, Stuart S. and William C. Strange. 2010. Small establishments/big effects: Agglomeration, industrial organization and entrepreneurship. In Edward L. Glaeser (ed.) *The Economics of Agglomeration*. Cambridge (MASS.): National Bureau of Economic Research, 277–302.
- Saraceno, Chiara. 1994. The ambivalent familism of the Italian welfare state. *Social Politics* 1(1):60–82.
- Shimer, Robert. 2005. The cyclicalities of hires, separations, and job-to-job transitions. *Federal Reserve Bank of St. Louis Review* 87(4):493–507.
- Tarantola, Anna Maria. 2011. Presentazione del rapporto l'economia delle marche nel 2010. Transcription of a public intervention in ancone (13 june 2011), Banca d'Italia.
- Venables, Anthony J. 2011. Productivity in cities: Self-selection and sorting. *Journal of Economic Geography* 11(2):241–251.
- Wheeler, Christopher H. 2006. Cities and the growth of wages among young workers: Evidence from the NLSY. *Journal of Urban Economics* 60(2):162–184.
- Wheeler, Christopher H. 2008. Local market scale and the pattern of job changes among young men. *Regional Science and Urban Economics* 38(2):101–118.

Appendix A. The Bank of Italy's Surveys

The Survey of Household Income and Wealth (SHIW)

This survey is conducted every 2 years by the Bank of Italy on about 8,000 households (24,000 individuals), distributed over about 300 Italian municipalities. The SHIW gathers information on income, savings, wealth and other socio-economic indicators.

The questionnaire for the 2006 wave (including its special section on local labor markets) can be downloaded at: http://www.bancaditalia.it/statistiche/indcamp/bilfait/docum/ind_06/Questionario/Quest_ing2006.pdf

Interviews are carried out by external professional interviewers. For the 2006 wave, details on methodology (sample design, questionnaire and data collection, data editing and imputation, non response, data quality, etc) are provided at: http://www.bancaditalia.it/statistiche/indcamp/bilfait/boll_stat/en_suppl07_08.pdf

The survey results are regularly published in the Bank of Italy's Reports. The data is freely available in an anonymous form for further elaboration and research. A full list of academic paper based on SHIW) data is available. Details can be found at: <http://www.bancaditalia.it/statistiche/indcamp/bilfait>.

The Survey on Industrial and Service Firms (SISF)

This survey is conducted annually by the Bank of Italy on about 3,000 industrial firms, 465 construction companies and 1,083 non-financial private service firms (representing 8.1%, 6.5% and 3.8% of their respective total reference populations). The SISF gathers information on status, organization, performance, and other economic indicators. The survey results are regularly published in the Bank of Italy's Reports. The data can be freely accessed, through the Remote Processing System BIRD, for further elaboration and research. A full list of academic paper based on SISF data is available. Details can be found at: <http://www.bancaditalia.it/statistiche/indcamp/indimpser>.

The questionnaire of the 2007 wave (including its special section on local labor markets) can be downloaded at: http://www.bancaditalia.it/statistiche/indcamp/indimpser/boll_stat/sb42_08/en_suppl_42_08.pdf

Interviews are carried out by Bank of Italy's employees (mostly by economists). The respondent is usually either the owner of the firm or a member of its top management, except for very large firms. Details on the methodology (sample design; data collection, questionnaire and response behaviour; data quality, checks and imputation of missing data; etc) are provided for the 2007 wave at: http://www.bancaditalia.it/statistiche/indcamp/indimpser/boll_stat/sb42_08/en_suppl_42_08.pdf.

Appendix B. List of variables

Dependent variables from SHIW

H1 Wages.

Log of hourly wages in Euro. Hourly wages are calculated by dividing the annual earnings (from any activity as payroll employee or 'fake' self-employed (see below for more details on the issue of fake self-employed), including fringe benefits, net of taxes and social security contributions) by the total amount of hours worked in a year (average hours worked per week \times months worked \times 4.3333).

H2 Change of employer or type of work.

Dummy variable that equals one if the worker changed employer or type of work in the last two years. It is taken from answers to the question "Have you changed employer or type of work in the last two years?" (question numbered R2.7 in the questionnaire).

H3 Change of employer but not of type of work.

Dummy variable that equals one if the worker changed employer but not type of occupation. It is taken from the question "What have you changed? Employer, type of work, or both?" (question R2.9).

H4 Learning.

Dummy variable that equals one if the worker acquired her skills informally from colleagues inside or outside the firm. It is taken from the question "Last year, by which of the following means did you acquire skills to improve your job performance?" (question R2.15).

H5 Useful past experience.

Dummy variable that equals one if the worker's previous experience in the same field is useful for the job held at the survey date. It is taken from answers to the question "Did you gain your previous experience in the same field you work in now?" (question R2.20).

H6 Training by firm.

Dummy variable that equals one if the worker received training by the firm. It is taken from the question "Last year, by which of the following means did you acquire skills to improve your job performance?" (question R2.15).

H7 Skill transferability.

Dummy variable that equals one if the worker's skills are totally or partially transferrable. It is taken from the question "If you were to leave your present employer, could the skills you have acquired be used in another job?" (question R2.16).

H8 Difficulty of finding a replacement by employer.

Dummy variable that equals one if the replacement of a worker is very difficult. It is taken from answers to the question "If you were to leave your job, how difficult/easy would it be for your employer to find a replacement (on a scale from 1 to 10)?" (question R2.17). The questionnaire variable has been re-scaled to ease interpretation. The re-scaled variable goes from 1 (very easy) to 10 (very difficult). The dummy variable is equal to 1 if the re-scaled variable takes a value higher than 9.

H9 Difficulty of finding an equivalent job.

Dummy variable that equals one if finding a new similar job in terms of salary or overall quality is very difficult. It is taken from answers to the question "If you were to lose your job, how difficult/easy would it be for you to find a similar job in terms of salary and overall quality (on a scale from 1 to 10)?" (question R2.14). The questionnaire variable has been re-scaled to ease interpretation. The re-scaled variable goes from 1 (very easy) to 10 (very difficult). The dummy variable is equal to 1 if the re-scaled variable takes a value higher than 9.

H10 Specialization.

Dummy variable that equals one if the worker judges her level of specialization as very high. It is taken from answers to the question "Comparing yourself with other people in Italy who perform the same job, how specialized is your work (on a scale from 1 to 10)?" (question R2.21). The questionnaire goes from 1 (not at all specialized) to 10 (very specialized). The dummy is equal to 1 if the questionnaire variable takes a value higher than 9.

H11 Appropriate experience.

Dummy variable that equals one if the worker's experience is appropriate for the employer requests. It is taken from answers to the question "In your opinion, does your job demand more work experience than you have, less work experience, the same amount of work experience?" (question R2.19).

H12 Appropriate skills.

Dummy variable that equals one if the worker's educational qualification is appropriate for the job. It is taken from answers to the question "Do you think your educational qualification is appropriate for the job you do?" (question R2.17).

*Dependent variables from SISF***F1 Output per worker.**

Log of the ratio between firm revenue in thousands of Euro (variable name: V210) and average workforce (variable name: V34) (see Bugamelli, Cristadoro, and Zevi, 2009, for further discussion).

F2 Share of terminations voluntary.

Share of terminations due to voluntary resignations (variable name: OCC2).

F3 Share of vacancies filled from the same sector.

Share of vacancies filled in by workers with previous experience in the same sector (variable name: OCC1).

F4 Number of days to train a key worker.

Number of days of formal training received on average by the firm's key worker (variable name: OCC6).

F5 Appropriate education and experience of new key workers.

Dummy variable that equals one if the worker's experience and education are enough for the job. It is taken from the question "Do you consider that, on average, your key workers are suitable for the tasks required from them?"; variable name: OCC5).

Explanatory variables

Males (SHIW).

Dummy variable that equals one for males.

Education (SHIW).

Number of years of studies required to achieve the highest qualification earned by the worker. The length of education is derived by assigning: 2 years to no qualification; 5 years to elementary school; 8 years to middle school; 16 years to an associate degree or other short course university degree; 18 years to a bachelor's degree; 20 years to a postgraduate qualification.

Experience (SHIW).

Difference between worker's age at the survey date and the age at first job held, which is available from the SHIW.

South (SHIW and SISF).

Dummy variable that equals one for residence in the South of Italy. South of Italy includes Abruzzi, Molise, Campania, Puglia, Basilicata, Calabria, Sicily, and Sardinia.

Density (SHIW and SISF).

Log of population density. Density is computed as the ration between population and area (km²) in 2001.

Industrial district (SHIW and SISF).

Dummy variable that equals one for industrial districts.

Age (SISF).

Age of the firm at the survey date.

Status: limited liability (SISF).

Dummy variable that equals one for limited liability firms.

Group (SISF).

Dummy variable that equals one if the firm is part of a group, i.e., a set of firms directly or indirectly controlled - through one or more chains of control by the same legal persons or the same public entity.

Manufacturing (SISF).

Dummy variable that equals one if the firm belongs to the manufacturing sector.

Appendix C. Data issues

Fake self-employed

A potential issue with the sample of workers is the presence of fake self-employed in the labor market. For tax reasons and taking advantage of loopholes in labor market regulations, a number of workers that are registered as self-employed are in fact payroll employees.

It could be that the presence of fake self-employed is higher in denser areas (therefore, limiting our sample to registered payroll employees might bias our results). In the 2006 SHIW questionnaire we introduced three questions to identify fake self-employed (see questions: R2.4, R2.5, and

R2.6). Basically, self-employed workers that i) work for just one firm/client; ii) at the firm/client's premises; and iii) observing the same working hours as the regular employees of their firm/client are taken to be fake self-employed and we treat them as regular payroll workers.

In questions H2, H3, H4, H6, H7, H8, and H9 we use answers from the sample of payroll employees and fake self-employed. In questions H10, H11, and H12 we use answers from all working individuals.

Key workers

A potential issue with the answers from SISF is that some questions need to distinguish between different types of workers. As firms employ different typologies to classify workers, we decided to identify those whom managers or owners believe make a significant difference to product quality or to competitiveness. This group is labelled 'key workers' and it is defined as workers. Questions F4 and F5 refer to key workers.

Sample sizes

For SHIW, we consider only workers aged between 24 and 60 and delete workers with a log wage above 3.92 or below 0.14 (corresponding to 1% extreme values). The sample size differs by question. Wages (H1) are constructed on all employed persons independently of the fact that they answered the special section questions (number of observations: 4,367). Questions on turnover (H2), learning (H4), training by the firm (H6), skill transferability (H7), difficulty of finding a replacement by the employer (H8) and difficulty of finding an equivalent job (H9) are asked to payroll employees and fake self-employed (number of observations: 1,287). Questions on appropriate experience (H11) and appropriate education (H12) are asked to all working individuals (number of observations: 1,606). Questions on useful past experience (H5) are asked to all working individuals who had more than one job in their lifetime (number of observations: 945). Questions on the change of employer (H3) are asked only to people who changed job in the last two years (number of observations: 117).

For SISF we delete observations with extreme values of log output per worker below 3.30 and above 7.38. We end up with 3660 observation. The number of observations fluctuates between 2452 and 2946 when we focus on local labor market variables.

Appendix D. Territorial units of reference

Local labor markets (or 'cities')

Local labor markets are defined by the Italian National Institute of Statistic (ISTAT, 1997). They are aggregations of two or more neighbouring municipalities based on daily commuting flows from place of residence to place of work as recorded in the 2001 Population Census. Local labor markets are thus largely 'self-contained': within a given unit, both the share of working residents working locally and the share of employees residing locally must be at least 75%.

This definition is consistent with standard definitions of cities in urban economics that define them through commuting patterns. In much of text we thus refer to these spatial units as cities. This definition is also consistent with the notion of ‘functional region’, defined as ‘a territorial unit resulting from the organization of social and economic relations in that its boundaries do not reflect geographical particularities or historical events’ (OECD, 2002). Italian local labor market also roughly follow the criteria used to define Metropolitan Statistical Areas in the us, Travel to Work Areas in the UK, or Metropolitan areas and employment areas in France.

Italian local labor markets span the entire national territory. In 2001, 686 of them were defined. They had an average population of 83,084 and a standard deviation of 222,418.

Industrial districts (IDs)

Through the ISTAT Cluster Mapping Project (ICMP) 156 cities (out of 686) are identified as IDs. Basically, IDs are cities with a prevailing specialization and a higher concentration of employment in small-sized manufacturing firms. To identify IDs, the ICMP uses four criteria (which all have to be met): (i) The share of manufacturing employment in total (non-farm) employment must be higher than the corresponding share at the national level. (ii) The share of small and medium enterprise manufacturing employment in total (non-farm) employment must be higher than the corresponding share at the national level. (iii) For at least one sector, the specialization index (the ratio between the share of sector employment in total manufacturing employment and the corresponding share at the national level) must be greater than one. (iv) In at least one sector for which the specialization index is greater than one, the share of small and medium enterprise employment in total employment must be higher than the corresponding share at the national level.