Available online at www.sciencedirect.com
ScienceDirect

ELSEVIER

Regional Science and Urban Economics 37 (2007) 1 – 21

regional SCIENCE & urban ECONOMICS

www.elsevier.com/locate/regec

Do spatial agglomeration and local labor market competition affect employer-provided training? Evidence from the UK

Giorgio Brunello [a,b,c,*], Francesca Gambarotto [a]

[a] *Department of Economics, University of Padua, Italy*
[b] *IZA, Germany*
[c] *CESifo, Germany*

Available online 31 July 2006

Abstract

In this paper we use British data to ask whether local employment density–which we take as a proxy of labor market competition–affects employer-provided training. We find that training is less frequent in economically denser areas. We interpret this result as evidence that the balance of poaching and local agglomeration effects on training is negative. The effect of density on training is not negligible: when evaluated at the average firm size in the local area, a 1% increase in density reduces the probability of employer-provided training by 0.014, close to 4% of the average incidence of this type of training in the UK.

JEL classification: J24; R12
Keywords: Training; Spatial economics; Britain

1. Introduction

When labor economists analyze the training decision, they usually overlook spatial factors, in spite of the relevance of the spatial agglomeration literature. The typical line of approach is that firms decide to invest in human capital when they can hold the trained worker and profit of her higher productivity, i.e. when the poaching risk is low. Poaching occurs because the skills learnt in a single firm are not wholly specific to that firm, but can be transferred to some extent to competitors. In most circumstances, the risk of poaching depends both on the type of skill and on the presence of local

* Corresponding author. Department of Economics, University of Padua, via del Santo 33, 35100, Padua, Italy.
E-mail address: giorgio.brunello@unipd.it (G. Brunello).

0166-0462/$ - see front matter
doi:10.1016/j.regsciurbeco.2006.06.006

competitors, who can find it profitable to hire the trained employee. If competitors are located far away, however, some workers may be discouraged by the expected mobility costs.

If we take a local labor market perspective, it is clear that dense labor markets, with more workers and more firms, present better opportunities to locate a better job than sparse labor markets. The higher risk of poaching typical of denser areas implies that firms located in these areas face an uncertain return to training and tends to under-invest in human capital. At the same time, however, we recognize that local density can also affect positively the training decisions of employers. When skills and technical knowledge are complements, trained workers, who are more capable of exploiting the positive spillovers associated to spatial proximity, are more productive in agglomerated areas. In imperfectly competitive labor markets, this productivity premium can translate into higher marginal benefits of training if it is not fully absorbed by higher wages.

It follows that firms located in denser areas, when deciding whether and how much to train their employees, need to take into account the effects of the higher risk of poaching and the effects associated to the complementarity of skills and local knowledge spillovers. When we compare similar firms in local labor markets with different density, employer-provided training incidence can be higher, or lower, in denser areas, depending on the direction and relative weight of these effects.

This paper is an empirical investigation of the relationship between local economic density and employer-provided training in British Nuts 2 groups of counties, based on longitudinal BHPS (British Household Panel Survey) data for the period 1994–2000[1]. While we are not aware of other empirical studies which address the same issue[2], our research is related to the growing number of studies which investigate the relationship between local economic density and productivity, starting with Ciccone and Hall (1996). This literature focuses mainly on knowledge spillovers and pecuniary externalities as key ingredients of local economic growth, but pays little attention to the potential link between agglomeration and productivity induced by the effects of local labor market competition on the incentives to train. Suppose that more local competition significantly affects the provision of training by firms. Since training is expected to increase productivity, the uncovered relationship between agglomeration and productivity across local areas could be partly driven by differences in the incentives to invest in the production of skills. We investigate this link, and find evidence that employer-provided training is lower in denser areas, which we take to suggest that the relationship between agglomeration and productivity would be even stronger were it not for the negative impact of density on training.

The data show that the incidence of employer-provided training varies significantly across British local labor markets. Regions with higher than average training levels are mainly urban areas: Greater London, the South East (Essex, Kent, Brighton) and the South West (Southampton, Oxford), the regions of Manchester and the West Midlands. Fig. 1 plots training incidence against employment density in each area, measured as the log of the number of employees in the private sector per squared kilometer. Density is a measure of spatial proximity, and has been used by Ciccone and Hall (1996), to capture the positive pooling externalities associated to close and repeated interactions among economic agents. Inspection of the figure does not reveal any clear pattern, but obviously a significant relationship could be obscured by the presence of numerous confounding effects, such as the industrial and occupational composition of labor, the average level of educational attainment and else.

[1] These data are included in the European Community Household Panel, December 2001 release (contract 14/99 with the Department of Economics, University of Padua).

[2] A recent exception is Brunello and De Paola (2004), who study the relationship between training and local economic density in a matching model and test the implications of such model on Italian data.

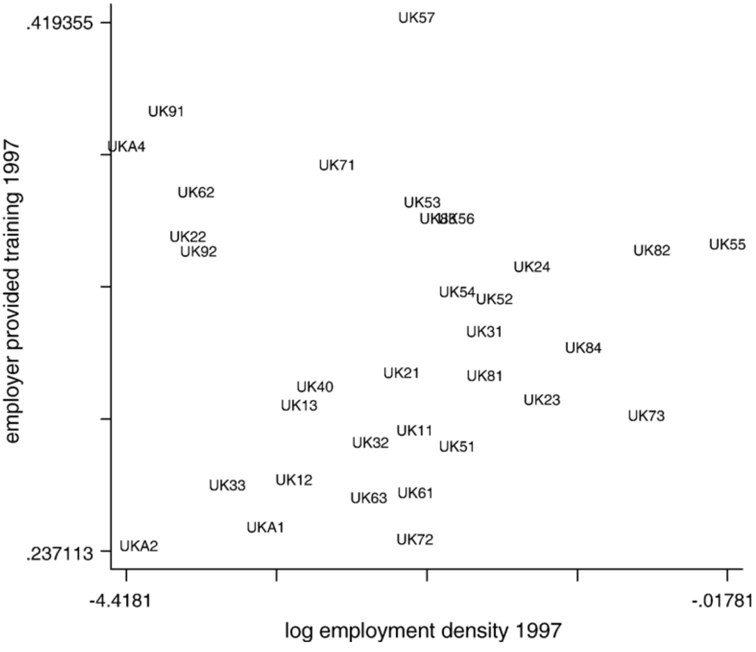

Fig. 1. Employer-provided training and log employment density, by NUTS 2 regions, 1997. UK11: Durham; UK12: Cumbria; UK13: Northumberland; UK21: Humberside; UK22: North Yorkshire; UK23: South Yorkshire; UK24: West Yorkshire; UK31: Derbyshire; UK32: Leicestershire; UK33: Lincolnshire; UK40: East Anglia; UK51: Bedfordshire; UK52: Berkshire; UK53: Surrey; UK54: Essex; UK55: Greater London; UK56: Hampshire; UK57: Kent; UK61: Avon; UK62: Cornwall; UK63: Dorset; UK71: Hereford; UK72: Shropshire; UK73: West Midlands; UK81: Cheshire; UK82: Greater Manchester; UK83: Lancashire; UK84: Merseyside; UK91: Clwyd; UK92: Gwent; UKA1: Borders; UKA2: Dumfries; UKA3: Highlands; UKA4: Grampian.

The paper is organized as follows. Section 1 contains a brief summary of the relevant literature, with special emphasis on agglomeration effects. Section 2 discusses the relationship between pooling and poaching effects and employer-provided training in local labor markets. Section 3 presents the empirical specification and Section 4 illustrates the data. The last two sections discuss the main results and some robustness exercises. Conclusions follow.

2. A brief summary of the literature

Labor market pooling as a Marshallian externality plays a crucial role in the location and spatial agglomeration theory of Krugman (1991). In this literature, geographic concentration produces spatial increasing returns and has a positive impact on the production/diffusion of technological innovation and knowledge (Audretsch and Feldman, 2004). Pooling externalities occur when the spatial concentration of workers fosters job turnover and improves the match between demand and supply, thus favoring the diffusion of ideas and increasing the productivity of firms located in the area (Rosenthal and Strange, 2004). The knowledge spillovers typical of agglomerated areas are closely linked to pooling externalities because knowledge is partly embodied in workers and its diffusion is driven by labor turnover (Combes and Duranton, 2006).

Particular attention has been given in this literature to the impact of knowledge externalities on the growth of cities. Glaeser et al. (1992) distinguish three types of local externalities: MAR

(Marshall-Arrow-Romer) externalities, driven by industrial specialization, Porter externalities, originated by specialization and strong competition among local firms, and Jacobs externalities, i.e. knowledge spillovers from diversity in the structure of local production. Their empirical evidence shows that urban productivity growth is increased by diversity and reduced by specialization. Henderson et al. (1995), observes that knowledge spillovers play a different role in traditional and high-tech industries and in different stages of growth. In high-tech sectors, Jacobs externalities stimulate growth when location takes place and MAR externalities are important for location persistence. In traditional sectors only MAR externalities are relevant.

Ciccone and Hall (1996) and Ciccone (2002) study the relationship between local economic density–measured as the number of employees per squared kilometer–and productivity and find that productivity is higher in denser areas, both in Europe and in the US. The density of economic activity positively affects productivity by reducing transportation costs, by increasing the interaction of firms because of spatial agglomeration and by fostering knowledge spillovers. Finally, Glaeser and Maré (2001), in their attempt to explain the urban wage premium, observe that labor matching works better in economically dense areas, such as cities. In addition, cities can provide opportunities for higher levels of interaction among agents, which fosters the accumulation of human capital. Moreover, the skill endowment of workers in urban areas can be quickly updated because the local context facilitates the learning process.

3. Local pooling, poaching and employer-provided training

Standard economic theory suggests that employers invest in training up to the point where the marginal benefits of the investment–in terms of higher labor productivity–are equal to the marginal (direct and opportunity) costs. The willingness of firms to pay for training depends on its degree of transferability. As argued by Becker (1964), in his classical study on human capital, the cost of general training is entirely borne by the employee if labor markets are perfectly competitive, because in this case the accumulated skills can be fully transferred to other firms. However, the influential work by Acemoglu and Pischke (1999) has shown that in the presence of information asymmetries, search costs and frictions in the labor market, firms may be willing to invest in (general) training, because imperfect competition drives a wedge between the productivity gains and the wage gains from training, which Acemoglu and Pischke define as wage compression.

In imperfect local labor markets, density has two key effects on employer-provided training. First, denser areas can offer better matching opportunities and a higher probability of re-employment. Therefore, firms located in these areas have better opportunities to hire skilled workers from the market and consequently a lower incentive to train (see Stevens, 1994, Rotemberg and Saloner, 2000, Brunello and Medio, 2001). Stronger labor market competition in denser areas can also favor poaching and further discourage employer-provided training. When the threat of poaching is too strong and there is harsh competition for skilled workers, firms can even decide to relocate in a less dense area. Almazan et al. (2003) suggest that relocation can be profitable for high-tech firms operating in science-based industries. For these firms the investment in human capital is crucial for production and the poaching risk need to be minimized, not only because it reduces the expected benefits of training, but also because it becomes a powerful vehicle of diffusion to competitors of new developed ideas and techniques (Combes and Duranton, 2006).

Second, density can affect training if the knowledge spillovers associated to local labor market pooling and the skills possessed by employees are complements. As briefly reviewed in Section 1 of the paper, a key tenet of the new economic geography is that localization economies are important for productivity and growth. In this approach, the higher concentration of individuals and firms in dense economic areas increases knowledge spillovers and fosters technological progress (Ciccone and Hall, 1996), but the ability of firms located in these areas to adapt new technologies and ideas is strictly related to the skills of their labor force (see Acemoglu, 2002). Training increases productivity for two reasons. First, the employee increases her skills in performing the relevant job. Second, she improves her ability to understand and process the flow of information from the productive environment where the firm is located and to translate this information into higher productivity on the job (see Rosenthal and Strange, 2003). This complementarity between local spillovers and skills suggests that the productivity gains from training are higher in denser areas.

Depending on the behavior of local wages, these higher productivity gains can translate into higher marginal benefits of training, and generate a higher incentive to train. To illustrate, let local wages be a linear combination of local productivity and the workers' outside option, a standard result when wages are determined by Nash bargaining between the parties. If the accumulated skills are only partially transferable to other firms, the wage gains from training are proportional to productivity gains, because the outside option is marginally affected by the investment in training. In this case, if the bargaining power of workers does not differ much between local areas, larger productivity gains can turn into larger benefits of training and–given marginal training costs–firms in denser areas can have a stronger incentive to invest in training.

If accumulated skills are easily transferable, the workers' outside option is affected by training to a higher extent, and larger productivity gains in denser areas need not imply larger marginal benefits of training. Define the difference between productivity and wage gains from training as wage compression. Acemoglu and Pischke (1999) show that wage compression is more relevant when labor market frictions are serious or when there are information asymmetries. Since it is not obvious that frictions and asymmetries are exacerbated by higher geographic concentration, wage compression can be less severe in denser areas when the skills provided by training are easily transferable. In this case, training in dense areas can be less profitable than in sparse areas. Overall, the pooling externalities associated to dense labor markets have the potential of affecting the returns to training, but the direction and size of this effect depends both on the nature of training and on the wage determination process.

Since the combination of pooling externalities and poaching effects can generate a trade-off in the training decisions taken by employers, the sign of the relationship between local economic density and training incidence is potentially ambiguous. A similar trade-off has been pointed out by recent applications of economic geography models to local labor markets, which have explained the localization decisions of firms as the outcome of contrasting positive pooling effects and negative poaching effects (see Combes and Duranton, 2006). In dense economic areas–with a relatively high concentration of workers and firms–labor market pooling improves the matching of workers and firms and facilitates the transmission of knowledge and innovative activities. As a result, expected labor productivity increases, which encourages localization. However, the risk of poaching is higher in dense areas. Since knowledge is partly embodied in workers, this risk discourages localization. We argue in this paper that the combination of pooling and poaching effects not only influences the localization decisions of firms and their productivity, but also their willingness to train employees.

The effect of density on employer-provided training is also affected by the structure of the local market. In general we expect that poaching effects are stronger in areas with a higher share of small and medium firms. The reason is that smaller firms "…may have higher training costs per employee than larger firms because they cannot spread the fixed costs of training over a large group of employees…" (Lynch, 2003). For these firms, poaching is relatively more attractive.

Equally dense areas which differ in the degree of industrial specialization can exhibit substantially different pooling and poaching effects. On the one hand, specialization can foster network externalities, as in the MAR concept introduced in Section 1 of the paper, and therefore increase beneficial pooling effects. On the other hand, a more specialized industrial structure, with a higher proportion of firms producing closely related products and using closely related processes, can favor within-industry mobility of trained employees, with a negative impact on training, especially if skills are industry specific (see Neal, 1995).

4. The econometric specification

The discussion in the previous section suggests that the relationship between employer-provided training and the density and specialization of economic activity in local labor markets is complex and cannot be signed a priori. In each area, positive pooling externalities interact with poaching externalities and labor turnover and affect training decisions. The overall effect of density and specialization on employer-provided training depends on the relative strength of these forces at play.

Our empirical specification assumes that the individual probability of receiving employer-provided training depends on individual, area-specific and aggregate effects. More in detail, we use the following probit specification

$$\mathrm{Prob}\{T_{ijt} = 1\} = \Phi\{\beta X_{ijt} + \gamma D_{jt} + \delta Y_{jt} + \sigma Z_t + u_i + \varepsilon_{ijt}\} \qquad (1)$$

where T is employer-provided training, X a vector of individual effects, D is log employment density, measured as the log of the ratio between employment in the area and the size of the area in squared kilometers, Y is a vector of confounding area-specific effects, Z is a vector of aggregate effects, ε is a normally distributed and serially uncorrelated error term, u_i is a normally distributed and time invariant individual effect, and the indices i, j and t are for the individual, the area and time respectively. As explained below, we identify the area with the Nuts 2 aggregation[3].

A potentially serious problem with (1) is that the error term includes unobserved individual heterogeneity. If unobserved individual ability and training are complements–as assumed by Acemoglu and Pischke (1998)–and if abler individuals concentrate in denser areas (Glaeser and Maré, 2001), then the estimated coefficient of density should be upward biased. Yatchew and Griliches (1985) discuss omitted variables bias in the context of the probit model and show that this bias exists even if the omitted variable–in our case unobserved ability or productivity–is uncorrelated with included variables.

We deal with this problem as follows. First, we control for unobserved individual ability by including among the explanatory variables in (1) both individual education–measured as a dummy equal to 1 if the individual has a high school or college degree and to 0 otherwise–and controls for the size of the firm, the type of labor contract, tenure–measured as a dummy equal to 1 if the individual was hired before 1991 and to 0 otherwise–and dummies for the occupation and

[3] Regional areas in the European Communities are organized into Nuts levels, depending on the degree of aggregation. For the UK, Nuts 1 regions correspond to Standard Regions, Nuts 2 areas to Groups of Counties, and Nuts 3 areas to Counties.

industry[4] where the individual is employed. Education captures important components of ability, and so do tenure and the allocation of workers to different jobs and labor contracts. Second, we use the Blundell and Smith (1986) test to verify whether local employment density–conditional on the controls for unobserved ability–can be treated as weakly exogenous in our sample. This test consists of two steps: in the former step local density D is regressed on the full set of exogenous variables as well as on additional instruments. In the second step, the residuals from this regression are included as an additional variable in the probit model. Assuming normality, the test of weak exogeneity is equivalent to a t-test on the residuals.

Given that density is measured at a higher level of aggregation than individual information on employer-provided training, the errors in (1) are likely to be correlated within clusters but independent between clusters. We correct the standard errors of the estimates for the presence of area, industry and time cluster effects in the error term, depending on the selected specification. Under the null hypothesis that density D is weakly exogenous, there is no need to adjust standard errors further to account for the fact that the first step residuals are generated regressors[5].

The vector Z includes both time and Nuts 1 dummies. The latter set of dummies captures all the unobserved effects common to aggregations of Nuts 2 local areas, but leaves enough cross-section variation for the identification of a significant relationship between density and training in Nuts 2 regions. The vector Y includes average years of schooling in the local Nuts 2 area, a measure of human capital spillovers (see Moretti, 2004), and the local unemployment rate, a proxy of local economic conditions. The training policy of firms can also be affected by local labor market policy. If employer-provided and publicly-provided training are substitutes, we expect the former to be lower, ceteris paribus, in areas where public provision is more widespread. We capture public policy effects with the two dummies Ob1 and Ob2, which indicate the Nuts 2 regions covered by European structural funds under the Objective 1 and Objective 2 programs.

Since we expect the relationship between local employment density and employer-provided training to vary with the structure of local industry, we finally add to the vector Y the area-specific index of industrial specialization S, computed as

$$s_{jt} = \sum_{k} \left(\frac{E_{kjt}}{E_{jt}} \right)^2 \tag{2}$$

where k is for the industry in the area[6] and E is employment, the area-specific average firm size in the manufacturing sector and its interaction with local employment density. As discussed in Section 2, this interaction verifies whether poaching effects are stronger in areas characterized by small and medium firms. By adding the vector Y to the list of explanatory variables, we depurate the estimate of the relationship between employment density and employer-provided training from the effects of other confounding area-specific factors.

[4] The industries are: mining and quarrying, manufacture of food products, manufacture of textiles, clothing and leather, manufacture of wood and paper products, manufacture of chemicals, manufacture of metal products and equipment, other manufacturing, construction, wholesale and retail trade, hotels and restaurants, transport and communication, financial intermediation and real estate, renting and business services. The occupations are: managers, professionals, technicians, clerks, service workers, craft workers, plant and machine operators and elementary occupations.

[5] See Blundell and Smith (1986, p.681). We are grateful to Guglielmo Weber for advice.

[6] Cingano and Schivardi (2004) use a similar variable.

5. The data

We use the *British Household Panel Survey* data included in the December 2001 release of the European Community Household Panel, a longitudinal household and personal survey modeled on the US Panel Study of Income Dynamics (PSID). As shown by Arulampalam et al., (2004), training participation in Europe is highest in Denmark and the UK. Since the perception of training events and the

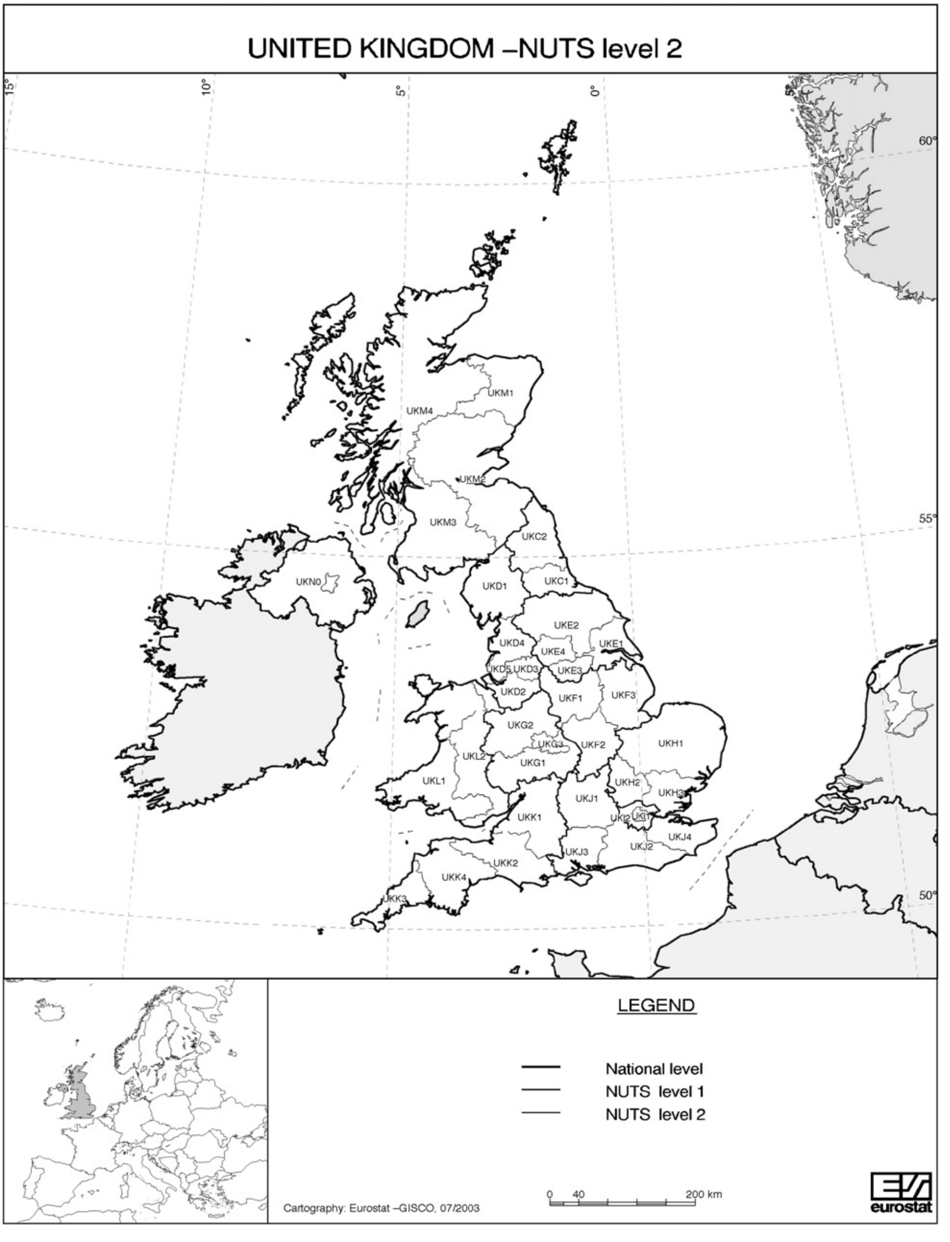

Fig. 2. Nuts 2 (group of counties) map in the UK.

Table 1
Summary statistics (BHPS, 1997)

	Mean	Standard deviation	Min	Max
Employer-provided training	0.317			
Gender	0.593			
Medium-sized firm	0.109			
Large-sized firm	0.416			
High school and higher	0.501			
Objective 1 dummy	0.012			
Objective 2 dummy	0.249			
Full time job	0.873			
Permanent contract	0.939			
Hired before 1991	0.129			
Age	35.568	10.687	17	59
Total employment in the Nuts 2 area (thousands)	607.243	373.269	123.030	1801.839
Firm size in manufacturing	25.916	7.483	12.674	41.907
Local unemployment rate	0.066	0.027	0.030	0.129
Employment density in the Nuts 2 area (thousands)	0.210	0.265	0.012	0.982
Average years of schooling in the area	12.722	0.373	11.664	13.337
Region specific index of specialization in the Nuts 2 area	0.131	0.016	0.097	0.164
Industry specific index of MAR externalities	0.139	0.077	0.017	0.320
Industry specific index of Jacobs externalities	7.320	2.780	3.580	17.184

interpretation of training questions can differ considerably across countries, we prefer to focus on a single country and on within-country variations. By selecting Britain, with more than 30 Nuts 2 areas and an average participation to employer-provided training close to 30%, we both have a significant number of local labor markets and reduce the risk of having too few training events in some local labor markets.

For each individual, the survey gives information on employer-provided training and on the area of residence of the household. This area, however, does not necessarily coincide with the area of employment, to which the concept of employment density discussed in the paper applies. The lack of coincidence between area of residence and area of work is a serious problem when we select relatively fine definitions of areas of residence, such as Nuts 3 or higher, because these regions do not necessarily correspond to the travel to work areas (TTWA) defined by commuting behavior. The mismatch between residence and work is less serious, however, when the areas of residence are broader, as in the case of Nuts 2 and Nuts 1. The natural choice in our context is the Nuts 2 aggregation (groups of counties). In the UK the average size of a group of counties is 6914 km^2, wide enough to have most residents working in the area. Broader or finer classifications such as Nuts 1 and Nuts 3 would be less appropriate, either because pooling externalities dissipate over larger regions (see Ciccone, 2002) or because the areas are too small to contain the relevant TTWA.

Fig. 2 shows a map of Britain divided into Nuts 2 areas. Even by choosing the Nuts 2 classification, we cannot completely rule out that, for some individuals in the sample, area of residence and area of work do not coincide. Therefore our empirical indicator of density measures true density with error. Under the conditions spelled out by Yatchew and Griliches (1985)[7], measurement error generates an attenuation bias in the estimated relationship between the probability of training and local density. We try to attenuate the measurement error associated to the mismatch between area of birth and area of residence by experimenting in the robustness section of the paper with an alternative measure of density, the average density of the region of residence and of the neighboring regions.

[7] The key condition is that the measurement error is normally distributed.

Table 2
Probit estimates of the probability of employer-provided training

	(1)	(2)	(3)
Age	0.002 (0.002)	−0.009*** (0.002)	−0.009*** (0.002)
Age squared * 100	0.006* (0.004)	0.010*** (0.005)	0.009*** (0.003)
Gender	0.049*** (0.007)	0.016* (0.008)	0.016* (0.008)
Log employment density in the Nuts 2 area	−0.014** (0.006)	−0.022*** (0.005)	−0.104*** (0.035)
High school and college degree		0.067*** (0.008)	0.067*** (0.008)
Full time job		0.049*** (0.010)	0.049*** (0.009)
Permanent contract		0.096*** (0.010)	0.096*** (0.008)
Hired before 1991		−0.027** (0.011)	−0.027** (0.011)
Medium-sized firm		0.037** (0.016)	0.037*** (0.016)
Large-sized firm		0.081*** (0.008)	0.082*** (0.007)
Average years of schooling in the Nuts 2 area			0.004 (0.012)
Local unemployment rate			−0.216 (0.287)
Area-specific index of industrial specialization in the Nuts 2 area			−0.387* (0.229)
Average firm size in the Nuts 2 area			0.052 (0.033)
Average firm size * log density			0.028*** (0.011)
Industry dummies	No	Yes	Yes
Occupation dummies	No	Yes	Yes
EU Objective 1 and 2 dummies	No	No	Yes
P-value of the *F*-test for the inclusion of confounding area specific variables			0.000
Number of observations	16,171	16,171	16,171
Pseudo *R* squared	0.058	0.120	0.121

Pooled cross-section time series data. Average partial effects. Dependent variable: *T*.
The regressions include a constant, year and Nuts 1 dummies. Cluster adjusted robust standard errors. One, two and three stars when the coefficients are significantly different from zero at the 10%, 5% and 1% level of confidence respectively. The *F*-test tests for the joint significance of the variables in vector *Y*.

The main question on vocational training in the data is as follows: “Have you at any time since January (in the previous year) been in any vocational education or training, including part-time and short-courses?”. Since the reference period may overlap with the reference period of the previous wave, long training events could be counted more than once. According to Arulampalam et al. (2004), however, there is little chance of double counting in Britain, because training events are generally very short. Conditional on a positive answer to the training question, the individual is asked whether training is paid for or organized by the employer. We consider such training as employer-provided and define the dummy T as equal to 1 if the individual has received employer-provided training since January of the year before the survey, and as equal to 0 if she has received no training. The treatment of the recipients of training not provided by the employer would require an additional category and a multinomial approach. However, since this group represents only 3.57% of the sample, we prefer to omit it from the key regressions. In the robustness section of the paper, however, we also present the estimates of a multinomial logit, which assigns the group to a separate category[8].

[8] The data also distinguish between on-the-job and off-the-job training. It is questionable whether such distinction can be used to separate general from firm-specific training, and we refrain to do so in this paper. See the discussion in Bassanini and Brunello (2003).

In general, having information on who pays for training only refers to who pays nominally. Workers who say that their employer pays for their training could also receive lower wages, and thus pay at least part of the costs. To check this possibility in our data, we regress log gross wages in year t on individual controls–industry, occupation, firm size, age, gender and tenure–and on a dummy equal to 1 if the individual received employer-provided training from January of the year t to the time of the survey in year $t+1$, and to 0 in the event of no training, using the fixed effect estimator to control for unobserved heterogeneity[9]. If trained individuals pay part of the training cost with a lower wage, we should find that the coefficient of the employer-provided training dummy attracts a negative sign. This is not the case, however, as the estimated coefficient is small, positive (0.005) and not statistically significant (standard error: 0.007)[10].

Our sample comprises men and women who are (i) between the ages of 17 and 59 years working at least 15 h/week; (ii) not employed in agriculture, the public sector or non-profit organizations. We pool all observations from the first (1994) to the last available wave (2000) and use time dummies to account both for aggregate effects and for the fact that the training question has been somewhat altered from 1998 onwards (see Arulampalam et al., 2004). Our measure of density is total employment in private industry and services per squared kilometer. Total employment in private industry and services in each Nuts 2 area and for each year is computed as follows: (a) we obtain aggregate employment from official Eurostat publications; (b) we use the cross-sectional BHPS stratification weights to compute for each available year the distribution of employment by local area and disaggregate aggregate data by Nuts 2 area[11]. Regional Nuts 2 variables such as the unemployment rate and average firm size in manufacturing are computed using Eurostat data from the International Statistical Yearbook[12] and the online information from the website www.nomisweb.co.uk[13].

Table 1 presents for the year 1997 the descriptive statistics of the main variables used in the empirical analysis. On average about 32% of the individuals in the sample has been involved in employer-provided training, a number which is broadly in line with official statistics (see OECD, 2003). Fifty percent of the sampled individuals have at least upper secondary education, and 53% are employed in medium and large firms. Average total employment in the Nuts 2 areas was 607.2 thousand employees in 1997, with a minimum of 123 thousand (North Yorkshire) and a maximum of 1801 thousand (Greater London). Average employment density was 210 employees per squared kilometer, ranging between 12 in South Western Scotland and 982 in Greater London, and average firm size in manufacturing was 25.92, with a range between 12.67 and 41.90. The unemployment rate in 1997 was on average 0.066, with a minimum of 0.03 (Oxfordshire) and a maximum of 0.129 (Merseyside), and the average years of education in each group of counties were on average 12.7, with a range between 11.6 and 13.3. Finally, the index of industrial specialization (MAR) in the same year was on average 0.139, with a range between 0.017 and 0.320, and the index of industrial diversity was on average equal to 7.320, with a minimum of 3.580 and a maximum of 17.184.

[9] Ideally, we would like to consider only contemporaneous training events. However, this is not possible with our data, which cover training events from the beginning of year t to the time of the survey in year $t+1$.

[10] Using wages at time $t+1$ does not change the sign of the correlation between training and wages.

[11] An alternative to (a) is to use Nuts 1 employment (source: www.nomisweb.co.uk) and disaggregate it by Nuts 2 area using (b). Results are very close to the ones obtained with the methodology described in the text.

[12] The data are available at the Department of Economics, University of Padova.

[13] Average years of schooling are computed using weighted BHPS data over the period 1994–1997, by assigning 11 years of school to individuals who have completed lower secondary education, 13 years to individuals with upper secondary education and 16 years to college graduates.

6. The results

We start the presentation of our results with Table 2, which shows the estimates of the probit model (1) based on the pooled sample. The numbers in the table are not the marginal effects, but the average partial effects of a unit change in each explanatory variable. The difference between these two measures can be illustrated as follows: let Prob$\{y=1|x\}=\Phi(x\beta)$ be the probit model, where Φ is the standard normal distribution of the error term ε. The marginal effect of a unit change in x_j is $\beta_j\phi(\overline{x}\beta)$, with the density evaluated at the mean value of x, $\overline{x}$. In the presence of neglected heterogeneity, captured by the term $u \approx N(0, \tau^2)$, the probit model in latent variable form is $y^* = x\beta + \gamma u + \varepsilon$, and $\text{Prob}\{y = 1|x\} = \Phi\left(\frac{x\beta}{\sigma}\right)$, where $\sigma^2 = \gamma^2\tau^2 + 1$. In this case, the average partial effect is $\frac{\beta j}{\sigma}\Phi\left(\frac{\overline{x}\,\beta}{\sigma}\right)$, which corresponds to the average marginal effect across the distribution of u in the population (Wooldridge, 2002).

Table 2 is organized in three columns. The specification in the first column is the most parsimonious, and excludes both individual educational attainment, type of contract, tenure, industry, occupation and firm size dummies, which we expect to control for unobserved individual heterogeneity, and the variables in the vector Y of confounding area-specific effects. The second column adds the controls for individual heterogeneity but omits the variables in Y. The third column adds also the variables in Y. We find that the individual controls attract the expected sign-negative for age and positive for male employees[14]. Moreover, employment in a full time job and with a permanent contract increases the probability of training, which is lower for individuals hired before 1991.

The estimated effect of log employment density on employer-provided training is negative and statistically significant in all specifications. The inclusion of controls for individual heterogeneity in column (2) leads to an increase–in absolute value–of the coefficient associated to log density, which suggests that the estimated contribution of density in column (1) is an upper bound. The richest specification in column (3) shows that training is higher in Nuts 2 areas with higher average firm size and lower in Nuts 2 areas with higher area-specific industrial specialization[15].

The interaction of log density with average firm size has a positive and statistically significant coefficient. Therefore, the negative correlation between employment density and employer-provided training is lower the higher the average firm size in the area. A natural interpretation of this result is that labor turnover is higher in areas where small firms prevail, which encourages these firms to hire the required skills from the market as an alternative to costly training[16]. When evaluated at the sample mean value of log firm size (3.208), the average partial effect of a 1% change in employment density on the probability of employer-provided training is equal to -0.014 $[(-0.1041+0.0281*3.207)]$.

We formally assess whether log employment density is weakly exogenous by applying the Blundell and Smith (1986) test to the specifications in columns (2) and (3) of Table 2. In the first step we regress log employment density on the set of instruments, which includes all the explanatory variables plus the log of the size of each Nuts 2 area, measured in squared kilometers, as the additional instrument. As discussed by Ciccone (2002), since the borders of Nuts 2 areas are set by administrative criteria, the size of the local area is a valid instrument, because it is correlated with density by construction but not correlated with employer-provided training, conditional on density.

[14] The negative impact of age of training emerges clearly in the last two columns of the table.

[15] Both the local unemployment rate and the average years of schooling attract statistically insignificant coefficients.

[16] The correlation between the average firm size in manufacturing and annual labor turnover in 1991 in Nuts 1 areas was 0.784.

Table 3
Probit estimates of the probability of employer-provided training, augmented with the residuals from the first step regression of log density on instruments

	(1)	(2)
Age	−0.009*** (0.002)	−0.009*** (0.002)
Age squared * 100	0.010*** (0.005)	0.009*** (0.003)
Gender	0.016* (0.008)	0.016* (0.008)
Log employment density in the Nuts 2 area	−0.023*** (0.005)	−0.115*** (0.043)
High school and college degree	0.067*** (0.008)	0.067*** (0.008)
Full time job	0.049*** (0.010)	0.049*** (0.009)
Permanent contract	0.096*** (0.010)	0.096*** (0.008)
Hired before 1991	−0.027** (0.011)	−0.027** (0.011)
Medium-sized firm	0.037** (0.016)	0.037*** (0.016)
Large-sized firm	0.081*** (0.008)	0.082*** (0.007)
Average years of schooling in the Nuts 2 area		0.004 (0.011)
Local unemployment rate		−0.135 (0.368)
Area-specific index of industrial specialization in the Nuts 2 area		−0.402* (0.213)
Average firm size in the Nuts 2 area		0.055* (0.030)
Average firm size * log density		0.031*** (0.012)
Residuals from first stage	0.004 (0.008)	0.006 (0.014)
Industry dummies	Yes	Yes
Occupation dummies	Yes	Yes
EU Objective 1 and 2 dummies	No	Yes
Number of observations	16,171	16,171
Pseudo R squared	0.120	0.121

Pooled cross-section time series data. Average partial effects. Dependent variable: T.
See Table 2.

The estimated coefficient of log size in the first step regression associated to the specification in the second column of Table 2 is equal to −1.272 and statistically significant–with a standard error of 0.006–which indicates that the additional instrument is not weak, according to the criteria discussed by Angrist and Krueger (2001)[17]. In the second step we add to the right hand side of (1) the residuals from the first step regression and verify whether they are significantly different from zero. Table 3 shows that they are not, which leads us to reject the hypothesis of no weak exogeneity of employment density[18].

The results in Tables 2 and 3 contrast with the positive correlation between local employment density and value added productivity found by Ciccone and Hall (1996) and Ciccone (2002) and suggest that the productivity gains associated to denser economic activity are not due to the fact that firms located in denser areas train more their employees. Ceteris paribus, firms in denser areas train less than firms in other areas. Following the discussion in Section 2 of this paper, we interpret this finding as evidence that the combination of pooling externalities and poaching effects generates a negative correlation between local economic density and employer-provided training.

This interpretation implies that turnover and poaching are higher in denser areas. One piece of evidence that labor mobility is higher in denser areas is that the correlation between labor turnover, as measured in the 1991 Employer's Manpower and Skill Practices Survey (see Martin, 1993), and log

[17] A similar estimate holds for the specification in the last column of Table 2.

[18] Under the null hypothesis of weak heterogeneity, the robust standard errors need not be adjusted further for the presence of generated regressors. We are grateful to Guglielmo Weber for advice on this point.

Table 4
Probit estimates of the probability of voluntary turnover

	(1)
Hired before 1991	−0.069*** (0.005)
Gender	−0.020*** (0.006)
High school and college degrees	0.014** (0.006)
Full time	0.043*** (0.008)
Permanent contract	−0.021 (0.016)
Training in the previous period	0.002 (0.009)
Training in the previous period * high density	0.029**(0.013)
Number of observations	9854
Pseudo *R* squared	0.052

Pooled cross-section time series data. Average partial effects. Dependent variable: dummy equal to 1 in the event of voluntary turnover and to 0 otherwise.
See Table 2. The number of observations is lower than in Table 2 because of the inclusion of lagged training.

employment density in British Nuts 1 areas is positive and equal to 0.49. Another piece is that the correlation between the percentage of unfilled skilled vacancies on total local employment and log employment density in Nuts 1 areas is negative and equal to −0.66[19]. The fact that denser areas have relatively fewer unfilled skilled vacancies as a percentage of local employment suggests that firms in these areas have less pressure to train employees because of the difficulties encountered in hiring the required skills from the market.

If poaching is higher in denser areas, we should find that in these areas employer-provided training has a positive effect on voluntary mobility. Our data set provides information on whether an individual has changed job in the reference period, defined as the year of the interview or the year immediately before, to obtain a better or more suitable job. We estimate a probit model, which associates the probability of turnover to individual characteristics, individual tenure and employer-provided training in the year before the reference period. In the estimates reported in Table 4 we define the dummy "high density" as equal to one if the local area has density higher or equal to median density and to zero otherwise, and interact this dummy with previous training. We find that the impact of previous training on turnover is not statistically different from zero, but that the coefficient associated to the interaction term is both statistically significant and positive. We interpret this as evidence that the effect of employer-provided training on voluntary turnover is positive in denser areas, which are more exposed to poaching effects.

The uncovered negative relationship between density and employer-provided training could be explained if this type of training and the training undertaken by the employee or provided by the local government are substitutes and the latter type of training is more frequent in denser areas. To check this, we have computed for each area and year the percentage of trained individuals–employed or not–who have been involved in training that was not employer-provided, and added this variable to the right hand side of (1). As shown in the first two columns of Table 5, this percentage attracts a positive and statistically significant coefficient, and its inclusion changes the estimated coefficient of log employment density only marginally. We conclude that areas where training decided by employees or provided by the government is high have also high employer-provided training. In the last two columns of the table, we add to the regressors the lagged dependent variable, to take into account the time persistency of training. The results show that the negative and statistically

[19] The data on unfilled vacancies by occupation are from the website www.nomisweb.co.uk. We classify as skilled the vacancies for managers, professional, technicians and craft workers.

Table 5
Probit estimates of the probability of employer-provided training

	(1)	(2)	(3)	(4)
Age	−0.009*** (0.002)	−0.009*** (0.002)	−0.007** (0.003)	−0.007** (0.003)
Age squared * 100	0.010*** (0.005)	0.009*** (0.003)	0.006 (0.004)	0.006 (0.004)
Gender	0.016* (0.008)	0.016* (0.008)	0.007 (0.010)	0.007 (0.010)
Log employment density in the Nuts 2 area	−0.021*** (0.005)	−0.103*** (0.034)	−0.014** (0.005)	−0.132*** (0.048)
High school and college degree	0.068*** (0.008)	0.067*** (0.008)	0.084*** (0.010)	0.085*** (0.010)
Full time job	0.049*** (0.010)	0.049*** (0.009)	0.043*** (0.014)	0.045*** (0.014)
Permanent contract	0.096*** (0.010)	0.096*** (0.008)	0.071*** (0.019)	0.069*** (0.019)
Hired before 1991	−0.027** (0.011)	−0.027** (0.011)	−0.036*** (0.010)	−0.036*** (0.010)
Medium-sized firm	0.037** (0.016)	0.037*** (0.016)	0.025* (0.016)	0.025* (0.016)
Large-sized firm	0.081*** (0.008)	0.081*** (0.007)	0.069*** (0.009)	0.070*** (0.009)
Average years of schooling in the Nuts 2 area		−0.003 (0.012)		−0.013 (0.015)
Local unemployment rate		−0.109 (0.265)		−0.224 (0.405)
Area-specific index of industrial specialization in the Nuts 2 area		−0.306 (0.233)		−0.615** (0.261)
Average firm size in the Nuts 2 area		0.050 (0.032)		0.036 (0.039)
Average firm size * log density		0.027** (0.011)		0.040** (0.015)
Percentage of trained individuals–employed or not–not trained by employer	0.820** (0.380)	1.001** (0.404)		
Lagged dependent variable			0.424*** (0.020)	0.423*** (0.020)
Industry dummies	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes
EU Objective 1 and 2 dummies	No	Yes	No	Yes
Number of observations	16,171	16,171	10,432	10,432
Pseudo R squared	0.120	0.121	0.248	0.250

Pooled cross-section time series data. Average partial effects. Dependent variable: T. Columns (1)–(2) include among the regressors the percentage of trained individuals–employed or not–not trained by the employers; columns (3)–(4) add the lagged dependent variable.
See Table 2. The reduction in the number of observations in the last two columns is due to the lagged dependent variable.

significant effect of log employment density on training remains. It is true that the introduction of the lagged dependent variable reduces the size of the impact of log density on training, but the long-term effect remains virtually unchanged[20].

[20] The long-term effect of density in column (3) of the table is −0.024 (−0.014/0.576).

Table 6
Probit estimates of the probability of employer-provided training

	(1)	(2)
Age	−0.006** (0.002)	−0.006** (0.002)
Age squared * 100	0.005* (0.003)	0.005* (0.003)
Gender	0.0108 (0.008)	0.0108 (0.008)
High school and college degrees	0.075*** (0.008)	0.075*** (0.008)
Full time	0.044*** (0.010)	0.044*** (0.010)
Permanent contract	0.093*** (0.013)	0.093*** (0.013)
Hired before 1991	−0.024*** (0.008)	−0.024*** (0.008)
Medium-sized firm	0.038*** (0.012)	0.038*** (0.012)
Large-sized firm	0.079*** (0.007)	0.079*** (0.007)
Area and industry specific index of industrial specialization	−0.107** (0.049)	−0.127 (0.094)
Area and industry specific index of industrial diversity		0.0006 (0.003)
Regional Nuts 2 dummies	Yes	Yes
Number of observations	13,347	13,347
Pseudo *R* squared	0.118	0.118

Pooled cross-section time series data. Average partial effects. With measures of industrial specialization and diversity. Dependent variable: *T*.
Each regression includes a constant, year and occupational dummies. Cluster adjusted robust standard errors. One, two and three stars when the coefficients are significantly different from zero at the 10%, 5% and 1% level of confidence. The number of observations is lower than in Table 2 because we only retain in the regressions year by industry by area clusters with at least 5 observations.

In the literature on local agglomeration effects and on the economics of cities, much emphasis has been placed on the concepts of MAR and Jacobs externalities. As discussed in the review of the literature, these concepts capture within-area industry specific agglomeration effects. The

Table 7
Multinomial logit estimates of the probability of training

	Employer-provided training	Other training
Age	−0.056*** (0.018)	−0.397*** (0.031)
Age squared * 100	0.058*** (0.002)	0.482*** (0.004)
Gender	0.103* (0.059)	0.227*** (0.073)
Log employment density in the Nuts 2 area	−0.147*** (0.037)	−0.038 (0.063)
High school and college degree	0.438*** (0.058)	0.556*** (0.075)
Full time	0.368*** (0.080)	−0.553*** (0.131)
Permanent contract	0.773*** (0.081)	−0.712*** (0.090)
Hired before 1991	−0.186** (0.086)	−0.268*** (0.086)
Medium-sized firm	0.243** (0.021)	0.228 (0.183)
Large-sized firm	0.527*** (0.051)	0.225* (0.132)
Industry dummies	Yes	Yes
Occupation dummies	Yes	Yes
EU Objective 1 and 2 dummies	No	No
Number of observations	16,770	16,770
Adjusted *R* squared	0.131	0.131

Pooled cross-section time series data. No training as the base outcome.
See Table 2. The number of observations is higher than in Table 2 because we include in the data individuals with training not provided by the employer.

former is an industry-specific index of industrial specialization, computed as (see Combes, 2001)

$$\mathrm{MAR}_{kjt} = \frac{E_{kjt}}{E_{jt}} \tag{3}$$

and the latter an industry-specific index of industrial diversity, computed as

$$J_{kjt} = 1/\left(\sum_{y \neq k}\left[\frac{E_{\gamma jt}}{E_{jt}}\right]^2\right) \tag{4}$$

Compared to the index S, which varies by region, indices MAR and J vary both by region and by industry. We test whether these indices affect employer-provided training by estimating the following version of (1)

$$\mathrm{Prob}\{T_{ijt} = 1\} = \Phi\{\beta X_{ijt} + \gamma RD_j + \delta Z_t + \theta W_{ikt} + \varepsilon_{ijt}\} \tag{5}$$

where W is a vector which includes MAR and Jacobs externalities. In this specification we exploit the fact that the indicators in W vary by region, time and industry and control for unobserved area

Table 8
Probit estimates of the probability of employer-provided training

	(1)	(2)	(3)
Age	−0.009*** (0.002)	0.001 (0.005)	−0.062*** (0.010)
Age squared*100	0.010*** (0.005)	−0.003 (0.006)	0.070*** (0.014)
Gender	0.026** (0.012)	0.015 (0.010)	0.081** (0.036)
Log employment density in the Nuts 2 area	−0.097*** (0.033)	−0.125*** (0.042)	−0.326** (0.130)
High school and college degree	0.052*** (0.011)	0.075*** (0.010)	0.342*** (0.031)
Full time	0.052*** (0.012)	0.064*** (0.013)	0.251*** (0.044)
Permanent contract	0.099*** (0.016)	0.098*** (0.014)	0.569*** (0.067)
Hired before 1991	−0.039** (0.016)	−0.036*** (0.012)	−0.091* (0.047)
Medium-sized firm	0.043* (0.026)	0.038** (0.016)	0.150*** (0.058)
Large-sized firm	0.111*** (0.012)	0.090*** (0.008)	0.318*** (0.033)
Average years of schooling in the Nuts 2 area	0.015 (0.012)	0.015 (0.013)	0.021 (0.051)
Local unemployment rate	−0.436 (0.406)	−0.332 (0.372)	−1.187 (1.077)
Area-specific index of industrial specialization in the Nuts 2 area	−0.327 (0.292)	−0.416* (0.249)	−1.623* (0.956)
Average firm size in the Nuts 2 area	0.034 (0.033)	0.080** (0.039)	0.111 (0.137)
Average size*log density	0.027*** (0.010)	0.036** (0.014)	0.087* (0.048)
Industry dummies	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes
EU Objective 1 and 2 dummies	Yes	Yes	Yes
Number of observations	9836	12,946	15,107
Adjusted R squared	0.091	0.133	0.091

Pooled cross-section time series data. Average partial effects in the former two columns. Robustness checks: (1) 1994–1997 only; (2) age 25 to 54 only; (3) training duration. Dependent variable: T.
See Table 2. The third column reports the coefficients of the ordered probit estimates of training duration, not the average partial effects. The number of observations in the last column is lower than in Table 2 because of missing values.

effects with Nuts 2 dummies. Table 6 shows that employer-provided training is significantly lower when industrial specialization is higher, which confirms our previous results. Conditional on specialization, we also find a positive but not statistically significant impact of industrial diversity. As discussed in Section 2 of the paper, industrial specialization can affect both local pooling and poaching effects and turnover effects. Our results suggest that poaching and turnover effects are stronger and/or the pooling effects weaker when the local industrial structure is more specialized.

7. Robustness

In this section we investigate the robustness of our results. First, we redefine the dependent variable T by assigning the value 0 to no training, 1 to employer-provided training, and 2 to other training, and estimate the specification in the second column of Table 2 with a multinomial logit. The results in Table 7 confirm the negative and statistically significant relationship between employer-provided training and density and the lack of such relationship for training not provided by the employer.

We also check whether changes in sample size and in the definition of employment density affect our key results. Table 8 replicates our estimates of the least parsimonious model in Table 2 on the sub-sample covering the years 1994–1997 (column (1) in the table); on the sub-sample of individuals aged 25 to 54 (column (2)); by using training duration as the dependent variable (column (3)). We exclude the years 1998–2000 in the first exercise because of a change in the wording of the question on training in the BHPS after 1997. We remove individuals aged between 17 and 24 because the training of this group is likely to include both initial vocational training as well as continuing training, which is typical of the older age group (see Arulampalam et al., 2004). Duration is an alternative measure of training. Since this variable is ordered in the range (0.3)[21], we use an ordered probit model. The results in Table 8 show that the sign of the relationship between log employment density and employer-provided training is robust to changes in the sample and in the definition of the dependent variable[22].

In an additional experiment, we add average productivity in the Nuts 2 area to the set of variables in the vector Y, as a further control for the local knowledge stock. The estimated coefficient turns out to be positive–but seldom statistically significant–in most specifications, and the relationship between local density and training remains negative and statistically significant[23].

Next, we experiment with alternative definitions of log employment density, our key explanatory variable. We have computed employment in the private non-agricultural sector at the Nuts 2 level by using the BHPS distribution of employment by local area to disaggregate private non-agricultural national employment. An alternative procedure is to use these weights to disaggregate Nuts 1 employment. We have done so, with no qualitative change of results. The measure of density used in the paper does not distinguish between skilled and unskilled employment, in line with the existing literature. One could argue, however, that the source of pooling externalities as well as of poaching effects is skilled rather than total employment. We have restricted our measure of local density to skilled employment, which we identify with the following occupations: managers, professionals,

[21] Duration is coded as 0 for no training, 1 for training lasting less than 2 weeks, 2 for training lasting from 2 to 9 weeks and 3 for training lasting longer than 9 weeks.

[22] The estimated effect in the first two columns of the table of a 10% increase in density on training is equal to −0.07 and −0.04 respectively.

[23] We are grateful to the Editor for suggesting this experiment. Results are available from the authors upon request.

technicians and craft workers. Again, we find that the relationship between density and employer-provided training is robust to these changes in the definition of density.

We have identified local labor markets with groups of counties, the Nuts 2 classification of regional areas, because this classification is wide enough to contain most relevant travel to work areas but not too large to determine the dissipation of pooling externalities. One potential problem here is that individuals who reside near the border of a group of counties could be employed across the border, in another group of counties. Furthermore, as argued by Ciccone (2002), there is no strong reason to believe that spatial externalities do not involve neighboring regions. We deal with these problems as follows. First, we replace density in each Nuts 2 area of residence with the average of this density and the density of neighboring regions, which share their borders with the area. By so doing, we are able to minimize the impact of any mismatch between area of residence and area of work, which remains after choosing a reasonably wide reference area, the group of counties. The results in the first column of Table 9 suggest that the negative relationship between employer-provided training and density is robust. Second, we augment the least parsimonious specification in Table 2 with an additional measure of density, the average employment density of neighboring areas, obtained by averaging the densities of the areas which share borders with each Nuts 2 region. The results in the second column of Table 9 show that both measures of density attract a negative and statistically significant coefficient. We find this result reassuring, because the negative correlation between employer-provided training and employment density is not affected by eventual misallocations of individuals to the relevant region of employment.

Table 9
Probit estimates of the probability of employer-provided training

	(1)	(2)
Age	−0.009*** (0.002)	−0.009*** (0.002)
Age squared * 100	0.009*** (0.003)	0.009*** (0.003)
Gender	0.016* (0.008)	0.016* (0.008)
Log employment density in the Nuts 2 area	−0.169*** (0.049)	−0.077*** (0.035)
Log employment density in the neighboring area	–	−0.019*** (0.006)
High school and college degree	0.067*** (0.008)	0.067*** (0.008)
Full time job	0.049*** (0.009)	0.049*** (0.009)
Permanent contract	0.096*** (0.008)	0.096*** (0.008)
Hired before 1991	−0.027** (0.011)	−0.027** (0.011)
Medium-sized firm	0.037*** (0.016)	0.037*** (0.016)
Large-sized firm	0.082*** (0.007)	0.082*** (0.007)
Average years of schooling in the Nuts 2 area	0.012 (0.011)	0.010 (0.012)
Local unemployment rate	−0.259 (0.257)	−0.298 (0.278)
Area-specific index of industrial specialization in the Nuts 2 area	−0.406** (0.204)	−0.448** (0.218)
Average firm size in the Nuts 2 area	0.067* (0.034)	0.030 (0.031)
Average firm size * log density	0.046*** (0.016)	0.020*** (0.010)
Industry dummies	Yes	Yes
Occupation dummies	Yes	Yes
EU Objective 1 and 2 dummies	Yes	Yes
P-value of the F-test for the inclusion of confounding area specific variables	0.000	0.000
Number of observations	16,171	16,171
Pseudo R squared	0.121	0.121

Pooled cross-section time series data. Average partial effects. (1) With average density; (2) with two measures of density—local and neighboring area. Dependent variable: T.
See Table 2.

8. Conclusions

The key finding of this paper is that employer-provided training in the UK is less frequent in economically denser areas. We have explained this result by arguing that poaching and turnover effects of agglomeration prevail on pooling effects. The size of their effect is not negligible: when evaluated at the average firm size in the area, a 1% increase in density reduces the probability of employer-provided training by 0.014, close to 4% of the average incidence of training in the UK during the sample period.

In a well-known paper, Ciccone and Hall (1996) find that higher density increases average productivity in the area by 5%. Our results suggest that this effect could have been even higher were it not for the negative impact of density on employer-provided training. Higher density affects productivity both directly, by facilitating the creation and diffusion of innovation, and indirectly, by affecting the composition of labor in the local area. Denser areas attract individuals with higher education, who are more productive and learn new skills faster. Faster learning encourages training. The same areas, however, are characterized by higher labor mobility, which reduces the incentive of firms to train. Overall, productivity can be higher in denser areas despite the fact that employer-provided training is lower.

Acknowledgements

We are grateful to an Editor and two anonymous referees, to Andrea Bassanini, Federico Cingano, Piero Cipollone, Maria De Paola, Peter Dolton, Luigi Guiso, Winfried Koeninger, David Jaeger, Patrizia Ordine, Gianmarco Ottaviano, Guglielmo Weber and the audiences in seminars at the Bank of Italy Research Department, the University of Calabria, Ente Luigi Einaudi, IZA Bonn and Paris II for comments and suggestions. This research was partly funded with a national research grant by the Italian Ministry of Research (MIUR). The usual disclaimer applies.

References

Acemoglu, D., 2002. Technical change, inequality and the labor market. Journal of Economic Literature 40, 7–72.

Acemoglu, D., Pischke, J., 1998. Why do firms train? Theory and evidence. The Quarterly Journal of Economics 113, 79–119.

Acemoglu, D., Pischke, J., 1999. The structure of wages and investment in general training. Journal of Political Economy 107, 539–572.

Almazan, A., de Motta, A., Titman, S., 2003. Firm location and the creation and utilization of human capital. NBER Working Paper, vol. 10106.

Angrist, J., Krueger, A., 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. Journal of Economic Perspectives 15, 69–85.

Arulampalam, W., Booth, A., Bryan, M., 2004. Training in Europe. Journal of the European Economic Association 2, 346–360.

Audretsch, D.B., Feldman, M.P., 2004. Knowledge spillovers and the geography of innovation. In: Thisse, J.F., Henderson, J.V. (Eds.), Handbook of Regional and Urban Economics, vol. 4. Cities and geography (North Holland, Amsterdam), pp. 2713–2739.

Bassanini, A., Brunello, G., 2003. Is training more frequent when wage compression is higher? Evidence from the European Community Household Panel. IZA Discussion Paper, vol. 839.

Becker, G., 1964. Human Capital. The University of Chicago Press, Chicago.

Blundell, R., Smith, R., 1986. An exogeneity test for a simultaneous equation tobit model with an application to labor supply. Econometrica 54, 679–685.

Brunello, G., De Paola, M., 2004. Training and the density of economic activity: Evidence from Italy. IZA Discussion Paper, vol. 1173.

Brunello, G., Medio, A., 2001. An explanation of international differences in education and workplace training. The European Economic Review 45, 307–322.

Ciccone, A., 2002. Agglomeration effects in Europe. The European Economic Review 46, 213–227.

Ciccone, A., Hall, R., 1996. Productivity and the density of economic activity. American Economic Review 86, 54–70.

Cingano, F., Schivardi, F., 2004. Identifying the sources of local productivity growth. Journal of European Economic Association 2, 720–742.

Combes, P., 2001. The economic structure and local growth: France 1984–1993. Journal of Urban Economics 47, 329–355.

Combes, P., Duranton, G., 2006. Labor pooling, labor poaching and spatial clustering. Regional Science and Urban Economics 36, 1–28.

Glaeser, E., Maré, D., 2001. Cities and skills. Journal of Labor Economics 19, 316–342.

Glaeser, E., Kallal, H., Scheinkman, J., Shleifer, A., 1992. Growth in cities. Journal of Political Economy 100, 1126–1152.

Henderson, V., Kuncoro, A., Tuner, M., 1995. Industrial development in cities. Journal of Political Economy 103, 1067–1090.

Krugman, P., 1991. Geography and Trade. MIT Press.

Lynch, L., 2003. Development intermediaries and the training of low wage workers, mimeo, Tufts University.

Martin, C., 1993. Labor Turnover: Evidence from UK establishments, Brunel University, mimeo.

Moretti, E., 2004. Human capital externalities in cities. In: Thisse, J.F., Henderson, J.V. (Eds.), Handbook of Regional and Urban Economics, vol. 4. Cities and geography (North Holland, Amsterdam), pp. 2243–2291.

Neal, D., 1995. Industry-specific human capital: Evidence from displaced workers. Journal of Labor Economics 13, 653–677.

OECD, 2003. Employment Outlook. Paris.

Rosenthal, S., Strange, W., 2003. Geography, industrial organization and agglomeration. The Review of Economics and Statistics 85, 377–393.

Rosenthal, S., Strange, W., 2004. Evidence on the nature and sources of agglomeration economies. In: Thisse, J.F., Henderson, J.V. (Eds.), Handbook of Regional and Urban Economics, vol. 4, Cities and geography (North-Holland, Amsterdam), pp. 2119–2171.

Rotemberg, J.J., Saloner, G., 2000. Competition and human capital accumulation: A theory of interregional specialization and trade. Regional Science and Urban Economics 30, 373–404.

Stevens, M., 1994. An investment model for the supply of training by employers. The Economic Journal 104, 408–419.

Yatchew, A., Griliches, Z., 1985. Specification errors in probit models. Review of Economics and Statistics 67, 134–139.

Wooldridge, J., 2002. Econometric analysis of cross section and panel data. The MIT Press, Boston.