# Marshallian labour market pooling: Evidence from Italy ☆

Monica Andini [a,*], Guido de Blasio [b], Gilles Duranton [c,d,e,f], William C. Strange [g]

[a] Bank of Italy, Branch of Napoli, Via Miguel Cervantes 71, 80133 Napoli, Italy
[b] Bank of Italy, Structural Economic Analysis Department, Via Nazionale 91, 00184 Roma, Italy
[c] Wharton School, University of Pennsylvania, 3620 Locust Walk, Philadelphia, PA 19104, USA
[d] Centre for Economic Policy Research, 77 Bastwick Street, London EC1V 3PZ, United Kingdom
[e] The Rimini Centre for Economic Analysis, Via Patara 3, 47900 Rimini, Italy
[f] Spatial Economics Research Centre, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, United Kingdom
[g] Rotman School of Management, 105 Saint George Street, Toronto, Ontario M5S 3E6, Canada

## ARTICLE INFO

## ABSTRACT

This paper employs a unique Italian data source to take a comprehensive approach to labour market pooling. It jointly considers many different aspects of the agglomeration — labour market relationship, including turnover, learning, matching, and hold up. It also considers labour market pooling from the perspective of both workers and firms and across a range of industries. Overall, the paper finds some support for theories of labour market pooling, but the support is weak. Specifically, there is a general positive relationship of turnover to local population density, which is consistent with theories of agglomeration and uncertainty. There is also evidence of on-the-job learning that is consistent with theories of labour pooling, labour poaching, and hold up. In addition, the paper provides evidence consistent with agglomeration improving job matches. However, the labour market pooling gains that we measure are small in magnitude and seem unlikely to account for a substantial share of the agglomeration benefits accruing to Italian workers and firms.

© 2013 Elsevier B.V. All rights reserved.

## 1. Introduction

As with most economic research on urban labour 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 – labour market pooling. The latter refers to the advantages for workers and firms deriving from sharing a labour market that is territorially limited to a small area: the local labour market. For instance, in a thicker local labour market workers might be able to find a job faster. Similarly, firms might fill vacancies faster. In addition, firms and workers are likely to find better matches 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 labour market pooling. The paper looks across all industries from the perspectives of both workers and firms, and it considers many different aspects of labour market pooling, including turnover, matching, hold up and learning. To our knowledge,

---

this is the first time that 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 labour 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 these data with data from the Italian National Institute of Statistics to assess the thickness of the labour 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 labour 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 works on agglomeration. There is consistent evidence of an urban wage premium. In addition, firm output per worker is positively related to population density.

The labour market pooling results that we find are, when taken as a whole, rather restrained in their support for the various sorts of labour market pooling that appear in 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 labour pooling, labour 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 labour market pooling.

There are several ways that one might interpret the modest magnitudes of our labour market pooling results. 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 labour 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 labour market pooling operates differently across different industries. For instance, it is common to consider the relationship between agglomeration and turnover for the computer industry. 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 that 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, looking across industries, 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 discusses the relevant literature and how our analysis arises from it. Section 3 presents the details of the paper's data sources. Section 4 includes the results of the estimates of the agglomeration–wage and agglomeration–productivity relationship. Section 5 contains the estimates of the relationship between agglomeration and turnover, learning, matching, and other aspects of labour market pooling. Section 6 assesses the importance of our measures of labour market pooling in the agglomeration–wage and agglomeration–productivity relationship. Section 7 concludes.

## 2. Literature

Marshall's insights have motivated a long line of research on labour market pooling as a microfoundation for agglomeration economies. This section reviews the theoretical and empirical contributions of the literature and shows how our analysis arises from it.

Theoretical research on labour 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 labour demands can generate an agglomeration economy. Strange et al. (2006) demonstrate that the firms who face greater difficulty in matching will locate in thick markets. 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 et al. (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 the 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 labour pooling include Gerlach et al. (2009), who consider the interaction between labour pooling and innovation, and Picard and Wildasin (2011), who consider the interaction with input sharing. A survey of the larger microfoundations literature, including labour market pooling, can be found in Duranton and Puga (2004).

The empirical literature on labour 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 et al. (2010) find that firms drawing from the same sorts of labour pool tend to coagglomerate. Jofre-Monseny et al. (2011) carry out a similar exercise and also find evidence consistent with labour market pooling.[1]

There is also a smaller but growing empirical literature that has looked specifically at labour market pooling. Papers in this literature have uncovered a number of instances where Marshallian labour market pooling seems to be at work. Fallick et al. (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. Looking across US industries, Bleakley and Lin (2012) show that workers change occupation and industry less frequently when population density is greater. With regard to matching, Andersson et al. (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. Using Canadian survey data, Strange et al. (2006) show that skill-oriented firms tend to choose locations with concentrations of activity in their own industry rather than locations with concentrations of

---

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

aggregate activity. Glaeser and Maré (2001), Wheeler (2006), and De la Roca and Puga (2012) all provide evidence that the urban wage premium rises with a worker tenure in a city, a finding consistent with learning. Bacolod et al. (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 labour market pooling reducing risk.

Our analysis builds naturally on the literature. First, we will determine if wages and other productivity measures are positively associated with agglomeration in our data. This is a necessary initial step, since looking for evidence of different sorts of labour market pooling is not likely to be fruitful unless there is some benefit from agglomeration that could potentially arise from labour market pooling. Of course, an observed relationship between agglomeration and wages or productivity could reflect any combination of agglomeration spillovers or sorting or selection. We will, therefore, estimate instrumental variable models in order to focus on the former. Second, we will ask whether turnover is greater in thicker markets. Third, we will consider the relationship between agglomeration and learning, bearing in mind that knowledge spillovers encourage learning, while holdup has the opposite effect. Fourth, we will also consider the relationship of agglomeration to matching, again bearing in mind the ambiguous relationship predicted by theory, 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.

## 3. 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.

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 labour 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 labour markets. The household survey contains 12 questions about the working of local labour markets and the firm survey contains another 5.[2] 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 labour markets while matching these outcomes with information about local labour markets.

Our sample includes 4367 workers (excluding government employees) and 3660 firms. Appendix C documents a number of data issues. In particular, the questions pertaining to labour 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 labour 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 primary 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 labour markets are organized around three themes:

### 3.1. Turnover and flexibility

To assess whether denser markets are associated with greater labour 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 labour market pooling is voluntary. We therefore measure turnover by the percentage of terminations due to voluntary 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.[3] 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 is voluntary. All these paint a picture with relatively modest labour market fluctuations. We return to this below.

### 3.2. 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 says that they have learned from informal contacts within the firm. We return to this issue below as well.

### 3.3. Matching

The theories discussed above establish an ambiguous relationship between agglomeration and a worker's risk of finding a job and a firm's risk of filling a vacancy. To examine this relationship, we ask workers to assess the ease of replacement faced by their employer, should the worker quit (H8). We also ask workers how easy it would be for them to find another job similar in terms of salary or overall quality, should they lose their current job (H9).[4] However, 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

---

[2] The original design of the local labour market sections of these two surveys contained more questions. For instance, questions on labour 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.

[3] See Shimer (2005) for US figures.

[4] These two questions do not allow distinguishing between effects of a sectoral shock and those of an idiosyncratic shock.

**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 | 0.26 | 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 Household Income and Wealth 2006 and Survey on Industrial and Service Firms 2007.

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 labour 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 the worker level, our controls include gender, education, experience, and its square. In a robustness check, we make use of confidential data on worker 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).

## 4. Agglomeration

Our interest in labour market pooling arises from its role in the generation of agglomeration economies. Before turning to labour 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 per worker.

Our main estimating equation is:

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

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. Finally, $\epsilon_i$ 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).[6]

Table 2 reports results for eight wage regressions. In column 1, we regress log wages on the log of city density alone.[7] 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 individual controls: a male dummy, years of education, labour market experience, and its square. In columns

---

[5] Sample 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.

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

[7] We use density rather than total population because this variable is more robust to the way boundaries are drawn. In particular, municipalities are part of the same local labour market only when both the share of working residents working locally and the share of employees residing locally are 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 dimensions. Our use of density should also ease comparisons since density is used more often than population in the recent literature (see Combes et al., 2011 for a longer discussion of these issues). In particular density is less sensitive to the modifiable areal unit problem (MAUP) than population.

**Table 2**
Wage equations.

| | (1) Raw | (2) Geog | (3) Geog + Indiv | (4) Males | (5) Old | (6) HighEd | (7) Movers | (8) IV |
|---|---|---|---|---|---|---|---|---|
| Log density | 0.040[a] | 0.035[a] | 0.027[a] | 0.026[a] | 0.035[a] | 0.049[b] | 0.027[b] | 0.022[a] |
| | (0.015) | (0.0081) | (0.0066) | (0.0078) | (0.0080) | (0.023) | (0.012) | (0.0085) |
| Indus. district | | −0.028 | −0.0016 | −0.0038 | −0.0048 | 0.067 | 0.0096 | 0.0024 |
| | | (0.021) | (0.018) | (0.022) | (0.024) | (0.053) | (0.026) | (0.018) |
| South | | −0.18[a] | −0.14[a] | −0.14[a] | −0.13[a] | −0.16[b] | −0.15[a] | −0.13[a] |
| | | (0.026) | (0.024) | (0.025) | (0.027) | (0.071) | (0.046) | (0.024) |
| Male | | | 0.15[a] | | 0.18[a] | 0.14[a] | 0.19[a] | 0.15[a] |
| | | | (0.012) | | (0.019) | (0.045) | (0.018) | (0.012) |
| Education | | | 0.041[a] | 0.043[a] | 0.044[a] | 0.059 | 0.042[a] | 0.041[a] |
| | | | (0.0024) | (0.0030) | (0.0035) | (0.036) | (0.0038) | (0.0025) |
| Experience | | | 0.026[a] | 0.027[a] | 0.017[b] | 0.045[a] | 0.022[a] | 0.026[a] |
| | | | (0.0024) | (0.0032) | (0.0076) | (0.0090) | (0.0040) | (0.0025) |
| Experience2 ($\times$100) | | | −0.038[a] | −0.037[a] | −0.024[c] | −0.073[a] | −0.033[a] | −0.037[a] |
| | | | (0.0055) | (0.0069) | (0.013) | (0.027) | (0.0091) | (0.0056) |
| Observations | 4367 | 4367 | 4367 | 2804 | 2293 | 398 | 1162 | 4252 |
| $R^2$ | 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.

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 simultaneously determined since we expect places that pay higher wages to be more densely populated.

Following earlier literature (e.g. Ciccone and Hall (1996) and Combes et al. (2010)), we instrument contemporaneous density with long population density lags, namely log 1871 and log 1921 population density. In our context, these density lags provide strong predictors for contemporaneous density. To be valid these instruments need to be correlated with the outcomes of interest only through contemporaneous density. The argument that 'residual productivity' (i.e., the error term in the wage regression) should be uncorrelated with our instruments relies on the fundamental changes that affected the Italian economy since 1871 and 1921. These include the two world wars, the fascist dictatorship, and the more general transformation from a largely rural economy at the start of the 20th Century. This has resulted in much more pronounced local differences. For instance, the North–South divide was not so large. Other research has also shown that these historical instruments yield results similar to alternative instruments based on geology (Combes et al., 2010) and are robust to the inclusion of many local characteristics. 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). We also note that, despite its possible weaknesses, using a standard approach to the identification of agglomeration allows us to compare our results more easily.

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 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 coefficients, 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.[8]

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:

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

and our dependent variable $Y_j$ is now the log of output per worker of firm $j$, $B_{c(j)}$ 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 reports results for the estimation of Eq. (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 consider are: age, a dummy variable for limited liability status, and a dummy variable for being part of a group.[9] 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 are 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.

---

[8] The 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) and Dalmazzo and de Blasio (2011), for works 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 Anna Maria (2011)).

[9] We do not include firm size since it could be a consequence of being more productive.

**Table 3**
Output per worker.

| | (1) Raw | (2) Geog | (3) Geog + Firm | (4) Old | (5) Group | (6) SME | (7) Manuf | (8) IV |
|---|---|---|---|---|---|---|---|---|
| Log density | 0.062[c] | 0.065[a] | 0.036[c] | 0.052[b] | 0.030 | 0.057[a] | 0.053[a] | 0.033[c] |
| | (0.033) | (0.022) | (0.019) | (0.025) | (0.028) | (0.017) | (0.013) | (0.019) |
| Indus. district | | 0.024 | 0.034 | 0.012 | 0.057 | 0.0012 | −0.020 | 0.038 |
| | | (0.039) | (0.037) | (0.063) | (0.050) | (0.042) | (0.036) | (0.039) |
| South | | −0.26[a] | −0.17[a] | −0.098 | −0.17[a] | −0.20[a] | −0.24[a] | −0.17[a] |
| | | (0.044) | (0.041) | (0.081) | (0.065) | (0.042) | (0.035) | (0.042) |
| Age | | | 0.0016[a] | 6.0e−07 | 0.0012[c] | 0.0015[b] | 0.0015[a] | 0.0014[b] |
| | | | (0.00057) | (0.00096) | (0.00066) | (0.00071) | (0.00050) | (0.00057) |
| Limited liability | | | 0.35[a] | 0.039 | 0.48[a] | 0.15 | −0.18[c] | 0.38[a] |
| | | | (0.099) | (0.14) | (0.15) | (0.10) | (0.094) | (0.099) |
| Group | | | 0.29[a] | 0.31[a] | | 0.30[a] | 0.37[a] | 0.30[a] |
| | | | (0.026) | (0.058) | | (0.031) | (0.030) | (0.027) |
| Observations | 3660 | 3660 | 3649 | 908 | 1518 | 2777 | 2645 | 3551 |
| R² | 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.

In a separate web appendix (Andini et al., 2013), we duplicate the regressions of Tables 2 and 3 using a full set of regional dummies (for NUTS 1 and NUTS 2 regions) and alternative measures of density (employment density and manufacturing employment density). We also duplicate the regressions of Table 3 replacing the dummy for being in an industrial district with a dummy for being in an industrial district with the same sectoral specialization as the firm. The results are generally very similar to those of Tables 2 and 3. These results are unsurprising given that most of the variation in log density takes place within regions rather than between regions and that employment and population density are highly correlated. For instance the correlation between log population density and log employment density is 0.95.

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., Bugamelli et al. (2009) and Glaeser and Resseger, (2010)), the estimated elasticity of wages with respect to density is slightly lower. It is about one to 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.[10] Overall we take these magnitudes as very close given differences in the data being used and differences in the estimation. Many detailed aspects of the wage findings are also in line with existing results.[11]

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).[12]

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 labour market conditions.[13]

Simple theoretical considerations as suggested by the Roback (1982) model imply 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 et al., 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.[14]

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) and 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 the same industry to capture them.[15] We failed to find robust results supportive of agglomeration effects taking place 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 et al. (2008) for a detailed comparison between these two types of agglomeration effects and the robustness of urbanisation effects to detailed sectoral controls).

---

[10] If 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.

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

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

[13] For instance, the IMF recently mentioned that: "Italy's wage setting system is an obstacle for growth" (International Monetary Fund, 2011).

[14] Clearly, 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 is 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 labour 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 et al., 2001) and public sector jobs (Alesina et al., 2001) deployed in lagging (Southern) regions in the last decades.

[15] For instance, including a dummy variable for manufacturing does not change anything in our results.

**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 | −0.0035 | −0.0027 | −0.0087 | −0.0042 | 0.025 | −0.019 | 0.0077 |
| | (0.0088) | (0.0091) | (0.0092) | (0.011) | (0.0083) | (0.023) | (0.016) | (0.011) |
| Observations | 1287 | 1287 | 1287 | 806 | 721 | 100 | 383 | 1251 |
| 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] | 0.12[b] | 0.13[b] | 0.10 | 0.25[b] | −0.00062 | 0.043 | 0.14[b] |
| | (0.058) | (0.057) | (0.060) | (0.070) | (0.11) | (0.0020) | (0.076) | (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 | – |
| | Raw | Geog | Geog + Firm | Old | Group | SME | Manuf | IV |
| *Panel C. Dependent variable: Firm share of terminations voluntary* | | | | | | | | |
| (Question F2), OLS and TSLS | | | | | | | | |
| Log density | 0.030[a] | 0.034[a] | 0.035[a] | 0.042[a] | 0.027[a] | 0.041[a] | 0.023[a] | 0.050[a] |
| | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| Observations | 2750 | 2750 | 2743 | 691 | 1165 | 2066 | 1987 | 2666 |
| 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.

## 5. Labour market pooling

### 5.1. Overview

In this section, we take advantage of the richness of the two surveys to explore the manifestations of labour 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 labour 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:

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

This specification mirrors regression (1) but considers, as dependent variable $Z_i$, a measure of labour market pooling for worker $i$. We consider a number of aspects of labour 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 primarily concerned with the relationship of the labour market pooling variables with agglomeration but, in estimating Eq. (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 labour 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. For binary dependent variables, estimation is carried out by probit and IV probit. We report marginal effects at the variable mean for continuous variables. For dummies, marginal effects are computed for the change of the dummy variable from 0 to 1. In a separate web appendix (Andini et al., 2013), we also report results from linear probability models.

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

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

This specification mirrors regression (2) but considers, as dependent variable $Z_j$, a measure of labour market pooling for firm $j$. These include the share of terminations that are voluntary (F2), the share of vacancies filled from workers previously employed in the same sector (F3), the numbers of days to train key workers (F4), and the appropriateness of a new worker in terms of education and experience (F5).

In estimating Eq. (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 includes 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.

### 5.2. Turnover

Table 4 presents results that relate to the turnover element of labour 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 labour market pooling, either positively or negatively.[16]

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. Importantly, we note that this lack of

---

[16] This 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 labour 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).

significance is not due to large standard errors. Instead the coefficients are precisely estimated. In all columns of panel A of Table 4 we can rule out that the marginal effect on worker job changes associated with density is 3% or more at conventional levels of significance. Except in column 6 where the number of observations is small, we can even rule out a coefficient of 2%. That is, our measured effect of density on the probability of changing jobs is lower than our measured effect of density on wages or output per worker. Put differently, if an average worker has a probability of 9% of changing jobs over a two-year period as indicated in Table 1, in a labour market twice as dense, our results indicate that this probability will remain below 9.13%.

These results are confirmed using linear probability models instead of probit for binary dependent variables (Andini et al., 2013). 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 types of employment. 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 labour market pooling in the spirit of Marshall. Similar results are obtained when adding a full set of dummies for NUTS 2 regions (Andini et al., 2013).

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 was voluntary. Together, these results mean that although Marshallian turnover 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., Faini et al. (1997) and Fallick et al. (2006)), 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.

### 5.3. Learning and holdup

Table 5 presents results that relate to the learning element of labour 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 spillovers. 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 consistent 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 1. 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 workers 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 labour 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 "labour 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. As with panel A of Table 4, we are again in the case of precisely estimated zeroes rather than coefficients lacking precisions. We also note that the results of Panels A–D are confirmed using linear probability models instead of probit (Andini et al., 2013).

The firm results presented in the last two panels of the table are noisier, but they ultimately tell a 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 labour 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 hinge on the difference between "key" workers and the rest of the workforce.

Duplicating these results with a full set of dummies for NUTS 2 regions in Andini et al. (2013) only makes small differences. The results of Panels A–C are slightly weaker whereas in Panels D and E, the coefficient on density is positive and significant in a majority of columns.

### 5.4. Matching

We now discuss the matching aspect of labour market pooling. As discussed above, while one expects matches to be better in thick markets, one also expects jobs where matching is difficult to be found in cities. The empirical relationship between matching and agglomeration will be a combination of these two effects.

Some of the previous results can be interpreted as bearing on this relationship. In particular, the results on the relationship of agglomeration to useful past experience in Panel B of Table 5 clearly bear on both matching and learning. The results show a robust positive relationship between a worker's useful experience and density. This is consistent with the idea that agglomeration improves matches but also with the idea that density leads to specialization.

Table 6 presents further results on the matching. Panel A presents results on a worker's assessment of the difficulty an employer is expected 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. Again, this is consistent with agglomeration improving matches.

The rest of the results is much weaker. The other worker estimates feature signs that are largely consistent with matching but with small coefficients. And the standard errors are small, suggesting that the

**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] | 0.0052 | 0.0047 | 0.0072[a] | 0.00063 | 0 | 0.000088 | 0.0071[c] |
| | (0.0041) | (0.0032) | (0.0030) | (0.0027) | (0.0048) | (0) | (0.0045) | (0.0041) |
| Observations | 1287 | 1287 | 1287 | 806 | 543 | 55 | 214 | 1251 |
| 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] | 0.076[a] | 0.079[a] | 0.073[a] | 0.061[a] | 0.096[c] | 0.095[b] | 0.079[a] |
| | (0.018) | (0.018) | (0.018) | (0.021) | (0.018) | (0.055) | (0.033) | (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 | −0.028[c] | −0.031[c] | −0.033[b] | −0.025 | −0.081 | −0.033 | −0.034 |
| | (0.020) | (0.017) | (0.017) | (0.015) | (0.022) | (0.069) | (0.028) | (0.021) |
| Observations | 1287 | 1287 | 1287 | 806 | 721 | 100 | 383 | 1251 |
| 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 | 0.010 | 0.011 | 0.0077 | 0.012 | −0.00040 | −0.0089 | 0.020 |
| | (0.015) | (0.015) | (0.014) | (0.013) | (0.014) | (0.030) | (0.020) | (0.013) |
| Observations | 1287 | 1287 | 1287 | 806 | 721 | 100 | 383 | 1251 |
| Pseudo-$R^2$ | 0.00 | 0.02 | 0.05 | 0.04 | 0.04 | 0.03 | 0.07 | – |
| | Raw | Geog | Geog + Firm | Old | Group | SME | Manuf | IV |
| *Panel E. Dependent variable: Firm share of vacancies filled by workers with same sector experience* (Question F3), OLS and TSLS | | | | | | | | |
| Log density | 0.010 | 0.0081 | 0.011 | 0.0053 | 0.026[b] | 0.0074 | 0.018 | 0.0060 |
| | (0.01) | (0.01) | (0.01) | (0.02) | (0.01) | (0.01) | (0.02) | (0.01) |
| Observations | 2452 | 2452 | 2443 | 604 | 1014 | 1878 | 1767 | 2376 |
| $R^2$ | 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 | 0.32 | 0.36 | 0.97 | 0.38 | 0.81 | −0.14 | 0.45 |
| | (0.48) | (0.49) | (0.49) | (1.01) | (0.56) | (0.53) | (0.56) | (0.64) |
| Observations | 2553 | 2553 | 2548 | 614 | 1060 | 1955 | 1839 | 2484 |
| $R^2$ | 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 Tables 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.

problem is not simply one associated with sample size. 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 insignificant coefficients 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 again small and insignificant.

There are two ways that these results can be taken.[17] One possibility is that in the Italian context, matching is not a broadly important source of agglomeration economies. This conflicts with other evidence suggesting that matching is important, such as Bleakley and Lin (2012), who use US data. The other possible interpretation is that the positive effect of

matching is obscured, at least somewhat, by the sorting of firms with difficult matches into high-density areas.

## 6. Labour market pooling and agglomeration

In this section we return to the agglomeration–wage relationship and to the agglomeration–productivity explored in Section 4 but consider density and our measures of labour 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.

More specifically, we expect the final outcomes, log wage or log output per worker, which we denote $Y$ (to remain consistent with our notations so far) to be a function of city characteristics $B_c$ (in particular log density) either through local labour market variables $Z$ or through a host of other channels such as technological spillovers or input–output linkages. That is, we expect a relationship of the following form:

$$Y = f(Z(B_c), B_c), \tag{5}$$

where the first argument of the function $f(.)$ captures the effect of urban agglomeration percolating through local labour markets and the second argument captures the effects of urban agglomeration

---

[17] Again, these findings are confirmed when using linear probability models instead of probit or when adding a full set of dummies for NUTS 2 regions (Andini et al., 2013). In the latter case we see a small decline in significance for the coefficient on log density in Panel A but some significant coefficients in Panels B and C.

**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] | −0.017[c] | −0.018[b] | −0.025[b] | −0.012 | 0.020 | −0.029 | −0.028[b] |
| | (0.010) | (0.0093) | (0.0091) | (0.010) | (0.0099) | (0.016) | (0.019) | (0.012) |
| Observations | 1287 | 1287 | 1287 | 806 | 721 | 100 | 383 | 1251 |
| 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 | −0.0093 | −0.013 | −0.017 | −0.0095 | −0.010 | −0.032 | −0.022 |
| | (0.012) | (0.012) | (0.012) | (0.013) | (0.015) | (0.034) | (0.020) | (0.016) |
| Observations | 1287 | 1287 | 1287 | 806 | 721 | 100 | 383 | 1251 |
| Pseudo-$R^2$ | 0.00 | 0.00 | 0.03 | 0.03 | 0.02 | 0.14 | 0.09 | – |
| *Panel C. Dependent variable: Worker specialization* | | | | | | | | |
| (Question H10), probit and IV probit | | | | | | | | |
| Log density | −0.0020 | −0.0049 | −0.0066 | −0.0062 | −0.0012 | 0.0055 | −0.0041 | −0.0038 |
| | (0.0076) | (0.0076) | (0.0074) | (0.0085) | (0.0073) | (0.031) | (0.015) | (0.0097) |
| Observations | 1606 | 1606 | 1606 | 1015 | 914 | 140 | 458 | 1558 |
| 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 | 0.0086 | 0.0087 | 0.0092 | 0.0011 | −0.010 | 0.023 | 0.0033 |
| | (0.012) | (0.012) | (0.012) | (0.011) | (0.015) | (0.026) | (0.023) | (0.014) |
| Observations | 1606 | 1606 | 1606 | 1015 | 914 | 140 | 458 | 1558 |
| 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 | −0.0085 | −0.0088 | −0.0083 | −0.0099 | 0.029 | 0.0073 | −0.011 |
| | (0.015) | (0.015) | (0.015) | (0.014) | (0.019) | (0.033) | (0.026) | (0.016) |
| Observations | 1606 | 1606 | 1606 | 1015 | 914 | 140 | 458 | 1558 |
| Pseudo-$R^2$ | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.05 | 0.01 | – |
| | Raw | Geog | Geog + Firm | Old | Group | SME | Manuf | IV |
| *Panel F. Dependent variable: Appropriate education and experience of new key workers* | | | | | | | | |
| (Question F5), probit and IV probit | | | | | | | | |
| Log density | 0.0094 | 0.012 | 0.0086 | 0.010 | 0.0070 | −0.0014 | 0.017 | −0.00052 |
| | (0.019) | (0.015) | (0.015) | (0.020) | (0.021) | (0.015) | (0.019) | (0.017) |
| Observations | 2946 | 2946 | 2939 | 731 | 1222 | 2252 | 2143 | 2862 |
| 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 Tables 2 and 3. Robust standard errors clustered by city in parentheses. [a], [b], [c]: significant at 1%, 5%, 10%. In all panels, columns 1 to 7 are estimated with probit. Column 8 is estimated by IV probit using 1871 and 1921 populations as instruments.

percolating through all other channels. Totally deriving $Y$ with respect to $B_c$ implies:

$$\frac{dY}{dB_c} = \frac{\partial f}{\partial Z}\frac{\partial Z}{\partial B_c} + \frac{\partial f}{\partial B_c}. \tag{6}$$

In simple terms and focusing on wages and density, the total effect of density on wages is the sum of the effect of density on local labour pooling times the effect of labour pooling on wages plus the direct effect of density on wages. The total effect of density on wages is measured by Eq. (1) for which the results are reported in Table 2. The effect of local labour market pooling on wages is measured by Eq. (3) for which the results are reported in Tables 4–6. To assess the importance of labour market pooling in agglomeration we still need to estimate the effect of labour market pooling on wages and the other effects of density on wages. We note that to assess the non-labour pooling effects of density on wages, we must condition out labour pooling. Conversely, Eq. (6) indicates that we should estimate the partial effect of labour pooling on wages, that is the effect of labour pooling on wages conditioning on density.

To estimate these two effects we thus consider regressions of the following form:

$$Y_i = \beta_0 + B_{c(i)}\beta_1 + X_i\beta_2 + Z_i\beta_3 + \epsilon_i, \tag{7}$$

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 labour market pooling variable measured for worker $i$. We also estimate the corresponding regressions for output per worker:

$$Y_j = \beta_0 + B_{c(j)}\beta_1 + X_j\beta_2 + Z_j\beta_3 + \epsilon_j, \tag{8}$$

where the our dependent variable $Y_j$ is now the log of output per worker of firm $j$, $B_{c(j)}$ 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 $Z_j$ is a labour 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 labour pooling at the worker level in turn. In the specifications reported in Panel C of Table 7, we also consider all our measures of labour pooling at the firm level. As can be seen from the results, only about half of the labour pooling variables is significant in our estimation of Eq. (7) for workers. On the other hand, all the labour pooling variables are significant in our estimation of Eq. (8) 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 as 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

**Table 7**
Labour market pooling and wages/output per worker.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| *Panel A. Dependent variable: log wage* | | | | | | |
| Labour market pooling variable: | Job change | Emp. change | Learning | Useful exp. | Training | Skill trans. |
| | (H2) | (H3) | (H4) | (H5) | (H6) | (H7) |
| Labour 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 | 1287 | 117 | 1287 | 945 | 1287 | 1287 |
| $R^2$ | 0.24 | 0.13 | 0.23 | 0.16 | 0.24 | 0.23 |
| *Panel B. Dependent variable: log wage* | | | | | | |
| Labour market pooling variable: | Difficu. Replac. | Find eq. Job | Spec. | Approp. Exp. | Approp. Edu. | |
| | (H8) | (H9) | (H10) | (H11) | (H12) | |
| Labour 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 | 1287 | 1287 | 1606 | 1606 | 1606 | |
| $R^2$ | 0.23 | 0.23 | 0.20 | 0.19 | 0.20 | |
| *Panel C. Dependent variable: log output per worker* | | | | | | |
| Labour market pooling variable: | Volunt. Turnover | Same Exp. | Training Days | Approp. Exp./edu. | | |
| | (F2) | (F3) | (F4) | (F5) | | |
| Labour 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 | 2743 | 2443 | 2548 | 2939 | | |
| $R^2$ | 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.

due to the smaller samples of workers who 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. As 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. First, as in previous regressions, density may be endogenous. In addition, some labour market pooling variables (e.g., changes in employer) are likely to be determined simultaneously with wages.

Combining the estimated relationship between our labour market variables and wages with previous estimations of the relationship between density and the labour market variables allows us to 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 labour market variables of Table 7. By doing that, we find that overall the labour market variables explain only a limited share of the urban wage/productivity premia.[18]

--------

[18] We also performed a principal component analysis to obtain a synthetic indicator of labour market pooling channels and use it as alternative explanatory variable in the estimation of Eqs. (7) and (8). The results confirm those presented in this section. Finally, we also confirmed that the results of Table 7 are robust to the introduction of regional dummies in Andini et al. (2013).

## 7. Conclusion

This paper looks at several different aspects of labour 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.

The paper's findings are broadly consistent with the many different sorts of labour 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 labour pooling, labour poaching, and hold up. In addition, the paper provides evidence consistent with agglomeration improving job matches.

The magnitudes, however, are relatively modest. The paper shows that labour market pooling gains are unlikely to account for a significant share of the agglomeration benefits accruing to workers 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 labour 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 labour 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.

## 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 8000 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 labour 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 3000 industrial firms, 465 construction companies and 1083 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 papers 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 labour 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 × months worked × 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 et al. (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 ratio between population and area ($km^2$) 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 labour market. For tax reasons and taking advantage of loopholes in labour 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: 4367). 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: 1287). Questions on appropriate experience (H11) and appropriate education (H12) are asked to all working individuals (number of observations: 1606). 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 observations. The number of observations fluctuates between 2452 and 2946 when we focus on local labour market variables.

## Appendix D. Territorial units of reference

*Local labour markets (or 'cities')*

Local labour markets are defined by the Italian National Institute of Statistic (Istituto Nazionale di Statistica, 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 labour 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 the 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' (Organisation for Economic Co-operation and Development, 2002). Italian local labour markets 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 labour 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.

### Supplementary material

Supplementary material for this article can be found online at http://dx.doi.org/10.1016/j.regsciurbeco.2013.04.003.

### References

Alesina, Alberto, Ichino, Andrea, 2009. L'Italia fatta in casa. Indagine sulla vera ricchezza degli italiani. Mondadori, Milano.

Alesina, Alberto, Danninger, Stephan, Rostagno, Massimo, 2001. Redistribution through public employment: the case of Italy. IMF Staff. Pap. 48 (3), 447–473.

Almazan, Andres, de Motta, Adolfo, Titman, Sheridan, 2007. Firm location and the creation and utilization of human capital. Rev. Econ. Stud. 74 (4), 1305–1327.

Andersson, Fredrik, Burgess, Simon, Lane, Julia I., 2007. Cities, matching and the productivity gains of agglomeration. J. Urban Econ. 61 (1), 112–128.

Andini, Monica, de Blasio, Guido, Duranton, Gilles, Strange, William C., 2013. Separate Appendices with Supplemental Material for Marshallian Labor Market Pooling: Evidence from Italy. Processed, University of Pennsylvania.

Anna Maria, Tarantola, 2011. Presentazione del rapporto l'Economia delle Marche nel 2010. Transcription of a public intervention in Ancona (13 June 2011). Banca d'Italia.

Bacolod, Marigee, Blum, Bernardo S., Strange, William C., 2010. Elements of skill: traits, intelligences, education, and agglomeration. J. Reg. Sci. 50 (1), 245–280.

Becattini, Giacomo, 1978. The development of light industry in Tuscany: an interpretation. Econ. Notes 7 (2–3), 107–123.

Becattini, Giacomo, 1979. Dal settore industriale al distretto industriale: Alcune considerazioni sull'unità d'indagine dell'economia industriale. Riv. Econ. Polit. Ind. 5 (1), 7–21.

Bleakley, Hoyt, Lin, Jeffrey, 2012. Thick-market effects and churning in the labor market: evidence from US cities. J. Urban Econ. 72 (2–3), 87–103.

Brunello, Giorgio, Lupi, Claudio, Ordine, Patrizia, 2001. Widening differences in Italian regional unemployment. Labour Econ. 8 (1), 103–129.

Brusco, Sebastiano, Paba, Sergio, 1997. Per una storia dei distretti industriali italiani dal secondo dopoguerra agli anni novanta. In: Barca, Fabrizio (Ed.), Storia del capitalismo italiano dal dopoguerra a oggi. Donzelli, Roma, pp. 265–334.

Bugamelli, Matteo, Cristadoro, Riccardo, Zevi, Giordano, 2009. La crisi internazionale e il sistema produttivo italiano: un'analisi su dati a livello d'impresa. Occas. pap. 58 (Bank of Italy).

Ciccone, Antonio, Hall, Robert E., 1996. Productivity and the density of economic activity. Am. Econ. Rev. 86 (1), 54–70.

Combes, Pierre-Philippe, Duranton, Gilles, 2006. Labour pooling, labour poaching, and spatial clustering. Reg. Sci. Urban Econ. 36 (1), 1–28.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, 2008. Spatial wage disparities: sorting matters! J. Urban Econ. 63 (2), 723–742.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, Roux, Sébastien, 2010. Estimating agglomeration economies with history, geology, and worker effects. In: Glaeser, Edward L. (Ed.), The Economics of Agglomeration. National Bureau of Economic Research, Cambridge (MA), pp. 15–65.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, 2011. The identification of agglomeration economies. J. Econ. Geogr. 11 (2), 253–266.

Dalmazzo, Alberto, de Blasio, Guido, 2011. Amenities and skill-biased agglomeration effects: some results on Italian cities. Pap. Reg. Sci. 90 (3), 503–527.

de Blasio, Guido, Di Addario, Sabrina, 2005. Do workers benefit from industrial agglomeration? J. Reg. Sci. 45 (4), 797–827.

De la Roca, Jorge, Puga, Diego, 2012. The Dynamic Earnings Premium of Dense Cities. Processed, CEMFI and IMDEA Social Sciences.

Di Addario, Sabrina, 2011. Job search in thick markets. J. Urban Econ. 69 (3), 303–318.

Di Addario, Sabrina, Patacchini, Eleonora, 2008. Wages and the city. Evidence from Italy. Labour 15 (5), 1040–1061.

Duranton, Gilles, Puga, Diego, 2004. Micro-foundations of urban agglomeration economies. In: Vernon Henderson, J., Thisse, Jacques-François (Eds.), Handbook of Regional and Urban Economics, Vol. 4. North-Holland, Amsterdam, pp. 2063–2117.

Ellison, Glenn, Glaeser, Edward L., Kerr, William R., 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. Am. Econ. Rev. 100 (3), 1195–1213.

Faini, Riccardo, Galli, Gianpaolo, Gennari, Pietro, Rossi, Fulvio, 1997. An empirical puzzle: falling migration and growing unemployment differentials among Italian regions. Eur. Econ. Rev. 41 (3–5), 571–579.

Fallick, Bruce, Fleischman, Charles, Rebitzer, James B., 2006. Job-hopping in Silicon Valley: some evidence concerning the microfoundations of a high-technology cluster. Rev. Econ. Stat. 88 (3), 472–481.

Freedman, Matthew, 2008. Job hopping, earnings dynamics, and industrial agglomeration in the software publishing industry. J. Urban Econ. 64 (3), 590–600.

Gerlach, Heiko A., Rønde, Thomas, Stahl, Konrad O., 2009. Labor pooling in R&D intensive industries. J. Urban Econ. 65 (1), 99–111.

Glaeser, Edward L., Gottlieb, Joshua D., 2009. The wealth of cities: agglomeration economies and spatial equilibrium in the United States. J. Econ. Lit. 47 (4), 983–1028.

Glaeser, Edward L., Maré, David C., 2001. Cities and skills. J. Labor Econ. 19 (2), 316–342.

Glaeser, Edward L, Resseger, Matthew G., 2010. The complementarity between cities and skills. J. Reg. Sci. 50 (1), 221–244.

Helsley, Robert W., Strange, William C., 1990. Matching and agglomeration economies in a system of cities. Reg. Sci. Urban Econ. 20 (2), 189–212.

Henderson, J. Vernon, 2003. Marshall's scale economies. J. Urban Econ. 53 (1), 1–28.

International Monetary Fund, 2011. Italy: 2011 Article IV Consultation Concluding Statement of the Mission, May 2011. Processed, IMF.

Istituto Nazionale di Statistica, 1997. I sistemi locali del lavoro 1991. ISTAT, Roma.

Jofre-Monseny, Jordi, Marín-López, Raquel, Viladecans-Marsal, Elisabet, 2011. The mechanisms of agglomeration. J. Urban Econ. 70 (2–3), 61–74.

Krugman, Paul R., 1991. Geography and Trade. MIT Press, Cambridge (MA).

Marshall, Alfred, 1890. Principles of Economics. Macmillan, London.

Matouschek, Niko, Robert-Nicoud, Frédéric, 2005. The role of human capital investments in the location decision of firms. Reg. Sci. Urban Econ. 35 (5), 570–583.

Mion, Giordano, Naticchioni, Paolo, 2009. The spatial sorting and matching of skills and firms. Can. J. Econ. 42 (1), 683–701.

Moulton, Brent R., 1990. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. Rev. Econ. Stat. 72 (2), 334–338.

Organisation for Economic Co-operation and Development, 2002. Redefining Territories. The Functional Regions. OECD, Paris.

Overman, Henry G., Puga, Diego, 2010. Labor pooling as a source of agglomeration: an empirical investigation. In: Glaeser, Edward L. (Ed.), The Economics of Agglomeration. National Bureau of Economic Research, Cambridge (MA), pp. 133–150.

Picard, Pierre M., Wildasin, David E., 2011. Outsourcing, labor market pooling, and labor contracts. J. Urban Econ. 70 (3), 47–60.

Puga, Diego, 2010. The magnitude and causes of agglomeration economies. J. Reg. Sci. 50 (1), 203–219.

Roback, Jennifer, 1982. Wages, rents and the quality of life. J. Polit. Econ. 90 (6), 1257–1278.

Rosenthal, Stuart S., Strange, William C., 2003. Geography, industrial organization, and agglomeration. Rev. Econ. Stat. 85 (2), 377–393.

Rosenthal, Stuart S., Strange, William C., 2004. Evidence on the nature and sources of agglomeration economies. In: Vernon Henderson, J., Thisse, Jacques-François (Eds.), Handbook of Regional and Urban Economics, Vol. 4. North-Holland, Amsterdam, pp. 2119–2171.

Rosenthal, Stuart S., Strange, William C., 2010. Small establishments/big effects: agglomeration, industrial organization and entrepreneurship. In: Glaeser, Edward L. (Ed.), The Economics of Agglomeration. National Bureau of Economic Research, Cambridge (MA), pp. 277–302.

Saraceno, Chiara, 1994. The ambivalent familism of the Italian welfare state. Soc. Pol. 1 (1), 60–82.

Shimer, Robert, 2005. The cyclicality of hires, separations, and job-to-job transitions. Fed. Reserv. Bank St Louis Rev. 87 (4), 493–507.

Strange, William C., Hejazi, Walid, Tang, Jianmin, 2006. The uncertain city: competitive instability, skills, innovation, and the strategy of agglomeration. J. Urban Econ. 59 (3), 331–351.

Wheeler, Christopher H., 2006. Cities and the growth of wages among young workers: evidence from the NLSY. J. Urban Econ. 60 (2), 162–184.

Wheeler, Christopher H., 2008. Local market scale and the pattern of job changes among young men. Reg. Sci. Urban Econ. 38 (2), 101–118.