LABOUR MARKET POOLING: EVIDENCE FROM ITALIAN INDUSTRIAL DISTRICTS

Guido de Blasio and Sabrina Di Addario[*]

1. Introduction

The classic argument for agglomeration is based on Marshall's three-pillar doctrine. According to Marshall (1890), there are three different reasons for the geographical concentration of a number of firms in the same industry. First, agglomeration creates a pooled market for workers with specialized skills. Second, it saves on transport costs owing to producers' proximity to input suppliers or final consumers. Third, it generates technological spillovers.

Following the contributions of Abdel-Rahman and Fujita (1990) and Krugman (1991), the new wave of theoretical and applied research on agglomeration developed Marshall's intuition in a number of directions. With reference to the second pillar, the impact of geographical concentration of industries on the availability of intermediate and final goods has been widely modeled (see Ottaviano and Puga, 1998 for a survey). The intuition, however, remains straightforward. On the one hand, a localized industry can support more specialized local suppliers. On the other hand, a localized industry implies a localized demand for final goods, which, in turn, makes localization more attractive to firms willing to save on shipping costs.

Regarding the third pillar, the importance for agglomeration of knowledge spillovers between nearby firms was well described by Marshall himself (1890): "*The mysteries of the trade become no mystery; but are as it were in the air…. Good work is rightly appreciated, inventions and improvements in machinery, in processes and the general organization*

---

[*] Bank of Italy, Economic Research Department We are grateful to Marco Bellandi, Andrea Brandolini, Matteo Bugamelli, Salvatore Chiri, Ivan Faiella, Marina Murat, Massimo Omiccioli, Giuseppe Parigi, Federico Signorini, William Strange, Sandro Trento, Eliana Viviano, Francesco Zollino and, in particular, to Luigi Cannari, for helpful discussions; we thank Antonia Mendolia for technical assistance. We also thank the participants at the CEPR conference "The Economics of Cities: Technology, Integration and Local Labour Markets" (6-8 June 2003, London).

of the business have their merits promptly discussed: if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus it becomes the sources of further new ideas". Notwithstanding Krugman's warning about the difficulties of measuring the third pillar ("*knowledge flows (...) are invisible; they leave no paper trail"*), the papers by Jaffe *et al.* (1993) and by Guiso and Schivardi (2000) provide some evidence on the relevance of the information-spillover motive.

The labour-market motive, the first pillar, has received a great deal of attention, but this has been focused on the theoretical side. Marshall's idea is that a pooled labour market benefits both firms and workers: *"A localized industry gains a great advantage from the fact that it offers a constant market for skill. Employers are apt to resort to any place where they are likely to find a good choice of workers with the special skill which they require; while men seeking employment naturally go to places where there are many employers who need such skill as theirs and where therefore it is likely to find a good market. The owner of an isolated factory, even if he has good access to plentiful supply of general labour, is often put to great shifts for want of some special skilled labour; and a skilled workman, when thrown out of employment in it, has no easy refuge"* (Marshall, 1890).

In his seminal work, Krugman (1991) shows that the efficiency gains from creating a localized industry with a pooled labour market are due to imperfectly correlated labour demand schedules for firms that may experience either "good times" or "bad times." Being in the same place would allow firms to take advantage of additional workers available during peak periods. At the same time, it would benefit workers, since the average rate of unemployment will correspondingly be lower. Krugman shows that agglomeration would drive wages up. In particular, because of efficiency gains, clustering would emerge as the outcome of a tug-of-war between firms, which prefer a less competitive labour market and hence dispersed production locations, and workers, who prefer a more competitive market and hence concentration. In contrast, Diamond and Simon (1990) show the importance of an insurance motive: workers could be willing to accept lower wages in locations where other firms stand by, ready to hire them.

In a vein similar to Krugman's tug-of-war, Rotemberg and Saloner (1990) suggest that workers – suppliers of industry-specific human capital, which is costly to acquire – might find it advantageous to locate where there are several potential firms that need such an input. In a pooled labour market, competition among firms would ensure a fair return to workers. In the absence of such competition, workers would be subject to the monopsony power of the firms. Anticipating such an outcome, workers would not choose to invest in industry-specific human capital. This model explains why the location decisions of firms and workers are interdependent and provides the prediction that wages within clusters should be higher. In a recent paper, Combes and Duranton (2001) propose a duopoly game where firms face a trade-off between the benefits of labour pooling (availability of workers whose knowledge helps reduce costs) and the costs of labour poaching (loss of some key workers to competition and the indirect effects of a higher wage bill to retain workers). The model combines the first and the second of Marshall's pillars, since workers have access to firm-specific knowledge and the pooled labour market acts as a conduit for spillovers, generating a set of predictions for wages and mobility. In particular, clusters should show higher wages and greater flows of workers between firms than isolated firms. Moreover, wages should be increasing over time, because experienced workers accumulate the kind of firm-specific knowledge that triggers poaching.

While a number of empirical analyses have been carried out for well-known clusters such as Silicon-Valley and Route 128 (or Prato and Biella in Italy) the lack of data has severely constrained the investigation of pooled labour markets *taken as a whole*. Poor data are the result of two shortcomings. First, macro-data are not fine enough to capture clusters (even when desegregated by region or smaller area), since the geographical extension of a cluster does not usually coincide with the administrative jurisdiction for which the data are available (e.g. municipalities, provinces, etc.). Second, in order to empirically investigate clusters taken as a whole, a definition providing a sensible singling out criterion is needed.

Following Signorini (2000), this paper proposes an empirical investigation of the labour market in Italian industrial districts (IIDs) that tries to overcome the above shortcomings by compounding the Bank of Italy's Survey of Household Income and Wealth micro-data with the official Istat-Sforzi algorithm that singles out IIDs. Our empirical strategy is composed of three blocks. First, we measure whether wages in IIDs are significantly different from those in isolated firms, controlling for the observable characteristics of workers, firms and geographical areas. In this context, we also analyze the role of the Mincerian determinants of wages. Second, we estimate the extent to which the probability of being self-employed, the likelihood of transiting from wage-and-salary to self-employment, and the worker's mobility across jobs are higher within IIDs. Finally, we provide a robustness check for the results based on sample restrictions and a finer definition of industrial districts. To this end, we undertake the same analysis on both superdistricts and a continuous district variable, as defined by Cannari and Signorini (2000). Superdistricts are an IID sub-sample in which the Istat-Sforzi district characteristics are highly emphasized. The continuous district variable is a variable that assigns each area (whether district or not) a value representing the degree of district features shown.

Our empirical investigation is linked to the recent applied literature on the returns to seniority and schooling. Neal (1995) and Parent (2000) show that the share of returns to seniority that could be attributed to firm-specific skills (which might represent a wage loss in case of displacement) is modest compared with the one that could be related to sector-specific skills (which are not lost as long as the worker remains in the same industry). Acemoglu and Angrist (2000) and Ciccone and Peri (2000) observe that the returns to schooling might be higher in agglomerations since clustering facilitates the exchange of ideas and triggers externalities that in turn raise private returns. Thus, the role of returns to seniority and schooling for cluster workers could shed light on these issues. Moreover, our work is related to a number of studies that concentrate on the peculiar functioning of the labour market within Italian industrial districts. This literature is mostly based on the analysis of specific case studies. Based on the evidence of Prato and Biella, Signorini (1994) suggests that average wages are higher in districts than elsewhere. In studies on Carpi, Solinas (1982; 1991) argues that districts are characterized by a wider role for firms in providing training to junior workers and by higher returns for skilled senior workers. This implies the existence of a peculiar IID wage curve, as workers are willing to accept reduced entry wages in exchange for on-the-job training, with the expectation of moving up the wage scale when they become senior workers and/or with the prospect of setting up on their own.[1] However, in a study on the provinces of Treviso and Vicenza Cingano (2003) does not find any evidence of a difference in the returns to

1 According to this literature, IID entrepreneurs encourage the most active employees to start an activity on their own as sub-contractors, in order to obtain for themselves an advantageous relationship with the sub-contracting firms and thus increase flexibility (Dei Ottati, 1992; Pyke *et al*., 1990).

seniority between IIDs and non-IIDs. In an analysis of the IID labour market using data from the Italian Social Security Institute, Casavola *et al.* (2000) do not find any clear evidence of a district wage premium. They argue, however, in favour of higher returns to seniority in IIDs, greater district worker mobility between jobs and higher district propensity to self-employment.[2] To our knowledge, no prior empirical analysis on the returns to education in IIDs has ever been undertaken.

Our results, based on IIDs taken as a whole, are noteworthy. We find a very limited role for geographical proximity. As for wages, we find only fragmentary evidence of a widespread wage premium within districts. Moreover, we find no evidence of district differentials for the returns to seniority, while district differentials for the returns to education might be negative. As for self-employment and labour mobility, we find that dwelling in a district has no impact on the probability of being self-employed and only a minor impact on the likelihood of transiting from wage-and-salary work to self-employment. Finally, there is no evidence of higher district worker mobility across jobs.

The paper is structured as follows. Section 2 provides some background information on the Italian industrial districts. Section 3 describes the dataset. Section 4 provides the results. Section 5 concludes.

2. Italian industrial districts

IIDs are geographically defined productive systems, characterized by the large number of firms that are involved, at various stages and in various ways, in the production of a homogeneous product (Pyke *et al.*, 1990). Most IIDs are located in the Centre and in the North of Italy (in particular in the North-East). District firms are mostly small and medium-sized enterprises (SMEs) specialized mainly in traditional sectors. Different IIDs specialize in different products of varying complexity and intended end-use. The role of IIDs can hardly be overstated: while SMEs

---

2 In order to analyze these issues, the authors compare the work experience and tenure of the average centre-north male manufacturing worker in IIDs to the corresponding ones in non-district areas. The difference they find between district and non-district workers is very small and could depend on the procedure adopted to impute past work-experience and tenure to the cohort of workers who were already in the sample at the beginning of the observation period (Casavola *et al.*, 2000, p. 61, footnote 4).

provide over 70 per cent of Italian manufacturing output (Eurostat, 1996), the IID share of total GDP is over 42 per cent (Istat, 1996). This latter share is even higher for sectors such as clothing, textiles, furniture, and leather. The best-known examples of IIDs are Prato, Carpi and Biella (specialized in textiles), Sassuolo (ceramic tiles), S. Croce sull'Arno and Solofra (leather), Martina Franca (furniture), and Barletta and Civitanova Marche (footwear).

A number of definitions have been proposed for IIDs (Becattini, 1990, and Brusco, 1990). These definitions refer to a set of elements of various nature: economic, institutional, sociological, and demographic. They extend Marshall's rationale to include the role played by the community and by the local institutions in favouring the diffusion of information and co-operative behaviour among agents. For example, Becattini, 1990 defines an IIDs as "a socio-territorial entity which is characterized by the active presence of both a community of people and a population of firms in one naturally and historically bounded area" where "community and firms tend to merge". These definitions are certainly intriguing, but they are difficult to relate to a quantitative benchmark, which is necessary to undertake empirical work. In the last decade, the shortcomings of the descriptive definitions led to a search for statistical criteria based on the productive system's structural characteristics, such as the specialization pattern and the presence of SMEs. These methodologies are necessarily partial with respect to the descriptive concept of industrial district. They are also discretionary, since minor changes in the key algorithm parameters would deliver different results. In this study (see the next section), in order to minimize the extent of discretion involved, we use three different measures of agglomeration.

3. Data

We use data from the 1998 Bank of Italy Survey on Households Income and Wealth (SHIW). This is a biannual survey that collects information on the economic behaviour of Italian families at the microeconomic level (detailed information on the SHIW can be found in

Brandolini, 1999).[3] In 1998 the SHIW surveyed 7,147 families, amounting to 20,901 individuals. In order to analyze the IID labour market, we focused on persons working in the non-farm private sector (excluding services to households) for a total of 4,665 observations.[4] Our sample thus comprises 3,161 employees and 1,504 self-employed workers, distributed in 63 different IIDs out of 217 local labour market areas.

Individuals were assigned to industrial districts by matching the 1998 SHIW with the 1991 Istat-Sforzi algorithm (ISA). According to the ISA, an area is defined as an industrial district if: a) it is a 'local labour market area', and b) the structure of its productive system is characterized by a dominant specialization and by the prevalence of SMEs.[5] A local labour market area (LLMA) is a self-contained geographical area, capable of offering employment to the majority of its resident population. The degree of self-containment is measured by the daily flows between production and residential sites (see Dalmazzo and de Blasio, in this volume).

IIDs are identified as LLMAs that satisfy the following four criteria:

1. the share of manufacturing employment in total non-farm employment must be higher than the corresponding share at the national level;
2. the share of SME manufacturing employment in total non-farm employment must be higher than the corresponding share at the national level;
3. for at least one sector, the specialization index must be greater than one. The specialization index is the ratio between the share of sector employment in total manufacturing employment and the corresponding share at the national level;

3 The SHIW considers households as the basic survey unit, and the sampling design is carried out in two stages: municipalities first, and then households. Data are collected in personal interviews by professionally trained interviewers, and are heavily processed to preserve data quality.

4 In particular, we excluded from the SHIW all the non-employed individuals and those employed as school teachers. We also excluded the following sectors: agriculture, hunting, forestry, and fishing; general government, defence, education, health and other public services; extraterritorial organizations and entities; domestic services provided to households and other private services.

5 According to the ISA definition, SMEs are firms with less than 250 employees. This ceiling has been deemed controversial (see for example Brusco and Paba, 1997 and Cannari and Signorini, 2000) on the grounds that it could be too high for IID firms. In Section 4, we consider only firms with less than 100 employees. to be SMEs.

4. in at least one sector for which the specialization index is greater than one, the share of SME employment in total employment must be higher than the corresponding share at the national level.

In 1991, there were 784 LLMAs and 199 IIDs (as defined by the ISA). In our 1998 sample, 908 employees and 411 self-employed workers belong to IIDs.

Sections 4.1 and 4.2 provide evidence based on the ISA definition of IIDs. Section 4.3 provides a robustness check for these results, using the Cannari and Signorini (2000) superdistricts and district continuous variable. Superdistricts represent the sub-sample of the Istat-Sforzi clusters where the ISA characteristics are highly emphasized. They are identified by a cluster analysis based on the four ISA criteria above. In particular, superdistricts typically display a very high incidence of manufacturing employment and SME manufacturing employment in total non-farm employment. Ninety-nine IIDs are classified as superdistricts. The district continuous variable associates to each LLMA a value representing the degree of district features shown by the area. It is calculated with a logit estimating the probability of each LLMA being classified as an IID according to the four ISA criteria. Thus, it represents an extension of the ISA methodology to the continuum.

4. Evidence

Our empirical strategy includes three building blocks. First, in Section 4.1 we measure whether wages within IIDs are significantly different from wages elsewhere, controlling for the observable characteristics of workers, firms and geographical areas. In this section we also analyze the role of the Mincerian determinants of wages. Then, in Section 4.2 we estimate the extent to which the probability of being self-employed, the likelihood of transiting from wage-and-salary to self-employment, and worker mobility across jobs are higher in IIDs than elsewhere. Sections 4.1 and 4.2 rely on the ISA definition. Section 4.3 provides a robustness check based on two alternative measures of agglomeration.

4.1 *Wages*

To test whether wages are higher in IIDs we estimate the following Mincerian[6] wage function:

$$logw_i = \alpha_0 + \alpha_1 SCHOOL_i + \alpha_2 EXP_i + \alpha_3 EXP_i^2 + \alpha_4 DISTRICT + Z_i \beta + u_i , \qquad (1)$$

where the dependent variable is the log of the hourly wage rate,[7] SCHOOL indicates the number of years of schooling, EXP denotes labour market experience, DISTRICT is a dummy variable for district workers, defined as those who reside in an IID irrespective of the size and branch of activity of the firm in which they work,[8] *Z* represents a vector of control variables for observable characteristics of firms and workers, and *u* is the error term.

A few features of our specification should be noted (see also the Appendix for the list of variables and the descriptive statistics)[9]. (i) The variable EXP is calculated as total number of years spent working[10]. To control for the potential endogeneity bias, we also use AGE instead of EXP. (ii) The vector of control variables includes some observable workers

---

6 See Mincer (1958) and Becker (1964). For a survey see Willis (1986) and Card (1999).

7 Earnings are measured after tax. We do not expect the use of net rather than gross earnings to significantly underestimate wage differentials, since tax structure is very similar across LLMAs. An additional problem is under-reported income: if the grey economy is more prominent in IIDs, then the omission of this income source might lead to underestimation of district differentials. We use hourly earnings in order to take into account irregular and overtime hours, which could be of some relevance in IIDs.

8 Using residence as the only identifying criterion for IID workers could, in principle, be criticized on the grounds that district firms are typically small sized and operate in the manufacturing sector, while we include all the workers residing in the IID area, as singled out by the ISA. If district wage premiums were only limited to SME manufacturing workers, the residence criterion could bias our results downwards. To lessen any fear that our results could depend on an erroneous identification criterion, in Tables 3 and 4 we restrict our sample to narrower characterizations of IIDs.

9 We also tested different specifications not shown here for the sake of conciseness. In particular, we estimated a version of Eq. (1) after decomposing EXP in two components: TENURE with the current employer, and PRIOR EXP, computed as EXP – TENURE (see below). Moreover, we replicated all the estimations with either EXP or AGE and TENURE. Finally, we also ran all the regressions on the log of annual wages rather than the log of hourly wages. Since this last set of specifications did not give particularly interesting results, we will not comment them any further.

10 EXP is calculated as the difference between current age and age at the first job held. Our proxy is thus more accurate than that computed by subtracting the years of schooling from age, since we do not erroneously attribute potential waiting unemployment to labour market experience. In contrast, AGE should be viewed as a more imperfect proxy of labour market experience, since it includes the years of schooling, possible unemployment periods, and labour market experience. When we replace EXP with AGE the IID wage differential lowers because AGE fails to recognize that IID workers enter the labour market earlier than non-IID workers.

and firms characteristics available in the SHIW dataset (dummy variables for FEMALE, SMEs, MANUFACTURING). Moreover, it includes two additional sets of controls: the LLMA unemployment rate from the 1996 Istat Labour Force Survey and the LLMA PAVITT specialization indexes, computed by the authors. This last set of controls allows us to provide some correction for the fact that the SHIW makes available only a breakdown between manufacturing and services and does not provide more detailed information about the branch of activity of the employee's firm.[11] (iii) The SHIW data set provides information about the employee's work status (blue-collar, office worker, junior manager, and manager). However, whether to control for work status is an open issue, since wages are likely to be correlated with status. Controlling for work status could therefore bias the education coefficient downwards. We tackle this issue on the empirical ground and provide estimates for both controlling and non-controlling for the employee's work status.

Table 1 shows the results.[12] The fit of the regression is quite good and all the variables are significant, with point estimates close to those found in previous studies using the SHIW (Cannari and D'Alessio, 1995 and Colussi, 1997),[13] even though returns to education turn out to be lower and returns to labour market experience higher than other authors' estimates. This is true regardless of the proxy used for labour market experience, and it is due to the fact that our sample excludes public sector workers, who have relatively high education levels and a compressed wage structure (Alesina *et al.*, 2001). Crucially, DISTRICT is positive and significant at the usual levels only in the regressions with EXP,[14] with an earning premium for cluster workers amounting to 3 per cent.[15]

---

[11] In particular, wages differ between low and high-intensity sectors. Thus, controlling for the extent to which the LLMA contains traditional versus high technology industries (PAVITT1-4), should help to offset the SHIW's lack of information on the firms' branches of activity.

[12] While the results presented in this paper are based on the 1998 SHIW, our conclusions are overwhelmingly confirmed when pooling 1995-1998 SHIW data.

[13] The aim of these papers is the analysis of the nationwide returns to education and to labour experience, with no reference to IIDs.

[14] In the specification in which we decompose EXP into TENURE and PRIOR EXP, we find that the nation-wide effect of TENURE is quite strong and comparable in size to the effect of PRIOR EXP. However, we do not find any district-specific effect.

[15] While, in principle, the presence of a centralized wage bargaining system could reduce earning differentials between areas, there is still considerable margin for wage differentials in Italy (according to Mauro *et al.*, 1999) wage differentials across regions, sectors and gender, vary between 10 and 30 per cent; see also Alesina, *et al.*, 2001).

Table 1

Earning functions: OLS estimates

Dependent variable: log of hourly wage rate				
	(1.1)	(1.2)	(1.3)	(1.4)-
DISTRICT	0.0285*	0.0339**	0.0235	0.0281
	(0.0167)	(0.0172)	(0.0167)	(0.0172)
EXP	0.0317***	0.0338***	-	-
	(0.0021)	(0.0022)	-	-
EXPSQR	-0.0005***	-0.0005***	-	-
	(0.0000)	(0.0000)	-	-
AGE	-	-	0.0499***	0.0506***
	-	-	(0.0046)	(0.0048)
AGESQR	-	-	-0.0005***	-0.0005***
	-	-	(0.0001)	(0.0001)
SCHOOL	0.0281***	0.0407***	0.0207***	0.0325***
	(0.0032)	(0.0028)	(0.0029)	(0.0026)
CONSTANT	1.7073***	1.7057***	0.9883***	0.9750***
	(0.0873)	(0.0913)	(0.1186)	(0.1227)
WSTATUS	yes	no	yes	no
R^2	0.39	0.36	0.39	0.36
No. obs.	3,129	3,129	3,129	3,129

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA PAVITT specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING.

To analyze the role of the Mincerian determinants of wages within clusters, we estimate a version of Eq. (1) that allows for interaction terms between RHS variables and DISTRICT.[16] Results are shown in Table 2. As for the district differential effects, no significant contribution of labour market experience is found (the interaction terms of EXP and AGE with DISTRICT are never significant). Moreover, there is evidence of negative cluster differentials for the returns to education (in column 2.1 the

16 Since the error disturbance is not significantly different across IIDs and non-IID LLMAs, the specification adopted here, in which data are pooled, is more efficient than running two separate regressions for the two sub-groups and then comparing the results (Greene, 2000).

reduction in the SCHOOL coefficient associated with cluster workers is about a half national average). The dummy DISTRICT, which is now meant to capture district wage differentials due to factors other than education and labour market experience, continues to be positive and significant only in the specifications with EXP. Summarizing, these results would suggest that IID wages might display a positive premium, which, however, does not reflect labour market experience. Moreover, this premium is eroded by a negative district differential for education, which penalizes relatively more the workers with higher human capital (for example, the results of column 2.1 would indicate that district wages are lower than non-district wages for workers with more than 13 years of schooling, which is the threshold for a high school diploma).[17]

There are a number of issues related to the choice of what to include within IIDs. The ISA provides a criterion to identify IIDs. However, once an LLMA is classified as an IID, the ISA leaves the question of which firms to include in the district quite open. In particular: (i) while the ISA is based on the prevalence of manufacturing, it is a matter of debate whether firms located within IIDs but belonging to sectors other than the industrial ones should be considered part of the district; (ii) while the ISA is based on the prevalence of SMEs, it is a matter of debate whether large firms located within IIDs should be considered part of the district; (iii) it is also a matter of debate whether our nation-wide sample should be replaced by the sub-sample of IIDs located in the Centre and North of Italy, which constitute a more homogeneous geographical area. In order to provide some robustness checks for the above issues, in Table 3 and Table 4 we show the results respectively for the specifications of column 1.1 and 2.1, estimated in different sub-samples. In the first column we exclude non-manufacturing firms from the sample. In the second column we keep only SMEs, lifting the restriction on the manufacturing sector. In the third column we apply the two restrictions simultaneously so that our sample comprises only manufacturing SMEs. Finally, we add a new restriction: we keep only the manufacturing SMEs located in the Centre and North. The main consequence of our check is that DISTRICT loses its significance

17 This result is consistent with our findings (not reported here) that within IIDs: (1) the average level of education is lower; (2) workers enter the labour market earlier. It is also consistent with the evidence presented by Casavola *et al*. (2000) on entrance to the IID labour market.

Table 2

Earning functions: OLS estimates with interaction terms

Dependent variable: log of hourly wage rate				
	(2.1)	(2.2)	(2.3)	(2.4)
DISTRICT	0.1865**	0.2243**	0.2929	0.2972
	(0.0931)	(0.1038)	(0.1933)	(0.1996)
EXP	0.0324***	0.0343***	-	-
	(0.0027)	(0.0027)	-	-
EXPSQR	-0.0005***	-0.0005***	-	-
	(0.0001)	(0.0001)	-	-
AGE	-	-	0.0533***	0.0538***
	-	-	(0.0058)	(0.0060)
AGESQR	-	-	-0.0005***	-0.0005***
	-	-	(0.0001)	(0.0001)
SCHOOL	0.0317***	0.0432***	0.0232***	0.0340***
	(0.0035)	(0.0031)	(0.0033)	(0.0029)
EXP*DISTRICT	-0.0014	-0.0004	-	-
	(0.0044)	(0.0044)	-	-
EXPSQR*DISTRICT	-0.0001	-0.0001	-	-
	(0.0001)	(0.0001)	-	-
AGE*DISTRICT	-	-	-0.0040	-0.0029
	-	-	(0.0096)	(0.0099)
AGESQR*DISTRICT	-	-	-0.0000	-0.0000
	-	-	(0.0001)	(0.0001)
SCHOOL*DISTRICT	-0.0150**	-0.0111*	-0.0124*	-0.0082
	(0.0073)	(0.0065)	(0.0067)	(0.0058)
CONSTANT	1.6727***	1.6642***	0.8878***	0.8748***
	(0.0912)	(0.0945)	(0.1371)	(0.1411)
WSTATUS	yes	no	yes	no
R^2	0.40	0.36	0.40	0.36
No. of observations	3,129	3,129	3,129	3,129

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA PAVITT specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – The additional controls have been interacted with DISTRICT.

Table 3

Earning functions: robustness

	(3.1)	(3.2)	(3.3)	(3.4)
Dependent variable: log of hourly wage rate				
	Manufacturing	SMEs	Manufacturing and SMEs	Manufacturing and SMEs in the Centre-North
DISTRICT	0.0142	0.0159	0.0078	-0.0059
	(0.0196)	(0.0197)	(0.2464)	(0.0245)
EXP	0.0305***	0.0318***	0.0310***	0.0291***
	(0.0027)	(0.0024)	(0.0034)	(0.0038)
EXPSQR	-0.0005***	-0.0005***	-0.0005***	-0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SCHOOL	0.0279***	0.0244***	0.0264***	0.0198***
	(0.0044)	(0.0039)	(0.0050)	(0.0068)
CONSTANT	1.6620***	1.5341***	1.5000***	1.7471***
	(0.1455)	(0.1163)	(0.1977)	(0.2058)
R^2	0.38	0.30	0.32	0.29
No. obs.	1,660	2,098	1,026	825

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA PAVITT specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING.

(Table 3). This is also due to the fact that restricting the sample towards narrower characterizations of IIDs makes the negative district differential in SCHOOL even more pronounced: it represents 72 per cent of the returns to education in 4.1 and above 80 per cent in 4.3 and 4.4.

We then turn to econometric issues. Since Griliches (1977), it has been well-known that there are a number of problems with the estimation of Eq. (1) by least squares using measured data. In particular, there are three issues. First, there could be an omitted variable problem, since ability, which is not observed, is presumably correlated with both Mincerian variables and wages. Second, there could be an endogeneity bias since human capital accumulation is the result of optimizing choice, taken by both individuals and their families in contexts where financial possibilities matter. Third, there could be a measurement error problem with both the education data, which are available as years of schooling, and proxies for labour market experience, which are measured as the number of years spent working.

Table 4

Earning functions: robustness with interaction terms

Dependent variable: log of hourly wage rate				
	(4.1)	(4.2)	(4.3)	(4.4)
	Manufacturing	SMEs	Manufacturing and SMEs	Manufacturing and SMEs in the Centre-North
DISTRICT	0.2025**	0.2023**	0.1459	0.1153
	(0.0930)	(0.0982)	(0.1080)	(0.1308)
EXP	0.0312***	0.0336***	0.0326***	0.0307***
	(0.0034)	(0.0030)	(0.0045)	(0.0056)
EXPSQR	-0.0004***	-0.0005***	-0.0005***	-0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SCHOOL	0.0338***	0.0282***	0.0323***	0.0289***
	(0.0049)	(0.0044)	(0.0061)	(0.0104)
EXP*DISTRICT	-0.0016	-0.0041	-0.0025	-0.0014
	(0.0052)	(0.0049)	(0.0061)	(0.0068)
EXPSQR*DISTRICT	-0.0001	-0.0000	-0.0001	-0.0001
	(-0.0001)	(0.0001)	(0.0001)	(0.0002)
SCHOOL*DISTRICT	-0.0244***	-0.0140*	-0.0275***	-0.0244**
	(0.0072)	(0.0086)	(0.0082)	(0.0118)
CONSTANT	1.5994***	1.4856***	1.4382***	1.6748***
	(0.1533)	(0.1209)	(0.2111)	(0.2320)
R^2	0.39	0.31	0.34	0.30
No. obs.	1,660	2,089	1,026	825

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA PAVITT specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. The additional controls have been interacted with DISTRICT.

Nevertheless, to provide some correction for the three econometric problems mentioned above we use instrumental variable estimates as suggested, for example, by Rosen (1977). This is also the preferred estimation strategy employed by Cannari and D'Alessio, 1995 and Colussi, 1997, so that their results could easily be compared with ours. In line with this work, we use family background variables as instruments: mother's and father's years of schooling and age. Table 5 and Table 6 show IV and OLS estimates for the sample of workers who provided information on age and educational attainment of the parents (the regressions correspond respectively to the specifications 1.1 and 1.2; and 2.1 and 2.2).

Table 5

Earning functions: OLS and IV estimates

Dependent variable: log of hourly wage rate				
	(5.1)	(5.2)	(5.3)	(5.4)
	OLS	IV	OLS	IV
DISTRICT	0.0259	0.0252	0.0312*	0.0288
	(0.0175)	(0.0176)	(0.0180)	(0.0182)
EXP	0.0322***	0.0327***	0.0342***	0.0348***
	(0.0022)	(0.0023)	(0.0022)	(0.0023)
EXPSQR	-0.0005***	-0.0005***	-0.0005***	-0.0005***
	(0.0000)	(0.0000)	(0.0001)	(0.0001)
SCHOOL	0.0303***	0.0370***	0.0425***	0.0530***
	(0.0022)	(0.0087)	(0.0030)	(0.0067)
CONSTANT	1.6526***	1.5728***	1.6486***	1.4927***
	(0.0941)	(0.1297)	(0.0990)	(0.1275)
R^2	0.40	0.40	0.36	0.36
No. obs.	2,777	2,777	2,777	2,777

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA PAVITT specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – Equations (5.3) and (5.4) do not include controls for employee's work status. – Instruments: age and educational qualifications of the parents. The Hausman test never allows us to reject the null hypothesis of exogeneity (at 1 per cent statistical significance). – The Sargan test never allows us to reject the hypothesis of orthogonality of the IV regression residuals and the instruments (at 1 per cent statistical significance). – The sample includes only workers who provided information on parental age and educational qualifications.

The endogeneity of education does not seem to be a problem in our data: the Hausman test does not enable us to reject the null hypothesis of exogeneity. This is somewhat contrary to what was expected. Indeed, the Hausman test indicated that there was an endogeneity issue with SCHOOL in both the papers by Cannari and D'Alessio, 1995 and Colussi, 1997. Our results are due to the fact that we use the 1998 survey of the SHIW, while the two papers refer to previous year surveys.[18] It is also worth mentioning that the Sargan test never enables us to reject the null hypothesis of orthogonality between the earning function residuals and the instruments. This implies that the variables of family background can be considered good instruments.

By and large, our findings remain confirmed.[19] In Table 5, while the point estimate of SCHOOL increases as expected, DISTRICT remains small in size and barely significant. In Table 6 the IV correction brings about a reduction in the statistical significance of both DISTRICT and SCHOOL*DISTRICT. Overall, the signs and sizes of the corrections resulting from IV estimates are minor, as expected given the results of the Hausman test.

Finally, it should be mentioned that our findings could, in principle, be affected by self-selection. If the hypothesis that the transition from wage-and-salary to self-employment is easier within IIDs turns out to be true, then our IID sample of observed wages could be biased downwards. Experienced and/or more talented workers would drop out from our sample, thus lowering average wages. To correct for such a problem we estimate a Heckman selection model. In order to determine whether the dependent variable is observed, the Heckman selection model calculates the likelihood of being an employee among a sample of employees and self-employed persons, using parents' educational attainment (MSCHOOL and FSCHOOL) and parents' work status (MWSTATUS and FWSTATUS) as selection variables. This set of selection variables is

---

[18] To understand the reasons for this discrepancy, we replicated the specification used by Colussi (1997), which excludes the LLMA unemployment rate and PAVITT specialization indexes, and all DISTRICT variables, both for 1993 (the year he analyzes) and 1998. Since Colussi's sample differs from ours (it includes, for instance, only heads of household, males, working full-time and all year), we also replicated his sample. We find that for the 1998 data the null hypothesis of no systematic difference between the IV and OLS coefficients cannot be rejected, even with the model specification and the sample used by Colussi.

[19] IV estimates (not shown here) were also run for all the models of Tables 1 and 2. Again, our results were broadly supported.

Table 6

Earning functions: OLS and IV estimates with interaction terms

Dependent variable: Log of hourly wage rate				
	(6.1)	(6.2)	(6.3)	(6.4)
	OLS	IV	OLS	IV
DISTRICT	0.1896*	0.0890	0.2354**	0.1320
	(0.0998)	(0.2591)	(0.1123)	(0.2739)
EXP	0.0326***	0.0330***	0.0345***	0.0350***
	(0.0027)	(0.0028)	(0.0027)	(0.0028)
EXPSQR	-00005***	-0.0005***	-00005***	-0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SCHOOL	0.0343***	0.0396***	0.0452***	0.0541***
	(0.0037)	(0.0095)	(0.0033)	(0.0073)
EXP*DISTRICT	-0.0001	0.0001	0.0006	0.0005
	(0.0047)	(0.0051)	(0.0047)	(0.0049)
EXPSQR*DISTRICT	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SCHOOL*DISTRICT	-0.0159**	-0.0076	-0.0120*	-0.0053
	(0.0078)	(0.0220)	(0.0071)	(0.0185)
CONSTANT	1.6161***	1.5509***	1.6044***	1.4727***
	(0.0977)	(0.1360)	(0.1016)	(0.1320)
R^2	0.40	0.40	0.37	0.36
No. obs.	2,777	2,777	2,777	2,777

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA PAVITT specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – The additional controls have been interacted with DISTRICT. – Equations (6.3) and (6.4) do not include controls for employee work status. Instruments: age and educational qualifications of the parents. – The Hausman test never allows us to reject the null hypothesis of exogeneity (at 1 per cent statistical significance). – The Sargan test never allows us to reject the hypothesis of orthogonality of the IV regression residuals and the instruments (at 1 per cent statistical significance). – The sample includes only workers who provided information on parental age and educational qualifications.

proven to be of key importance in the case of Italy due to children's propensity to follow their father's profession (Barca and Cannari, 1997). The likelihood-ratio test for correlation between the regression and the selection equations always allows us to reject the null hypothesis of no correlation (at 1 per cent statistical significance), which justifies the Heckman selection model with our data.

Table 7

Earning functions: OLS and Heckman selection model estimates

Dependent variable: log of hourly wage rate				
	(7.1)	(7.2)	(7.3)	(7.4)
	OLS	Heckman selection model	OLS	Heckman selection model
DISTRICT	0.0277	0.0275	0.0330*	0.0345*
	(0.0173)	(0.0173)	(0.0179)	(0.0179)
EXP	0.0323***	0.0319***	0.0344***	0.0347***
	(0.0022)	(0.0022)	(0.0022)	(0.0022)
EXPSQR	-0.0005**	-0.0005***	-0.0005***	-0.0005***
	(0.0001)	(0.0000)	(0.0000)	(0.0001)
SCHOOL	0.0298***	0.0279***	0.0419***	0.0433***
	(0.0034)	(0.0033)	(0.0029)	(0.0029)
CONSTANT	1.6611***	1.6506***	1.6555***	1.6420**
	(0.0923)	(0.0922)	(0.0968)	(0.0977)
R^2	0.40	-	0.36	-
No. obs.	2,809	4,045	2,809	4,045

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA ATECO specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – Equations (7.3) and (7.4) do not include controls for employee work status. – Selection variables: educational qualifications and work status of the parents. – The likelihood-ratio test for correlation between the regression and the selection equations always allows us to reject the null hypothesis of no correlation (at 1 per cent statistical significance). – The sample includes all the employees and self-employed persons who provided information on educational qualifications and work status of the parents. – The Heckman selection model estimates the likelihood of earning a wage (that is, of observing the dependent variable) by using 4,045 observations (2,809 employees and 1,236 self-employed).

Table 8

Earning functions: OLS and Heckman selection model estimates with interaction terms

Dependent variable: log of hourly wage rate				
	(8.1)	(8.2)	(8.3)	(8.4)
	OLS	Heckman selection model	OLS	Heckman selection model
DISTRICT	0.1792*	0.1755*	0.2259**	0.2223**
	(0.0991)	(0.0983)	(0.1115)	(0.1120)
EXP	0.0327***	0.0324***	0.0346***	0.0349***
	(0.0026)	(0.0026)	(0.0027)	(0.0026)
EXPSQR	-0.0005***	-0.0005***	-0.0005***	-0.0005***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SCHOOL	0.0337***	0.0317***	0.0445***	0.0456***
	(0.0037)	(0.0036)	(0.0033)	(0.0032)
EXP*DISTRICT	-0.0002	-0.0003	0.0006	0.0005
	(0.0047)	(0.0046)	(0.0047)	(0.0047)
EXPSQR*DISTRICT	-0.0001	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
SCHOOL*DISTRICT	-0.0154**	-0.0146**	-0.0114	-0.0109
	(0.0077)	(0.0076)	(0.0070)	(0.0070)
CONSTANT	1.6282***	1.6189***	1.6149***	1.6046***
	(0.0959)	(0.0956)	(0.0996)	(0.1003)
R^2	0.40	-	0.37	-
No. obs.	2,809	4,045	2,809	4,045

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA ATECO specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – The additional controls have been interacted with DISTRICT. – Equations (8.3) and (8.4) do not include controls for employee work status. – Selection variables: educational qualifications and work status of the parents. – The likelihood-ratio test for correlation between the regression and the selection equations always allows us to reject the null hypothesis of no correlation (at 1 per cent statistical significance). – The sample includes all the employees and self-employed who provided information on educational qualifications and work status of the parents. – The Heckman selection model estimates the likelihood of earning a wage (that is, of observing the dependent variable) by using 4,045 observations (2,809 employees and 1,236 self-employed).

Table 7 and Table 8 show the results (again, the specifications shown correspond to 1.1, 1.2 and 2.1, 2.1). The Heckman selection model does not provide any significant correction (for example, in columns 7.1 and 7.2 the average predicted dependent variable is equal respectively to 2.60 and 2.64). As will be made clear in the next section, this comes as no surprise, since the hypothesis of quicker district move to self-employment does not receive empirical support.

4.2 *Self-employment and labour mobility*

Is self-employment made easier within IIDs? This is a crucial issue for Italy, whose economy is more reliant on SMEs than other OECD countries. It is therefore important to understand the relative role of agglomerations, which are deemed to be areas prone to entrepreneurship. We single out two categories of self-employment: entrepreneur (business owner, owner or assistant in family business, active shareholder or partner) and free-lance (who runs his/her own activity with no employees). In the Appendix, Table A.2 contains some descriptive evidence, highlighting the fact that the share of entrepreneurs in the total sample (and to a lesser extent, that of free-lance workers) is higher in IIDs than in non-IIDs. Table 9 reports the empirical results from logit estimation, where the dependent variable is a binary indicator equal to l if a respondent pursues: an entrepreneurial activity (column 9.1); a free-lance activity (9.2); either of the two (9.3). The sample includes 4322 persons, 989 of whom are self-employed. We add the dummy DISTRICT to the specification adopted by Barca and Cannari, 1997. In the first column, the coefficients for AGE and AGESQR display the expected sign and high significance. The coefficient for SCHOOL is not significant, highlighting the negligible role of education in the chance of becoming an entrepreneur. The dummies FMANEX and FENTFL (equal to 1 respectively for those with a father manager or executive and those with a father entrepreneur or free-lance) are large in size and significant, stressing the role of inter-generational links in Italy. Surprisingly, the IID dummy variable shows no impact on the probability of becoming an entrepreneur. In the second column the likelihood of working free-lance does not vary much with age, decreases with education, and is less affected by inter-generational persistence. The district dummy again shows no effect. The results of the third column confirm these findings.

Table 9

Self-employment

Dependent variable: probability of being entrepreneur or free-lance

	(9.1)	(9.2)	(9.3)	(9.4)
	Dependent variable: ENT	Dependent variable: FL	Dependent variable: ENT and FL	Dependent variable: BOTH
DISTRICT	0.0917	0.0987	0.1325	0.1155
	(0.1922)	(0.1790)	(0.1487)	(0.1440)
AGE	0.1602***	0.0567*	0.1057***	0.2002***
	(0.0437)	(0.0327)	(0.0298)	(0.0373)
AGESQR	-0.0013***	-0.0002	-0.0006*	-0.0019***
	(0.0005)	(0.0004)	(0.0004)	(0.0004)
SCHOOL	0.0035	-0.0474**	-0.0331*	0.1052***
	(0.0225)	(0.0214)	(0.0182)	(0.0233)
FMANEX	0.5636*	-0.2276	0.1983	0.2959
	(0.3122)	(0.2959)	(0.2484)	(0.2602)
FENTFL	0.6556***	0.3415**	0.6317***	-0.0014
	(0.1522)	(0.1358)	(0.1183)	(0.1212)
CONSTANT	-9.7052***	-8.5283***	-8.5072***	-5.0649
	(1.4160)	(1.1954)	(1.0212)	(1.0286)
No. obs.	4,322	4,322	4,322	4,322
Wald χ^2	125.96	167.87	277.51	270.12
Prob > χ^2	0.000	0.0000	0.0000	0.0000
Pseudo R^2	0.15	0.16	0.22	0.13
Log likelihood	-1196.3807	-1415.1103	-1820.5003	-1726.6561

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA ATECO specialization indexes, employee work status, age at the time of the first job, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – The sample includes 4,322 persons, (3,333 employees and 989 self-employed) who provided information on parents' work status. Persons who worked BOTH as employees and self-employed workers number 757.

These results have been double-checked in a number of experiments. First, in analogy with Table 2, we logit estimate equations that allow for interactions between RHS variables and DISTRICT. Moreover, we run regressions without FMANEX and FENTFL. No differential district effects are ever found. Second, we provide sensitivity analysis for these results, in analogy with Tables 3 and 4. Again, the irrelevance of agglomeration for self-employment is strongly confirmed.

To test more directly the hypothesis that the transition from wage-and-salary to self-employment is easier within IIDs, we run two additional sets of logit estimates. First, the SHIW provides the binary variable BOTH, which is equal to 1 for those who have worked both as employees and as self-employed (757 in our sample). We estimate the probability of BOTH=1 in column 9.4). We find again no role for DISTRICT, while this probability is affected by SCHOOL and does not depend on the family work status. Second, we estimate the probability of being self-employed only for the sub-sample of those who had at least one work experience as employees (the descriptive evidence of Table A.2 shows that the share of entrepreneurs who previously worked as employees is substantially higher in IIDs). This reduces our original sample to 3,545 observations, 397 of which self-employed. Results are displayed in Table 10, which corresponds to the first three columns of Table 9. The likelihood of transiting from employee to entrepreneur is highly correlated with AGE and it is not driven by inter-generational links. Instead, the transition from employee to free-lance is not affected by AGE, while SCHOOL makes some (small) contribution. In both estimates, the dummy for agglomeration is never significantly different from zero; in the third column, however, we find a positive (and barely significant) contribution of DISTRICT.

Once more, these findings have been verified both by estimating equations that allow for interactions and by moving to narrower definitions of IIDs. The only notable upshot is that DISTRICT affects positively the likelihood of transiting from employee to free-lance (but not to entrepreneur) when the sample consists of only manufacturing SMEs.

Is worker mobility across jobs higher within agglomerations? To check this claim we estimate a Poisson maximum-likelihood regression for the number of activities, including temporary ones, performed up to 31st December, 1998. Results are shown in Table 11. Column 11.1 provides the results for the sub-sample of employees. Surprisingly, DISTRICT has a highly significant negative sign. We then check this result on several grounds. First, to accommodate the fact that within IIDs the distinction

Table 10

The transition from wage and salary work to self-employment

Dependent variable: probability of being an entrepreneur or free-lance			
	(10.1)	(10.2)	(10.3)
	Dependent variable: ENTWITH	Dependent variable: FLWITH	Dependent variable: ENTWITH and FLWITH
DISTRICT	0.2053	0.3422	0.3369*
	(0.2797)	(0.2471)	(0.2008)
AGE	0.2758***	0.0534	0.1435***
	(0.0651)	(0.0552)	(0.0424)
AGESQR	-0.0024***	-0.0001	-0.0009*
	(0.0008)	(0.0007)	(0.0005)
SCHOOL	0.0470	0.0575*	0.0572*
	(0.0373)	(0.0338)	(0.0308)
FMANEX	0.6824	0.2814	0.5714
	(0.5489)	(0.5054)	(0.4248)
FENTFL	0.3417	0.3186	0.3852**
	(0.2190)	(0.2005)	(0.1604)
CONSTANT	-14.9048***	-8.7028***	-10.8878***
	(2.0006)	(1.7956)	(1.4685)
No. obs.	3,545	3,545	3,545
Wald χ^2	150.54	140.80	240.34
Prob > χ^2	0.0000	0.0000	0.0000
Pseudo R^2	0.22	0.19	0.24
Log likelihood	-558.0304	-687.4855	-960.8347

Notes: All regressions are weighted to population proportions. – White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA ATECO specialization indexes, employee work status, age at the time of the first job, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – The sample includes 3,545 persons (of which 397 entrepreneurs and free-lance with previous experience as employees) who provided information on parents' work status.

between types of activities could be misleading, since employees might move to self-employment owing to particular circumstances,[20] the dependent variable is estimated for the wider sample of employees and self-employed (11.2). In this case, DISTRICT is less significant but still negative. Second, as before, we estimate for both specifications 11.1 and 11.2 equations that allow for interaction terms between RHS variables and DISTRICT, and we increasingly restrict our sample according to the sensitivity analysis proposed in Tables 3 and 4.[21] In these experiments, we never find any DISTRICT effect.

Table 11

Worker mobility across jobs: Poisson estimates

Dependent variable: number of activities held up to 12-31-1998		
	(11.1)	(11.2)
	Employees	Employees and self-employed
DISTRICT	-0.0941***	-0.0546*
	(0.0374)	(0.0310)
EXP	0.0321***	0.0309***
	(0.0041)	(0.0032)
EXPSQR	-0.0004***	-0.0004***
	(0.0001)	(0.0001)
SCHOOL	0.0074	0.0047
	(0.0056)	(0.0037)
CONSTANT	0.7636***	0.7877***
	(0.1697)	(0.1434)
No. obs.	3,015	4,343
Wald χ^2	302.46	359.71
Prob > χ^2	0.0000	0.0000

Notes: All regressions are weighted to population proportions. White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA ATECO specialization indexes, work status, and the following dummy variables: FEMALE, SOUTH, and SMEs, MANUFACTURING. – The sample includes all the people (4,343, of which 3,015 employees) who provided information on both the number activities carried on in their working lives and their age at the first job.

[20] For example during business cycle peaks.

[21] Moreover, we run regressions on the sub-samples of blue-collar workers and of employed people under 50.

4.3 *Beyond the ISA definition of clusters*

The results of Sections 4.1 and 4.2 are based on the identification of agglomerations provided by the official ISA criterion. The extent to which the crucial features of IIDs are accurately captured by this criterion has raised a number of controversies (Brusco and Paba, 1997 and Cannari and Signorini, 2000). In particular, the ISA splits up the LLMAs into two groups: districts and non-districts. The underlying idea is that non-linearities play a role: district effects materialize only above the ISA thresholds.

In this section we test whether our results could be due to the specific characteristics of the ISA. First, we verify whether the ISA thresholds are set too low. In other words, the fact that in a number of respects cluster labour markets do not differ from non-cluster ones could be due to the ISA classifying too many LLMAs as IIDs. To this end, we replicate the analysis of the above two sections for superdistricts (DISTRICT_S) rather than for the ISA IIDs. Second, we test whether our results could be due to the ISA being too tight. That is, IIDs and non-IIDs might indeed display similar characteristics to some extent, and their being split into two groups by the ISA may be too artificial a device. For this purpose, we extend the ISA methodology to the continuum by substituting DISTRICT with a continuous variable (DISTRICT_C) that associates each LLMA with a value for the degree of district features shown.[22]

Both DISTRICT_S and DISTRICT_C take the ISA as starting point. This is a nice feature of these indicators. In practice, different algorithms to single out districts could be proposed. However, abandoning the ISA criterion would not be without costs: to facilitate cross-country comparisons, ISA-type criteria are now being established by a number of OECD countries (OECD, 2001). Moreover, the ISA criterion is now a cornerstone of regional policy in Italy.

Table 12 gives a taste of our analysis. In this table we replicate models 1.1 and 2.1, replacing DISTRICT with the two alternative indicators of agglomeration. The results of the previous sections are extremely well-supported: since DISTRICT_S and DISTRICT_C are never

[22] As in Cannari and Signorini (2000).

Table 12

Sensitivity checks with superdistricts and district continuous variable

Dependent variable: hourly wage rate				
	(12.1)	(12.2)	(12.3)	(12.4)
DISTRICT_S	-0.0023	-	0.0234	-
	(0.0202)	-	(0.1177)	-
DISTRICT_C	-	0.0365	-	0.0757
	-	(0.0256)	-	(0.1178)
EXP	0.0316**	0.0317***	0.0315***	0.0324***
	(0.0021)	(0.0021)	(0.0023)	(0.0030)
EXPSQR	-0.0005***	-0.0005***	-0.0005***	-0.0005***
	(0.0000)	(0.0000)	(0.0001)	(0.0001)
SCHOOL	0.0281***	0.0283***	0.0273***	0.0299***
	(0.0032)	(0.0032)	(0.0035)	(0.0039)
EXP*DISTRICT_S	-	-	0.0026	-
	-	-	(0.0059)	-
EXPSQR*DISTRICT_S	-	-	-0.0001	-
	-	-	(0.0001)	-
EXP*DISTRICT_C	-	-	-	-0.0026
	-	-	-	(0.0057)
EXPSQR*DISTRICT_C	-	-	-	-0.0000
	-	-	-	(0.0001)
SCHOOL*DISTRICT_S	-	-	0.0064	-
	-	-	(0.0082)	-
SCHOOL*DISTRICT_C	-	-	-	-0.0072
	-	-	-	(0.0082)
CONSTANT	1.7077***	1.7141***	1.7062***	1.7036***
	(0.0875)	(0.0870)	(0.0889)	(0.0941)
R^2	0.39	0.39	0.39	0.40
No. obs.	3,129	3,129	3,129	3,129

Notes: All regressions are weighted to population proportions. White-robust standard errors in brackets. – * (**) [***] denotes statistical significance at 10 (5) [1] per cent level. – The additional controls included in the regressions are LLMA unemployment rate, LLMA ATECO specialization indexes, employee work status, and the following dummy variables: FEMALE, SOUTH, SMEs, MANUFACTURING. – In (12.3) and (12.4) the additional controls are also interacted with DISTRICT_S (C).

significant in our earning functions, with the notable exception of negative superdistrict differential returns to education (which appear only when we restrict our sample to manufacturing). Analogously, there are no different results for self-employment and labour mobility determinants. The only exception to the lack of a specific agglomeration effect lies is the likelihood of transiting from wage-and-salary worker to free-lance, as the dummy DISTRICT_C in this equation is positive and significant.

5. Conclusions

The theoretical literature on agglomeration place considerable emphasis on the labour market motive. Empirically, this literature suggests that labour market pooling affects wages, entry into self-employment, and worker mobility between jobs. The evidence presented in this paper amounts to a call for caution. We find little evidence of cluster differential effects for either wages or worker mobility. We also find no evidence of agglomeration differential returns to seniority, while the only cluster-specific effect that our data reveal is a negative differential for the returns to education. Finally, as for self-employment, the only role that agglomeration seems to play is limited to the transition from wage-and-salary employment to free-lance work.

At first glance, our results indicate that the first pillar of Marshall's story does not contribute much to the explanation of agglomeration in Italy. However, a different story could be proposed. Our results do not deny that the gains from labour market pooling could be important. Rather, they challenge the view that those gains benefit both firms and workers, since the advantages from participating in a pooled labour market seem to be quite limited for employees, at least in terms of wages and labour mobility.

It should be reiterated that, unlike case-studies, our conclusions rely on a general empirical investigation. That is, our analysis focuses on pooled labour markets taken as a whole, as singled out by the ISA cluster classification criterion (and by the two alternative criteria of Section 4.3). Consequently, our analysis does not preclude the possibility that the theoretical predictions are true in specific cases.

Finally, a number of suggestions for future research could be derived from our results. First, the relative bargaining strength of firms and workers within clusters seems to be an interesting topic, since the benefits of labour market pooling do not appear to be reaped by workers. In particular, workers who accumulate sector-specific human capital seem to do so at the expense of education. Accordingly, it would be interesting to analyze whether this under-accumulation of generalist human capital might weaken workers' bargaining power (perhaps by lowering their chances of finding another job). Second, the fact that worker mobility across jobs does not appear to be higher in agglomerations than in the rest of the country casts doubts on the idea that district firms' labour demand schedules are imperfectly correlated; a closer look at the characteristics of demand shocks for cluster firms would add an important piece of evidence. Finally, our results suggest that within IIDs there could be a relation between the modest role of education (for both workers and entrepreneurs) and the districts' specialization in traditional sectors. It is extremely important to understand the features and the consequences of such a relation, partly in view of the recent findings of Bils and Klenow (2000), who take a critical view of the traditionally proposed growth-enhancing role of schooling.

APPENDIX

List of variables

AGE. Age was computed as the difference between the survey year and the individual's year of birth.
AGESQR. Age squared.
AGE1JOB. Age at first job.
BOTH. This is a dummy variable that equals one if the individual has had work experience both as an employee and as a self-employed worker.
DISTRICT. This is a dummy variable that equals one if the LLMA is an IID according to the ISA classification.
DISTRICT_C. This is a continuous variable denoting the extent to which district characteristics are present in an LLMA, as in Cannari and Signorini (2000).
DISTRICT_S. This is a dummy variable that equals one if the LLMA is a superdistrict, as in Cannari and Signorini (2000).
ENT. This is a dummy variable that equals one if the individual is an entrepreneur.
ENTWITH. This is a dummy variable that equals one if the individual is an entrepreneur who had at least one work experience as an employee.
EXP. Work experience was computed as the difference between current age and age at first job.
EXPSQR. Work experience squared.
FAGE. Father's age.
FEMALE. This is a dummy variable that equals one if the individual is female.
FL. This is a dummy variable that equals one if the individual is a free-lance worker.
FLWITH. This is a dummy variable that equals one if the individual is a free-lance-worker who had at least one work experience as an employee.
FENTFL. This is a dummy variable that equals one if the individual's father is an entrepreneur or a free-lance worker.
FMANEX. This is a dummy variable that equals one if the individual's father is a manager or an executive.
FSCHOOL. Father's educational attainment.
FWSTATUS. Father's work status.
LWAGE. Hourly wages were calculated by dividing the annual earnings (from any activity as employee, including fringe benefits, net of taxes and social security contributions) by the total amount of hours worked in a year. In the analysis we used the natural logarithm of hourly wages:

$$lwage = \log\left(\frac{annualearnings}{no.hours * no.months * 4.3333}\right)$$

MAGE. Mother's age.
MANUFACTURING. This is a dummy variable that equals one if the individual works in a manufacturing firm.
MSCHOOL. Mother's educational attainment.
MWSTATUS. Mother's work status.
NUMJOBS. Number of jobs held.
PAVITT1-4. These four variables denote the LLMA PAVITT specialization indexes for the following categories: high technology, specialization, scale intensive, and traditional sectors. For each PAVITT category *j*, the specialization index is the ratio between the share relative to the LLMA and the share relative to the country of category *j*'s employees in total manufacturing industry employees:

$$I_{SP} = \frac{(\frac{Nj}{Nm})_{LLMa}}{(\frac{Nj}{Nm})_{\mathrm{ITA}}}$$

where N is the number of employees, j refers to Pavitt's four categories, m to manufacturing, LLMA to local labour market area, and ITA to Italy. PAVITT1-4 are computed by adapting the PAVITT classification, originally made for the 1981 ATECO system, to the 1991 ATECO system.
SCHOOL. The information on education available in the survey refers to the highest qualification earned by the individual. We derived the length of education by assigning: 0 years to no qualification; 5 years to elementary school; 8 years to middle school; 11 years to professional secondary school diploma; 13 years to high school; 16 years to a university diploma or other short-course university degree; 18 years to a bachelor's degree; and 20 years to a postgraduate qualification.
SMEs. This is a dummy that equals one if the firm has less than 100 employees.
SOUTH. This is a dummy that equals one if the individual resides in the South of Italy.
UNEMPLOYMENT RATE. The LLMA unemployment rate is calculated as the ratio of job seekers in the total labour force, using the 1996 Istat Labour Force Survey.
WSTATUS. This is variable that assumes the following values: 1 – Blue-collar workers 2 – Office worker; 3 – Junior manager; 4 – Manager; 5 – Member of the professions; 6 – Business owner; 7 – Free-lance; 8 – Owner or assistant of a family business; 9 – Active shareholder or partner.

Table A.1

Summary statistics

	Total sample			IIDs		
	No. obs.	Average	Std. dev.	No. obs.	Average	Std. dev.
AGE	4665	39.012	11.628	1319	38.553	11.715
AGE1JOB	4663	19.118	4.774	1319	18.509	4.406
AGESQR	4665	1657.144	952.701	1319	1623.453	948.541
BOTH	4665	0.177	0.382	1319	0.208	0.406
DISTRICT	4665	0.283	0.450	1319	1.000	0.000
DISTRICT_C	4665	0.296	0.350	1319	0.755	0.251
DISTRICT_S	4665	0.101	0.302	1319	0.359	0.480
ENT	4665	0.123	0.328	1319	0.136	0.343
ENTWITH	3743	0.120	0.325	1108	0.151	0.358
EXP	4663	19.896	12.759	1319	20.043	12.510
EXPSQR	4663	558.590	587.875	1319	558.111	561.148
FAGE	4423	70.742	13.942	1265	70.152	14.021
FEMALE	4665	0.308	0.462	1319	0.355	0.479
FL	4665	0.130	0.336	1319	0.124	0.329
FLWITH	3743	0.058	0.234	1108	0.080	0.272
FENTFL	4464	0.393	0.488	1273	0.414	0.493
FMANEX	4464	0.066	0.248	1273	0.049	0.217
FSCHOOL	4502	6.139	4.093	1283	6.089	3.639
FWSTATUS	4464	4.256	3.416	1273	4.313	3.479
LWAGE	3129	2.446	0.425	909	2.458	0.381
MAGE	4209	67.745	13.385	1194	67.098	13.287
MANUFACTURING	4665	0.419	0.493	1319	0.516	0.500
MSCHOOL	4294	5.475	3.600	1215	5.570	3.361
MWSTATUS	4257	7.288	3.052	1205	6.683	3.386
NUMJOBS	4665	2.058	1.650	1319	2.071	1.438
PAVITT1	4665	0.620	0.893	1319	0.290	0.530
PAVITT2	4665	0.984	0.578	1319	1.139	0.627
PAVITT3	4665	0.890	0.582	1319	0.847	0.647
PAVITT4	4665	1.277	0.530	1319	1.408	0.541
SCHOOL	4665	10.373	3.650	1319	10.299	3.505
SMEs	4519	0.769	0.422	1282	0.782	0.413
SOUTH	4665	0.257	0.437	1319	0.022	0.147
UNEMPLOYMENT RATE	4665	10.890	7.307	1319	6.182	2.579
WSTATUS	4665	3.635	3.143	1319	3.604	3.219

Table A.2

Frequency of entrepreneurs and free-lance workers in the sample

	IIDs		NON-IIDs		TOTAL	
	No.	%	No.	%	No.	%
			Overall sample			
ENT	138	11.2	269	8.7	407	9.4
FL	159	12.9	423	13.7	582	13.5
ENT-FL	297	24.1	692	22.4	989	22.9
Total	1232	100.0	3086	100.0	4322	100.0
			Sample of those with previous experience as an employee			
ENTWITH	75	29.4	101	20.1	176	23.2
FLWITH	76	29.8	145	28.9	221	29.2
ENTWITH-FLWITH	151	59.2	246	49.0	397	52.4
Total	1057	100.0	2488	100.0	3545	100.0

REFERENCES

Acemoglu, D. and Angrist, J. (2000), "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws," in B.S. Bernanke and K. Rogoff (eds.), *NBER Macroeconomics Annual 2000*, Cambridge MA, MIT Press.

Abdel-Rahman, H. and Fujita, M. (1990), "Product Variety, Marshallian Externalities, and City Sizes," *Journal of Regional Science*, Vol. 30, pp. 165-183.

Alesina, A., Danninger, S. and Rostagno, M. (2001), "Redistribution Through Public Employment: The Case of Italy," *International Monetary Fund, Staff Papers*, Vol. 48, No. 3, pp. 447-473.

Barca, F. and Cannari, L. (1997), "Imprenditori e investitori: il ruolo dell'istruzione", in N. Rossi (ed.), *L'istruzione in Italia: solo un pezzo di carta?*, Bologna, Il Mulino.

Becattini, G. (1990), "The Marshallian Industrial District as Socio-Economic Notion," in Pyke *et al.* (1990).

Becker, G.S. (1964), *Human Capital*, New York, Columbia University Press for NBER.

Bils, M. and Klenow, P. (2000), "Does Schooling Cause Growth?," *American Economic Review*, Vol. 90, No. 5, pp. 1160-1183.

Brandolini, A. (1999), "The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality," Banca d'Italia, Temi di discussione, No. 350.

Brusco, S. (1990), "The Idea of the Industrial Districts: Its Genesis," in Pyke *et al.* (1990).

Brusco, S. and Paba, S. (1997), "Per una storia dei distretti industriali italiani dal secondo dopoguerra agli anni novanta," in F. Barca (ed.), *Storia del capitalismo italiano dal dopoguerra ad oggi*, Rome, Donzelli.

Cannari, L. and D'Alessio, G. (1995), "Il rendimento dell'istruzione: alcuni problemi di stima," Banca d'Italia, Temi di discussione, No. 253.

Cannari, L. and Signorini, L.F. (2000), "Nuovi strumenti per la classificazione del sistemi industriali," in Signorini (2000).

Card, D. (1999), "The Causal Effect of Education on Earnings," in O. Ashenfelter and D. Card (eds.), *Handbook of Labour Economics*, Amsterdam, North Holland.

Casavola, P., Pellegrini, G. and Romagnano, E. (2000), "Imprese e mercato del lavoro nei distretti industriali," in Signorini (2000).

Ciccone, A. and Peri, G. (2000), "Human Capital and Externalities in Cities," Centre of Economic Policy Research, Discussion Paper, No. 2599.

Cingano, F. (2003), "Returns to Specific Skills in Industrial Districts," *Labor Economics*, Vol. 10, No. 2, pp. 149-164.

Colussi, A. (1997), "Il tasso di rendimento dell'istruzione in Italia," in N. Rossi (ed.), *L'istruzione in Italia: solo un pezzo di carta?*, Bologna, Il Mulino.

Combes, P. and Duranton, G. (2001), "Labour Pooling, Labour Poaching and Spatial Clustering," CERAS Working Paper 01-03.

Dei Ottati, G. (1992), "Fiducia, transazioni intrecciate e credito nel distretto industriale," *Note Economiche*, Vol. 22, No. 1-2, pp. 1-30.

Diamond, C.A. and Simon, C.J. (1990), "Industrial Specialization and the Returns to Labour," *Journal of Labor Economics*, Vol. 8, No. 2, pp. 175-201.

Eurostat (1996), *Enterprises in Europe*, Fourth Report, Bruxelles.

Greene, W.H. (2000), *Econometric Analysis*, New York, Prentice Hall International.

Griliches, Z. (1977), "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica*, Vol. 45, No. 1, pp. 1-22.

Guiso, L. and Schivardi, F. (2000), *Information Spillover and Factor Adjustment*, Banca d'Italia, Temi di discussione, No. 368.

Istat (1996), *Rapporto Annuale. La situazione del Paese nel 1995*, Istituto Poligrafico dello Stato, Roma.

Jaffe, A., Trajtenberg, M. and Henderson, R. (1993), "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *Quarterly Journal of Economics*, Vol. 108, pp. 577-598.

Krugman, P. (1991), *Geography and Trade*, Cambridge, Mass., MIT Press.

Marshall, A. (1890), *The Principles of Economics*, London, Macmillan, 1920.

Mauro, P., Prasad, E. and Spilimbergo, A. (1999), "Perspectives on Regional Unemployment in Europe," International Monetary Fund, Occasional Paper No. 177.

Mincer, J. (1958), "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, Vol. 66, No. 4, pp. 281-302.

Neal, D. (1995), "Industry-specific Human Capital: Evidence from Displaced Workers," *Journal of Labor Economics*, Vol. 13, No. 4, pp. 653-677.

OECD (2001), *Territorial Reviews: Italy*, Paris, OECD.

Ottaviano, G. and Puga, D. (1998), "Agglomeration in the Global Economy: A Survey of the 'New Economic Geography'," *World Economy*, Vol. 21, No. 6, pp. 707-731.

Parent, D. (2000), "Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics," *Journal of Labor Economics*, Vol. 18, No. 2, pp. 306-323.

Putnam, R.D. (1993), *Making Democracy Work: Civic Tradition in Modern Italy*, Princeton, Princeton University Press.

Pyke, F., Becattini, G. and Sengenberger, W. (1990), *Industrial Districts and Inter-firm Co-operation in Italy*, Geneva, International Institute for Labor Studies.

Rosen, S. (1977), "Human Capital: A Survey of Empirical Research," in H. Ehrenberg (ed.) *Research in Labor Economics*, Greenwich, Conn., JAI Press.

Rotemberg, J.J. and Saloner, G. (1990), “Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade,” National Bureau Economic Research Working Paper, No. 3228.

Signorini, L.F. (1994), “Una verifica quantitativa dell’effetto distretto,” *Sviluppo Locale*, No. 1.

Signorini, L.F. (2000), *Lo sviluppo locale*. *Un’indagine della Banca d’Italia sui distretti industriali,* Rome, Donzelli.

Solinas, G. (1982), “Labour Market Segmentation and Workers’ Careers: The Case of the Italian Knitwear Industry,” *Cambridge Journal of Economics*, Vol. 6, No. 4, pp. 331-352.

Solinas, G. (1991), “L’anzianità aziendale: alcune considerazioni sul caso emiliano,” *Politica economica*, Vol. 7, No. 3, pp. 409-447.

Willis, R.J. (1986), “Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions,” in O. Ashenfelter and R. Layard (eds.), *Handbook of Labour Economics*, Vol. 1, Amsterdam and New York, North Holland.