UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche “Marco Fanno”

TRAINING AND ECONOMIC DENSITY: SOME EVIDENCE FROM ITALIAN PROVINCES

GIORGIO BRUNELLO
Università di Padova, IZA e Cesifo

MARIA DE PAOLA
Università della Calabria

November 2006

“MARCO FANNO” WORKING PAPER N.30

Training and Economic Density: Some Evidence from Italian Provinces^

Giorgio Brunello (University of Padova, IZA and Cesifo)

Maria De Paola* (University of Calabria)

Forthcoming in *Labour Economics*

* Corresponding author: Department of Economics and Statistics, University of Calabria, via P. Bucci, 87036 Arcavacata di Rende, Italy; e-mail: m.depaola@unical.it

^We are grateful to Juan Dolado, two anonymous referees, Giovanni Costa, Francesca Gambarotto, Vincenzo Scoppa, and the audiences at seminars in Modena, Rome and San Francisco (EALE-SOLE 2005) for comments and suggestions. The usual disclaimer applies.

Abstract

In this paper we use a search and matching model to investigate the economic relationship between training and local economic conditions. We identify two aspects of this relationship going in opposite directions: on the one hand, the complementarity between local knowledge spillovers and training generates a positive correlation between training and local density; on the other hand, higher wages and labor turnover in denser areas reduce training. Overall the relationship can be either positive or negative, depending on the relative strength of these two effects. Our empirical analysis, based on a sample of Italian firms, shows that training is lower in provinces with higher labor market density, measured as the number of employees per squared kilometer.

Key words: training, local labor markets, Italy

JEL Code: J24, R12

Introduction

The productivity gains associated to local economic density are documented by an increasing body of empirical research (Henderson, 1986; Ciccone and Hall, 1996: Ciccone, 2001). Understanding the sources of these gains is important for policy, especially in the light of the Lisbon strategy, which aims at making of Europe a highly competitive and productive region of the world. One source identified by the literature is the positive spatial externalities associated to the physical proximity of workers and firms, which more than offset the negative congestion effects originated by the intense use of capital and labor (Ciccone and Hall, 1996). An additional channel linking density to productivity operates via training and its positive influence on productivity[1]: on the one hand, denser areas encourage firms to invest in human capital, because knowledge spillovers are better exploited by skilled workers, and trained employees are more productive. On the other hand, density increases wages and turnover, which discourage training investments and reduce productivity. If the former effect prevails on the latter, higher density is associated to more training, and thus to higher productivity.

The purpose of this paper is to explore this second channel, both theoretically and empirically. We believe that the study of the relationship between density and training is important because it helps us understand how local agglomeration patterns influence productivity. The complementarity between knowledge spillovers, innovation and skills has already been emphasized in the relevant literature. Acemoglu, 2002, for instance, argues that the ability of each firm to adapt new technologies and ideas developed by other firms is strictly related to the skills of its own labor force. On the empirical side, Moretti, 2004, finds that productivity gains from human capital spillovers are relevant for manufacturing establishments in the US.

Skills after labor market entry are generated by training and experience. Training is important both because it increases the ability to perform well the relevant task and because it improves the ability to understand and process the flow of information from the productive environment, and to translate this information into higher productivity on the job (Jaffe et al. 1993; Anselin, et al. 1997). The complementarity between skills and local knowledge spillovers suggests that firms located in denser areas have stronger incentives to invest in training: by increasing the skills of the labor force,

[1] The empirical evidence showing that training increases productivity is reviewed by Bassanini et al., 2005. In this paper, we restrict our attention to employer-provided training.

firms benefit more from the positive externalities associated to density and are more productive.

This complementarity, however, is not sufficient to generate a positive relationship between local economic density and training incidence, because denser areas have also higher wages (see Glaeser and Mare, 2001, and Ciccone and Cingano, 2003, for a review) and higher labor turnover, which discourage investment in training[2]. It follows that the relationship between local agglomeration effects, measured by local employment density, and employer-provided training, can take either sign. In their empirical analysis of UK data, Brunello and Gambarotto, 2004, show that the balance of positive and negative effects is tilted in favor of the latter, and that training incidence is lower, ceteris paribus, in denser economic areas.

In this paper, we present a model of employer-provided training in search equilibrium which illustrates the steady – state relationship between training, density and productivity in local labor markets. The model also provides guidance to the empirical investigation which follows. We use data on more than 1000 Italian manufacturing firms, drawn from the Survey of Italian Manufacturing (*Indagine sulle Imprese Manifatturiere*) conducted by Mediocredito Centrale, and estimate the relationship between training incidence, measured as the percentage of trained employees in each sampled firm during the year of reference, and local labor market density, where the local market is identified with the province.

We find that the estimated relationship is negative and statistically significant. While the nature of the data at hand warn us against easy generalizations, we confirm the qualitative findings obtained by Brunello and Gambarotto, 2004, for a different country and with a different dataset. The combined evidence of a negative relationship between training and local density and a positive relationship between productivity and density suggests that the productivity gains to geographic proximity are not attained because firms in denser areas train more their employees. Since higher density influences productivity both directly – by stimulating innovation – and indirectly – by altering skills, our results suggest that the uncovered relationship between density and productivity would be even stronger in the absence of the negative impact of density on training.

[2] The effects of labor market conditions on training have been studied in the relevant literature. Acemoglu and Pischke, 1999, for instance, argue that firms operating in local labor markets characterized by high unemployment rates have more incentives to invest in training, because worker have lower bargaining power. An alternative view is that high local unemployment, by increasing the availability of skilled employees in the local labor market, reduces matching costs and the incentives to train (see Brunello and Medio, 2001; Stevens, 1996).

One of the advantages - identified by authors such as Marshall, Arrow and Romer - which accrue to firms locating near other producers in the same industry is that geographic proximity helps spreading information and the exchange of ideas, the discussion of solutions to problems, and the awareness of other important information (Feldman, 1993). In Marshall's words "..The mysteries of the trade become no mysteries, but are as they were in the air". While this view focuses on the intra-industry transmission of knowledge, according to Jacobs, 1969, the most important source of knowledge spillovers is external to the industry in which the firm operates, and knowledge externalities are especially promoted by the variety of the local economic system[3].

Italian small and medium firms have often been scrutinized because of the important role played by industrial districts, or clusters of firms involved in the production of homogeneous goods. We find that firms operating in such districts, and belonging to the same industrial sector characterizing the district, invest more in training. On the other hand, specialization, measured by the ratio of employment in the own industry and area and employment in the area, does not seem to have any significant additional effect on training. We interpret these results as evidence that, conditional on local density, the production of skills is favored by the marked cooperative behavior typical of industrial districts (see Brusco, 1982, and Becattini et al., 1990), which reduces the risk of poaching and increases the returns to training.

Turning to the policy implications, the natural question to ask is whether institutional design can affect the negative impact of local density on training. Our results suggest that the development of institutions that foster the combination of competition and cooperation, a typical feature of industrial districts, can help reducing the risk of poaching and the negative congestion effects of local agglomeration, and by so doing promote training and productivity.

The paper is organized as follows. Section 1 illustrates the theoretical model. Section 2 describes the data used in the empirical analysis and provides some descriptive statistics. In Section 3 the empirical model is specified. The main results are discussed in Section 4, and a few extensions are presented in Section 5. Conclusions follow.

[3] On the relevance of Marshall-Arrow-Romer and Jacobs externalities see also Glaeser et al., 1992, Henderson, et al., 1995, Cingano and Schivardi 2004,

1. Training and local economic density

We find it convenient to model the relationship between training and economic density in two steps: first, in sub-sections 1.1 to 1.6 we develop the basic model by considering a single local labor market with a given number of workers and firms; second, in sub-section 1.7 we allow workers and firms to move between local labor markets and explore how results should be amended with respect to the basic framework.

1.1 The Setup

We start by considering a local labor market populated by a large number of identical and risk neutral firms and workers, who discount the future at the common rate *r*. Firms cannot change their location and worker flows are only between the single market and inactivity. Workers and firms match in a market characterized by frictions (Pissarides, 2000; Mortensen and Pissarides, 1999; Blanchard and Diamond, 1989). To be productively employed, each unskilled worker needs workplace training, which is assumed to take place at the start of the employment spell.

Training is general and increases the worker's productivity both in the incumbent firm and in other firms. When the labor market is imperfectly competitive, firms may be willing to pay training costs (see Acemoglu and Pischke, 1999). Here we assume, for the sake of simplicity, that these costs are borne entirely by the employer[4]. Training increases individual productivity for two reasons. First, the employee increases her skill in performing the relevant job; second, and most important for the purposes of this paper, 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 in higher productivity on the job.

While the first effect is standard, the second effect is based on the view that skills and knowledge spillovers are complements (Rosenthal and Strange, 2004), which suggests that the positive externalities of agglomeration are better exploited by skilled labor. Since knowledge spillovers are associated with geographic proximity - see Ciccone and Hall, 1996 - the intensity of these externalities is bound to increase with the density of economic activity in the local labor market.

[4] Most training is employer-provided, as shown by Bassanini et al., 2005, in a comparative perspective.

With a perfect capital market and conditional on training, the asset value of a job filled by a trained worker, $F_E^T(t)$, satisfies the following Bellman equation

$$rF_E^T = y(t,s) - w(t,s) + (q+j)(F_v - F^T{}_E) \qquad [1]$$

where *y* is productivity, t is training (intensity), σ is local economic density, *w* are wages, F_V is the asset value of a vacant job slot, *q* is the rate of exogenous job separation and j represents the probability of becoming inactive.

The filled job yields a net income flow equal to productivity minus the wage paid to the worker. Productivity is influenced by knowledge spillovers deriving from economic density, $y_s \geq 0$[5]. Both wages and productivity are positively affected by training[6]. In addition, we characterize the complementarity of skills and knowledge spillovers as stronger in denser economic areas by assuming $y_{ts} \geq 0$. Since firms are identical and the wage rate does not vary across firms, there is no endogenous turnover. With exogenous probability $q+j$ the worker – firm pair separates and a job vacancy is created. The associated change of state yields a loss equal to $F_V - F_E^T$. The asset value of a vacant job satisfies the condition

$$rF_V = -d + h\left[pF_E^T + (1-p)(F_E^T - c(t)) - F_V\right] \qquad [2]$$

The vacant job costs *d* per unit of time and changes state according to a Poisson process with rate *h*, the probability that the firm finds a new worker[7]. If the firm meets a skilled worker with probability *ph*, where *p* is the proportion of skilled workers in the unemployment pool, the asset value of the match is F_E^T. If it meets an unskilled worker, with probability *(1-p)h*, the new match is worthless unless the firm sinks the training cost $c(t)$ after the match. The investment in training transforms the unskilled employee into a perfect substitute of any skilled worker. Therefore, the value of the match to the firm in such event is $F_E^T - c(t)$.

---

[5] We denote the partial first and second derivatives of a function $Z(x,y)$ as Z_x and Z_{xy}.

[6] Hence we have $y_t(t,s) \geq 0$ and $w_t(t,s) \geq 0$. A review of the evidence on the effects of training on wages is Bassanini et al, 2005.

[7] Search is undirected, as in Albrecht and Vroman, 2002.

Since all profit opportunities for new jobs are exploited in equilibrium, the free-entry condition is $F_V = 0$. Substituting this condition in [2] yields

$$F_E{}^T = \frac{d}{h} + c(t)(1-p) \tag{3}$$

1.2 Wage Determination

The wage is determined after training by Nash bargaining between the firm and the worker. Assuming that the parties have equal bargaining power $b = \frac{1}{2}$ and that the asset value of inactivity is zero, the expected return to being employed is equal to

$$rE(t) = w(t) + q[U(t) - E(t)] - jE(t) \tag{4}$$

where *E* and *U* are the asset values of employment and unemployment respectively. This return is equal to the wage gained in the current period plus the expected loss from separation, which occurs at the rate $q+j$. Letting unemployment benefits be equal to zero, the asset value of unemployment for a skilled worker with training t is

$$rU(t) = f[E(t) - U(t)] - jU(t) \tag{5}$$

where *f* is the hazard rate from unemployment. Using expressions [4] and [5], the surplus is equal to

$$E(t) - U(t) = \frac{w(t)}{r+q+f+j} \tag{6}$$

Taking into account the zero profit condition, the surplus from a filled job is

$$F_E{}^T - F^V = \frac{y(t,s) - w(t)}{r+q+j} \tag{7}$$

The Nash bargaining solution and the hypothesis of equal bargaining power imply that the wage is chosen to split equally the total surplus. Therefore, $E(t)-U(t)=F_E{}^T(t)-F_V$ and

[8] $$w(t,s,f)=\frac{[y(t,s)](r+q+f+j)}{2(r+q+j)+f}$$

1.3 Training

Firms invest in training to maximize $p=F_E{}^T-c(t)$, or

[9] $$p=\frac{y(t,s)}{2(r+q+j)+f}-c(t)$$

where $c_t(t)>0, c_{tt}(t)>0, c_t(0)=0$. The optimal level of training t^* satisfies the following first order condition:

[10] $$\frac{y_t(t^*,s)}{2(r+q+j)+f}-c_t(t^*)=0$$

Training is positive when $2(r+q+j)+f<\infty$. In the following sub-sections we investigate how training is affected by local economic density s.

1.4 Equilibrium

All firms in the economy share the same matching technology, which takes the following form

[11] $$x(u,v)=mu^{1-a}v^a$$

where x is the number of matches divided by the labor force, L, m is a parameter indicating the efficiency of matching, u denotes the unemployment rate, v the ratio of vacancies to the labor force and a is a parameter smaller than one. With this technology we have:

[12] $$h = \frac{x}{v} = m\left(\frac{v}{m}\right)^{a-1} = mq^{a-1}$$

[13] $$f = \frac{x}{u} = m\left(\frac{v}{m}\right)^{a} = mq^{a}$$

where θ is the ratio of vacancies to unemployment, our measure of labor market tightness.

The labor flows characterizing this economy can be described as follows: employed skilled labor quits at the rate *q* and becomes inactive at the rate j ; skilled unemployed labor either becomes inactive at rate j or finds a job at the constant rate *f*. Finally, unskilled labor either becomes inactive or fills in jobs at the rate *f*. The constant outflow from the local labor market needs to be replenished with fresh resources entering the market. With a single labor market, new entries are unskilled and enter only via the unemployment pool at the rate j [8]. They match and become skilled employees after receiving training.

The flow equations which determine the equilibrium values of the unemployment rate *u* and of the proportion of trained labor in the unemployment pool, *p*, are

[14] $$(q + j)(1 - u) = fu$$

[15] $$q(1 - u) = pu(f + j)$$

Combining these equations we get

[16] $$u = \frac{q + j}{q + j + f}$$

[17] $$p = \frac{qf}{(q + j)(f + j)}$$

Notice that $p_q > 0$: when the labor market tightens, more skilled labor flows into the unemployment pool, because employment is higher, and is partly replaced by unskilled labor, because search is undirected. Therefore, the percentage of skilled workers in the unemployment pool increases. The equilibrium value of the vacancy-unemployment ratio q is obtained by using [3] in [1], which yields

[8] With more than one local labor market, new entries in a single market can also be skilled workers moving in from another local market. See sub-section 1.7 for a discussion.

[18] $$w(t,s,q) = y(t,s) - (r+q+j)\left[\frac{d}{h(q)} + c(t)(1-p(q))\right]$$

In equilibrium [18] must be equal to the wage equation [8]. Using [12], [13] and [17], we can write this equality in implicit form as

[19] $$A(t,q,s) = y(t,s) - [2(r+q+j) + f(q)]\left[\frac{d}{h(q)} + c(t)(1-p(q))\right] = 0$$

Similarly, the first order equation with respect to training can be written as

[20] $$B(t,q,s) = y_t(t,s) - [2(r+q+j) + f(q)]c_t(t) = 0$$

where *B* is the training equation in implicit form. Equations [19] and [20] determine the equilibrium values of t and ? as functions of s and other exogenous parameters.

1.5 Comparative Statics

Differentiation of [19] and [20] yields

[21] $$\begin{aligned} A_t \partial t + A_q \partial q &= -y_s \partial s \\ B_t \partial t + B_q \partial q &= -y_{ts} \partial s \end{aligned}$$

where

$$A_t = y_t(t,s) - c_t(t)[1-p(q)][2(r+q+j) + f(q)] > 0$$

$$B_t = y_{tt}(t) - [2(r+q+j) + f(q)]c_t(t) < 0$$

$$A_q = -f_q(q)\{d + c(t)(1-p(q))\} + [2(r+q+j) + f(q)]\left[c(t)p_q(q) + \frac{d}{h(q)^2}h_q(q)\right]$$

$$B_q = -c_t(t)f_q(q) < 0$$

and y_{ts} is the derivative of the marginal productivity of training y_t with respect to σ. Notice that the sign of A_q is uncertain because p_q is positive. Using Cramer's rule, we get the key comparative statics

[22] $$\frac{\partial t}{\partial s} = \frac{y_{ts}\, A_q - y_s\, B_q}{A_t\, B_q - B_t\, A_q}$$

[23] $$\frac{\partial q}{\partial s} = \frac{y_s\, B_t - y_{ts}\, A_t}{A_t\, B_q - B_t\, A_q}$$

The determinant $\Delta = A_t B_q - B_t A_q$ can be unambiguously signed by imposing that the steady state equilibrium is locally stable. A first order Taylor approximation of [19] and [20] yields

$$B(t,q) \approx B_t\,(t - t^*) + B_q\,(q - q^*)$$

$$A(t,q) \approx A_t\,(t - t^*) + A_q\,(q - q^*)$$

The Jacobian of this system, J, evaluated at the equilibrium point t^*, q^* is

$$\begin{bmatrix} B_t & B_q \\ A_t & A_q \end{bmatrix}$$

The pair t^*, q^* is a stable fixed point if $tr\ J < 0 < \det \mathrm{J}$. Therefore, stability requires that $B_t A_q - B_q A_t > 0$, and implies that the denominator in the comparative statics [22] and [23] is negative.

An increase in local economic density σ leads to an increase in q, since both the determinant $\Delta = A_t B_q - B_t A_q$ and the numerator of [23] are negative: the positive externalities associated to a higher value of s make job creation more profitable and lead to a higher equilibrium ratio of jobs to workers. The effect of economic density on q is stronger when the complementarity between skills and knowledge spillovers – measured by y_{ts} - is higher. Conversely, the effect of s on training cannot be signed, because of the presence of conflicting effects: on the one hand, an increase in s generates higher productivity, which encourages firms to invest more in training; on the other hand, a higher value of s leads to a higher q, which reduces the incentive to train, because wages are higher and profits are lower. The positive effect of economic density on training depends on the relative strength of the complementarity between training and knowledge spillovers. When this is low, the negative influence of higher wages and higher turnover tends to prevail. When $y_{ts}=0$, there is no complementarity and the overall relationship between density and training is negative.

An additional link between local labor market tightness and training not explicitly modeled here is that turnover is higher in tighter markets. Higher turnover discourages training. In Pissarides, 2000, a higher ? implies faster job destruction because the worker is more likely to find another job and quit. In our simple framework we can introduce this link in an ad-hoc way by assuming that the job separation rate q is an increasing function of tightness ?. Nothing of substance changes in the model, except that training is now discouraged both by higher wages and higher labor market turnover.

Economic density also reduces the unemployment rate u and increases the availability of skilled workers in the labor market, denoted by p. On the one hand, unemployment is a decreasing function of q, and q increases in s [9]. Therefore $\frac{\partial u}{\partial s}=\frac{\partial u}{\partial q}\frac{\partial q}{\partial s}<0$. On the other hand, as shown by [17], the fraction of skilled workers in the labor market increases with q, which in turn increases with s. Hence, $\frac{\partial p}{\partial s}=\frac{\partial p}{\partial q}\frac{\partial q}{\partial s}>0$.

The effect of density on wages is more complex, and corresponds to $\frac{\partial w}{\partial s}=\overset{+}{\frac{\partial w}{\partial s}}+\overset{+}{\frac{\partial w}{\partial q}}\overset{+}{\frac{\partial q}{\partial s}}+\overset{+}{\frac{\partial w}{\partial t}}\overset{?}{\frac{\partial t}{\partial s}}$. The first term is positive, because higher density increases productivity; the second term is also positive, and works via the effect of density on labor market tightness; the last term cannot be signed a priori, because it depends on the effect of density on training.

1.6 Endogenous Local Economic Density

In the previous sections we have defined σ as local economic density, with the view that local knowledge spillovers are more intense in denser local areas. According

---

[9] The effect of economic density on unemployment is less clear when we consider that the rate of separation may be negatively influenced by labor market tightness $q_q<0$, in this case: $\frac{\partial u}{\partial s}=\frac{\partial u}{\partial q}\frac{\partial q}{\partial s}+\frac{\partial u}{\partial q}\frac{\partial q}{\partial q}\frac{\partial q}{\partial s}$. While the first term is negative, the second is positive and the overall effect uncertain.

to Hall and Ciccone, 1996, denser local labor markets provide more opportunities of interaction between individuals and a faster diffusion of knowledge. They measure local density as the number of employees per squared kilometer, $s = \frac{N}{K}$, where *K* is the size of the local labor market. Since $N = (1-u)L$, local density is given by $s = \frac{(1-u)L}{K}$, an endogenous variable in our model because the unemployment rate depends on labor market tightness ?. We show in the Appendix that the comparative statics illustrated in the previous section continue to hold in a qualitative way in this case.

1.7 Endogenous Local Labor Force

The simple model illustrated above considers a single local labor market. Do our results change when mobility between markets is allowed? Consider the case of two markets. With free entry, firms are bound by the zero profit condition for vacancies and therefore are indifferent between geographical areas. On the other hand, workers can choose to move between local markets to maximize their expected utility. This implies that the labor force in each local area cannot be considered as fixed. Since in our model workers enter the labor market via unemployment, the following arbitrage condition has to be satisfied, where 1 and 2 indicate two different geographical areas and r and j are assumed to be identical in the two areas:

[24] $$U_1 = U_2$$

It turns out that this arbitrage condition is satisfied when:

[25] $$\frac{f_1(q_1)[y_1(t_1, s_1)]}{2(r+q+j)+f_1(q_1)} = \frac{f_2(q_2)[y_2(t_2, s_2)]}{2(r+q+j)+f_2(q_2)}$$

Let local density be defined as in Hall and Ciccone, 1996. Since the total population is given and equal to $L = L_1 + L_2$, it follows that

[26] $$\frac{L}{K} = \frac{K_1 s_1}{K(1-u_1)} + \frac{K_2 s_2}{K(1-u_2)}$$

These two additional equations determine the values of s_1 and s_2. The remaining endogenous variables, u_1, u_2, q_1 and q_2 are determined by equations similar to [16] and [19][10].

2. The Data

In the empirical application we use firm level data drawn from the 8th wave of the survey "Indagine sulle Imprese Manifatturiere", conducted by Mediocredito Centrale in 2001 with a questionnaire administered to a representative sample of Italian manufacturing firms with at least 10 employees[11]. The sample is stratified for firms with up to 500 employees and covers the universe of firms with higher employment. Strata are based on the geographical area, the industry and the firm size. The survey collects data on firm location, its age, the industry it belongs to, the total number of employees (with a breakdown between white and blue-collar employees), and the share of college and high school graduates[12], as well as accounting and financial data, such as the total value of sales, the amount spent in R&D investment, and access to public subsidies. These data are collected in 2001 for the reference period 1998 to 2000. Since measurement error and missing information increase with the length of time separating the reference year and the year of the interview, we focus our attention exclusively on the year before the survey.

The only available question on training in these data is formulated as follows: "How many employees - during year 2000 - have participated to training courses run by public or private organizations specialized in the provision of training?".[13] Since the questionnaire is filled up by management rather than by single employees, it is reasonable to expect that answers refer to training paid for or organized by the firm. By design the question focuses on formal off-the-job training and excludes both on-

[10] If the arbitrage condition is not satisfied, people migrates from the area characterized by a lower expected return from unemployment to the area with a higher return. For example if $U_1 > U_2$ people migrates from area 2 to area 1, which increases unemployment in the latter area and reduces it in the former. Given the number of vacancies, labor market tightness falls in area 1 and increases in area 2, and the arbitrage condition converges to equality.

[11] Mediocredito Centrale is an Italian investment bank. Detailed information about the data can be found in the Mediocredito Centrale's web site, www.mcc.it. Brunello and Gambarotto, 2004, use individual data from the European Community Household Panel to study the relationship between training and local economic density in the UK. We cannot use these data for Italy, however, because the available regional information for this country is only at the NUTS 1 level, too aggregate for our purposes. See Ciccone, 2001, for a discussion of the proper level of aggregation.

[12] The data contain also information on the number of workers with temporary and part-time contracts. Unfortunately, this information cannot be used in our regressions because of the large number of missing values.

[13] There is no information in our data on training hours and training costs.

the-job and informal training[14]. This could be a problem if firms in dense labor markets substitute off-the-job training with other forms of training, including on-the-job training, perhaps in an attempt to reduce turnover. In this case the uncovered effects of density on off- the-job training need not hold for training in general.

There are two reasons, we believe, which make this problem not so relevant. First, since our data include mainly firms with 10 to 50 employees, which cover 68% of the sample of reporting firms (and 81% of the total sample), the substitution between formal off-the-job training and formal on-the-job training is less likely to take place. Usually, these firms cannot spread over a large number of participants the fixed costs of training facilities and, as a consequence, rely more on external courses organized by specialized institutions including schools, training institutes, equipment suppliers and employer's organizations[15]. Second, the scope of substitution between different types of training is likely to be limited, because both general and specific skills are valuable to firms and the optimal mix depends on technological and organizational factors. Substituting general skills with specific skills is not an easy task, otherwise firms would provide only specific training, which is in contrast with empirical evidence suggesting that large part of the training financed by firms is general (Bishop, 1997, Loewenstein and Spletzer, 1999). Even if firms were able to substitute off-the-job training with on-the-job and informal training, the risk of turnover is not less severe if on-the-job training develops general skills. Training could be general even if courses are carried out on the firm premises and via informal activity based on the mentoring by experienced workers.

The sample we use in the estimates consists of 3517 manufacturing firms. Of these, only 1124 provide information on the number of employees participating in formal training courses. As shown in Table 1, the average percentage of trained employees in the sample of reporting firms is 12.2. While comparisons are obviously difficult, this percentage is not very different from that reported by employee surveys

[14] Measured training usually refers to formal training activities – such as training courses – and ignores informal training – such as learning by doing – which is very subjective and hard to measure. A recent review of the issues associated to measurement error in training is contained in Bassanini et al, 2005. On-the-job training differs from off-the-job training because the location of training is the workplace rather than the classroom. While it is difficult to individuate differences in terms of content, both off-the-job training and on-the-job training involve a mix of specific and general skills. As documented by Bishop, 1997, a growing number of firms are training their workers in completely general skills. Similarly, OECD, 2003, reports that courses occurring outside the workplace impart essentially general skills.

[15] Evidence on this is provided for Italy by Montanino (2000). According to the second Continued Vocational Training Survey, conducted by Eurostat in 1999, the percentage of firms with 10-19 and 20-49 employees which carried out internal CVT courses was 8 and 20 percent respectively.

such as the European Community Household Panel for Italy in the same year (9 percent), in spite of the fact that our survey only considers off-the-job training.

The vast majority of the firms reporting training in our sample is small: about 29 percent have between 10 and 20 employees, 39 percent have 21 to 50 employees and the rest is bigger. About 74 percent of these firms are localized in the industrialized Northern part of the country, and only 8 percent are located in the South. Table 1 shows the means and standard deviations of the firm-specific and area-specific variables used in the empirical analysis, separately for firms reporting and non reporting training information.

About 30 percent of the employees in firms reporting training are white collars and have a high school diploma or higher education, and 10 percent of the reporting firms belong to employer associations and networks (*consorzi*), which have been established for financial reasons and to promote exports and innovation. The percentage of reporting firms which belong to an industrial district and have received public subsidies, national and local, in the year 2000 is close to 41 and 48 percent respectively. Non reporting firms are typically smaller than reporting firms (average size equal to 34.59 versus 66.11 of reporting firms), are younger and have fewer employees with higher education, spend less in R&D and participate less both to networks and to industrial districts.

The large number of non – reporting firms, as well as the differences with reporting firms outlined in Table 1, clearly suggests that we should be careful when interpreting empirical results based only on the sub-sample of reporting firms. Since such sub-sample is unlikely to be a random draw from the population of firms, we model explicitly in the empirical analysis the endogenous selection of firms into each sub-sample.

We identify the local labor market with the province, which corresponds to the Nuts 3 Eurostat classification, in line with the existing empirical literature (see Ciccone, 2001).[16] We use the data from the Italian National Statistical Institute (The Labor Force Survey, 2000; The Industrial Census, 1996) to compute local economic density and different measures of agglomeration. Average total employment in the Nuts 3 areas in 2000 was 509 thousand employees, with a minimum of 57.5 thousand (Aosta – Valle d'Aosta) and a maximum of about 2 millions (Milan - Lombardy). Average industrial employment density for the sub-sample of reporting firms – measured as industrial employment per squared kilometer - was 74.5 employees per

[16] An alternative definition, the travel to work area (LLMA), is in our too small to take into account all relevant knowledge spillovers.

squared kilometer, ranging between 313.3 and 3.3[17]. These figures suggest that there are large differences across Italian local labor markets. While these differences are related to the well-known disparities between the North and the South of the country, there is also diversity among provinces belonging to the same region. For instance, industrial density in Lombardy, the largest Italian region, is on average equal to 164 employees per squared kilometer, with a range from 7 to 313.3.

3. The Empirical Specification

The model in Section 1 shows that the relationship between training and the density of local labor markets is complex and cannot be determined a priori. This result is derived by considering a stylized economy with a continuous - time infinite horizon and identical workers and firms, who invest in training with the same intensity. In the real world of heterogeneous workers and firms, one key dimension of training intensity is the percentage of workers receiving training, *T*, or training incidence. This is the variable available in our dataset. Since percentages are bounded between zero and one, we use the logistic transformation $I = \ln\frac{T}{1-T}$. We characterize the empirical relationship between training and density with the following three equations model

[27] $$I = bW + lD + e_1$$

[28] $$D = pZ + e_2$$

[29] $$P = 1(dY + e_3 > 0)$$

where $e_i \approx N(0, s_i^2)$, and we allow arbitrary correlation among the errors e_i[18].

The first equation relates training incidence *I* to log density *D* and the vector of exogenous variables *W*; the second equation is the linear projection of *D* on the vector *Z* of selected instruments; the third is a selection equation, where *P* is a dummy equal to 1 for firms reporting information on training – for which $dY + e_3 > 0$ - and to zero otherwise, and *Y* is a vector which includes all exogenous variables in *Z* and *W* plus at

[17] Average density, measured as total employment per squared kilometer, for the sub-sample of reporting firms was 238.7.
[18] The specification and the estimation strategy follow closely Wooldridge, 2002, Section 17.4.2.

least one additional variable, excluded from [27] and [28] (see Wooldridge, 2002)[19].

Considering the three equations in turn, the vector *W* in equation (27) includes both firm-specific and area-specific variables. Firm-specific effects control for differences in firm-specific productivity, and include a measure of average educational attainment (the percentage of employees with at least upper secondary education), the percentage of white collars, firm size and 18 industry dummies within the manufacturing sector. Unfortunately, we have no data on the age composition of the workforce, but we try to proxy this information with a dummy equal to 1 if the firm was established after 1990, with the idea that younger firms are more likely to have a younger workforce[20]. Since training and innovation are complements (see Acemoglu, 2002), we capture the degree of innovative activity with the share of R&D expenditure on total sales in 1999. We use data for 1999 rather than for the year 2000 to reduce the problems associated to the potential endogeneity of this variable.

Employer – provided training can be encouraged by public policy with the provision of incentives, tax breaks and subsidies. While we do not have information on the specific target of each incentive measure, we know whether each firm has received subsidies of any kind from the central and local government during the year 2000 and use this qualitative information as an additional dummy in the empirical specification.

The vast majority of firms in the sample is composed of small and medium firms, and these firms are often related in Italy by formal and informal ties, which could affect the incentives to train. On the one hand, small Italian firms may participate in production and financial networks (*consorzi)*, which provide technical and financial assistance and support, with potential spillovers on training activity. We capture the influence of these network with a dummy equal to 1 if the firm participates to a *consorzio,* and to 0 otherwise. On the other hand, many firms may be members of industrial districts. Industrial districts, defined as clusters of firms involved in the production of homogeneous goods, are an interesting feature of Italian medium and small firms (see Signorini, 2000). They are identified by the Italian National Statistical Institute (ISTAT) on the basis of two criteria: a) the area is a local labor market with a high degree of self-containment; b) the productive system in the area has a dominant specialization and is composed mainly of small and medium firms

[19] According to Wooldridge, as a practical matter the model should have at least two variables in Z and Y that are not in W. When there is only one excluded variable, technical identification of the parameters in [27] is still possible, but "..it is unlikely to work well in practice because of severe multicollinearity among the IVs.." (Wooldridge, 2002, p. 570).

[20] We do not have information on the composition of the workforce by gender. Since manufacturing sectors have different shares of female labor in total employment, the gender effect should be partly captured by industry dummies.

(see De Blasio and Di Addario, 2002). We classify a firm in our sample as belonging to an industrial district if it is located in an industrial district – as defined above - and if its line of production is coherent with the industrial specialization of the district.

Local areas differ in their degree of industrial specialization, which we measure for each firm with the ratio of employment in the own industry and area and employment in the area. If labor turnover takes place mainly within sectors (see Neal, 1995), then training can be higher when the industrial structure is less specialized because turnover is lower. On the other hand, the relevance of agglomeration externalities can vary with the degree of local specialization.

Area-specific variables include a measure of economic density, our key explanatory variable, and a set of regional dummies, which are expected to capture the confounding effects of omitted area – specific variables, including the local unemployment rate[21]. Our measure of density is the log of the ratio between industrial employment and the size of the province expressed in squared kilometers, but we also experiment with a broader measure of density, based on total employment. While industrial density is appropriate because the firms in our sample belong to the manufacturing sector, a broader definition captures inter-sectoral spillovers. Since our measures of density are at the level of the province and regions include several provinces, the inclusion of regional dummies implies that we focus on the correlation between training and density within each region.

The relevant empirical literature has pointed out that density in equation [27] is an endogenous variable. As argued by Glaeser and Mare, 2001, one of the main difficulties in estimating agglomeration effects from labor market data hinges on the fact that denser areas are expected to be populated by abler workers. If we consider the mobility of individuals among local areas, the higher density of some geographical areas may depend on the fact that these areas are more productive and attract talented individuals. To the extent that unobserved ability and productivity are not fully captured by our controls, OLS estimates are biased.

Suppose that firms located in dense areas employ workers of higher ability than those employed by firms in sparse areas. One the one hand, “better” workers are more likely to be trained; on the other hand, “better” workers might be subject to higher turnover rates, due for example to better and more frequent offers from competitors, and this would reduce the incentive to train in dense versus sparse areas. In either case one can find that density affects the intensity of training even

[21] A well known stylized fact of Italy is that the bulk of the variation of local unemployment rates is at regional and macro-regional level (North, Centre and South).

with no agglomeration forces at work. Unless the parameter of interest (?) is correctly identified, the estimated coefficient could simply capture differences in labor force composition, rather than the net effect of density on training outlined in the theoretical model.

Our solution to this serious identification problem is to estimate [27] using instrumental variables (IV). IV estimation is desirable also because density can be measured with error. To illustrate, our data provide information on the location of the firm, and we match to this location the relevant province and the associated measure of density. When the firm is multi-plant, however, we cannot exclude that some of these plants are located in a different province. In this case density is measured with error because the province does not necessarily encompass the location of all plants. The potential presence in our data of multi-plant firms, with plants located in different provinces[22], has induced us to exclude from the main analysis all firms with more than 500 employees, under the reasonable assumption that the probability of having more than one plant is significantly higher in larger firms. By so doing, we hope to reduce measurement error.

Following Ciccone and Hall, 1996, Combes et al., 2004, and Mion and Naticchioni, 2005, we instrument density with deeply lagged values of population density, going back to 1871 and 1911. These lags are eligible as instruments if the main sources of agglomeration in the 19^{th} and early 20^{th} century are not related to the residuals in equation [27]. As argued by Ciccone and Hall, 1996, this is equivalent to assuming that the early patterns of agglomeration in Italy do not reflect factors which significantly contribute to unobserved productivity in the year 2000, but "..have a remaining influence mainly through the legacy of agglomeration" (Ciccone and Hall, p. 61).

While one might argue that current agglomeration patterns reflect the long run effects of the economic miracle which took place in Italy after the Second World War – which triggered substantial migration from the agricultural South to the industrialized North of the country - it is difficult to believe that these effects have worked their way backwards so as to influence the distribution of the population at the beginning of the 20^{th} century or even earlier, in the 19^{th} century. In their study of the US, Ciccone and Hall, 1996, use late 19^{th} century data to instrument current density. We use data on the resident population from the Italian Censuses of 1871 and 1911[23]. With two

[22] We have no information on whether a firm in our sample has more than one plant and on the location of these plants.
[23] Since the Italian government has instituted 10 new provinces in the past 10 years, with an almost even distribution between the Northern, Central and Southern areas of the country, we use data at the level of the municipality in 1871

instruments for a single endogenous variable – density - we can verify instrument validity using Hansen J test.

The selection equation [29] is motivated by the fact that, in our data, slightly less than one firm out of three provides information on *I*. Since we have little reason to expect that reporting training information is random, we use Heckman's correction to take into account endogenous selectivity. Following Wooldridge, 2002, we start by estimating a Probit equation for the probability that firms report information on training, and compute the inverse Mills ratio. Next, we add this ratio to the list of exogenous regressors in equation [27] and estimate the augmented specification by instrumental variables, so as to take into account the endogeneity of density. Equation [29] includes all the exogenous variables in equations [27] and [28] plus the interaction between regional dummies and a dummy equal to 1 if the firm has less than 30 employees and to 0 otherwise, which combines two facts: non reporting firms are smaller than reporting firms, and the distribution of reporting firms varies across regions, with higher reporting rates in the North and lower rates in the South[24].

4. The Results

Our main findings are illustrated in Table 2, where we report in the first two columns the IV estimates of the training equation – with and without correction for endogenous selection, and in the third column a less parsimonious specification, which includes the index of own industry specialization – measured for each firm by the ratio of employment in the own industry and area and employment in the area.

The last row in the table reports the F test of the hypothesis that the selected instruments – population density in 1871 and 1911 - are jointly different from zero in the first stage regression of log density – our endogenous variable – on all the exogenous variables in the model, including the selected instruments. The value of the test is above 100, much higher than the rule of thumb value of 10, suggested by Staiger and Stock, 1997, to reject the hypothesis that the instruments are weak. We also report in each column the p-value of the Hansen J test on instrument validity,

and 1911 to assign the population to the current list of provinces.

[24] Notice that in equation (27) we include regional dummies and control for firm size with the number of employees in each firm. Therefore our implicit and rather mild assumption is that the effect of regional dummies varies with firm size in the selection equation but not in the training equation.

which suggests that such validity is not rejected by our data[25].

We find that training incidence is higher when the percentage of white-collar workers and the share of employees with at least upper secondary education are higher. There is evidence that firms which spend more, as a percentage of their sales, in research and development, are more likely to engage in training, which points to the complementarity between innovative activity and the development of an adequate human capital. Training is higher among firms which have received economic incentives from the central or the local government, because part of these subsidies – such as those provided under the European Social Fund - is targeted to encourage training programs.

Once we control for endogenous selection into the sample of reporting firms, there is no statistically significant relationship between training and firm size. Training, however, is more frequent among firms established after 1990, which are likely to have a younger labor force. Perhaps more interestingly, we also find that training is higher among firms which belong both to formal networks and associations of firms, which specialize in financial assistance, export promotion, research and development and else, and to industrial districts. This might depend on the fact that all these associations – both formal and informal - make member firms more productive and able to capture profit opportunities and knowledge spillovers (see Hirschman, 1977; Podonly and Baron, 1997)[26]. The positive - and statistically significant at the 10 percent level of confidence - coefficient associated to industrial districts can be explained with the distinct cooperative behavior which characterizes the firms belonging to these organizations, resulting in the reduction of the risk of poaching and the increase in the returns to training (see Brusco, 1982, and Becattini et al. 1990).

The coefficient of the inverse Mills ratio attracts a positive and statistically significant sign, which points to the positive correlation between the error terms e_1 and e_3: firms which are more likely to report their training investment have also a higher training incidence. As shown in Table 3, which presents the results for the probit estimate of [29], reporting is more likely for larger firms, which invest more in R&D, receive public incentives, belong to networks and have a higher share of educated labor. We test whether the interactions between regional dummies and a

[25] Standard errors are computed by bootstrapping (200 replications). In each replication we re-estimate both the probit model and the structural equation.
[26] Several sociological studies point out that networks, by promoting social interaction, favour the transfer of tacit knowledge.

dummy equal to 1 when firms have less than 30 employees are jointly statistically significant, and reject the null hypothesis of no significance[27].

The key result of Table 2 is that the estimated coefficient of log density is always negative, which suggests a negative relationship between training and employment density. The inclusion of the inverse Mills ratio in the IV regression increases the absolute value of the impact of log density on training, which also becomes statistically significant at the 5 percent level of confidence. The further addition in Column 3 of the index of specialization does not alter the coefficients reported in Column 2, as the additional variable attracts an imprecise coefficient.

While the relationship between density and training is statistically significant, the size of the effect is rather small: focusing on the second Column of Table 2 and evaluating at the sample average percentage of trained employees in all firms with at most 500 employees, we find that a 10 percent increase in local density reduces the percentage of trained employees by 1.46 percent[28]. Table 4 is organized as Table 2 and reports the results of the estimates when we use total rather than industrial employment to compute local density. The impact of density on training remains negative and statistically significant, but the absolute value of the estimated coefficient slightly declines.

The uncovered negative correlation between training and local economic density is not new in the literature. Brunello and Gambarotto, 2004, for instance, use individual rather than firm data for the UK and find qualitatively similar results. The theoretical model presented in this paper provides a framework for the interpretation of our and their findings. We argue that higher density influences training incidence in two ways. First, the more intense knowledge spillovers increase productivity and the marginal benefits of training, which are higher for any value of labor market tightness θ. Second, larger positive externalities increase job creation and vacancies, which raise in turn both wages and turnover. While the former effect increases training, the latter effect reduces it. Our results suggest that, in the Italian economic environment, it is the latter effect which prevails.

Additional support to this conclusion comes from Table 5, where we provide some direct evidence on the relationship between voluntary turnover and local employment density, using the 2001 wave of the European Community Household Panel. We define a dummy variable equal to 1 in the event of voluntary turnover

[27] The p-value of the test is equal to .000
[28] The elasticity is obtained by multiplying the estimated coefficient of log density reported in the table by (1-T), where T is evaluated at its sample mean.

during the reference period – defined as turnover motivated by better economic conditions - and to 0 otherwise, and estimate a probit model which includes among the regressors educational attainment, gender, age, age squared and log employment density – measured at the NUTS 1 level of regional aggregation. While this definition of density is open to question because of the broadness of regional areas – which explains why we did not use these data for our main analysis - it does provide a useful benchmark to study the relationship between density and turnover. As shown in the first column of Table 5, voluntary turnover is higher in denser areas, which confirms the relative importance of poaching effects as deterrents to training.

In the same table we also show that the probability of filling a temporary job is negatively and significantly related to density. It is well known that the probability of receiving training differs among temporary and permanent workers. Since we cannot control for the type of jobs in the estimates reported in Tables 2 and 4 – because the available control has too many missing values – one might argue that the uncovered negative relationship between density and training is spuriously driven by the fact that firms in dense provinces use temporary workers to a higher extent. Our evidence in the second column of Table 5 points to the contrary, and shows that the probability of being a temporary employee is lower, not higher, in dense economic areas.

5. Extensions

So far we have assumed in our empirical specifications that the relationship between local density and training is invariant with firm size. Since both agglomeration and labor turnover effects could vary with the size of the firm, we relax this assumption in the first Column of Table 6, where we interact local density with two dummies, one for firms with less than 30 employees, and another for firms with at least 30 employees. It turns out that the coefficient associated to density is always negative, but statistically significant only for smaller firms. The higher sensitivity of training to density among smaller firms suggests that poaching effects may be stronger for this group of firms.

Since the cooperative climate which characterizes industrial districts might affect the relationship between density and turnover, we interact density with a dummy equal to 1 if the firm belongs to an industrial district, and report the results in the second Column of Table 6. It turns out that the negative and statistically significant impact of density on training is confirmed only for the firms which do not belong to an industrial district. For the rest of the firms, we cannot reject the hypothesis that the impact of density of training is equal to zero[29]. The natural interpretation of this result is that the formal and informal institutions which regulate the interactions of firms belonging to an industrial district reduce the risk of poaching and the negative congestion effects of agglomeration[30].

6. Conclusions

Economic density, according to a growing body of research, encourages shared learning among individuals and firms and generates positive knowledge spillovers. Since the ability to translate information in real economic advantages is strictly related to the skills of the labor force, firms located in dense labor markets are encouraged to invest in training. However, density not only provides benefits, it also generates economic costs - identifiable with higher wages and a higher turnover risk. In the model proposed in this paper, these effects are analyzed in a matching and search framework.

While from the theoretical point of view the effect of economic density on training investment can be either negative or positive, our empirical analysis suggests the prevalence of a negative effect. Using data on a sample of Italian manufacturing firms, we show that training is higher in provinces with lower employment density. This result confirms the evidence presented by Brunello and Gambarotto, 2004, for the UK, and more recently by Muehleman and Walter, 2006, for Switzerland.

Our findings refer to off-the-job training provided by external organizations specialized in training provision. An open question is whether they can be extended to all training. While our data do not allow a direct answer, we speculate that the exclusion of on-the-job training should not substantially modify our results for two reasons: first, our sample includes mainly small firms with less than 50 employees, which find it too costly to provide their own training facilities; second, there is no clear correlation between the type of knowledge transmission (on-the-job versus off-the-job, formal versus informal) and the transferability of knowledge. While firms might substitute general with specific training, the empirical evidence shows that to a

---

[29] In the training regression, we instrument the interaction of density with the industrial district dummy with the interactions of population density in 1897 and 1911 with the industrial district dummy, and test whether the sum of the coefficients attracted by density and the interaction term is equal to zero. Since the p-value of the test is 0.688, we cannot reject the hypothesis of a null effect of density on training for the firms belonging to an industrial district.

[30] In a different but related context, Soskice, 1994, argues that the institutions facilitating the cooperation among firms and reducing poaching are key to understanding the German “high skill equilibrium”

large extent on-the-job training is general, so that the threat of turnover applies as well.

Our results suggest that the productivity gains associated to local economic density are not attained because firms in denser areas train more their employees. If anything, they train less, we believe because congestion effects – such as higher wages, poaching and turnover - are at work which dampen the positive effects of clustering on training. For the viewpoint of policy, the key question is whether firms in dense economic areas are training too little. It is hard to say. The economic literature on the under-provision of training is rather inconclusive, mainly because of relevant measurement issues: to establish a case for under-provision requires information on the private and social costs and returns to training, a formidable empirical task with the data at hand[31].

References

Acemoglu, D., [2002], Technical Change, Inequality and the Labor Market, *Journal of Economic Literature*, 40, 1, pp. 7-72.

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

Albrecht, J. and Vroman, S., [2002], A matching model with endogenous skill requirements, *International Economic Review,* 43, 283-305.

Anselin, L., Varga, A., and Acs, Z., [1997], Local Geographic Spillovers between University Research and High Technology Innovations, *Journal of Urban Economics,* 42, 422-48.

Bassanini, A., Brunello, G., Booth, A., De Paola, M. and Leuven, E., [2005], Workplace Training in Europe, forthcoming in Brunello, G., Garibaldi, P. and Wasmer, E., (eds), 2006, *Education and Training in Europe*, Oxford university Press,

Blanchard O. and Diamond P., [1989], The Beveridge Curve, *Brooking Papers on Economic Activity*, vol.1, 1-60.

Becattini G., Sengenberger W., Pyke F., [1990], (eds.), Industrial Districts and Inter-Firm Co-operation in Italy, *International Labour Office*, Geneva.

Bishop, J.H., [1997], What We Know about Employer-Provided Training: A Review of the Literature, *Research in Labor Economics*, 16, 19-87.

Brunello, G. and Medio, A., [2001], An explanation of international differences in education and workplace training, *European Economic Review*, 307-322.

Brunello, G. and Gambarotto, F., [2004], Agglomeration Effects on Employer-Provided Training: Evidence from the UK, forthcoming in *Regional Science and Urban Economics*

Brusco, S., [1982], The Emilian Model: Productive Decentralization and Social Integration, *Cambridge Journal of Economics*, 6(1), 167-184.

Ciccone, A., [2001], Agglomeration-Effects in Europe and in the USA, *CREI working paper,* n. 9.

Ciccone, A., and Hall, R. E., [1996], Productivity and the Density of Economic Activity, *American Economic Review,* 86, 54-70.

[31] See Bassanini et al. 2005.

Ciccone, A. and Cingano F., [2003], Skills and clusters, in R. Söltwedel, J. Brocker and D. Dohse (ed.) *Innovation Clusters and Interregional Competition*, Advances in spatial Science, Springer

Cingano, F. and Schivardi, F., [2004], Identifying the sources of local productivity growth, *Journal of the European Economic Association*, 2(4), 720-742.

De Blasio G. and Di Addario S., [2002], Labor market pooling: evidence from Italian industrial districts, *Banca d'Italia, Temi di discussione*, n. 453.

Feldman, M. P. ,[1993], An examination of the geography of innovation, *Industrial and Corporate Change,* 4, pp. 451-470.

Glaeser, E., Kallal, H., Scheinkman, J. and Shleifer, A., [1992], Growth in Cities, *Journal of Political Economy*, 100, pp.1126-52,

Glaeser, E. and Mare D., [2001], Cities and Skills, *Journal of Labor Economics*, 19(2), 316-342.

Hirschman, A.O., [1977], A generalized linkage approach to development, with special reference to staples, *Economic Development and Cultural Change* (Suppl.) 25, pp. 67–98.

Henderson, V., [1986], Efficiency of Resource Usage and LLS Size, *Journal of Urban Economics* 18, pp. 23-56.

Henderson V., Kuncoro A. and Turner M., [1995], Industrial Development of Cities, *Journal of Political Economy*, 103, pp.1067-1090.

Loewenstein M., and Spletzer, J., [1999], General and specific training: Evidence and Implication, *The Journal of Human Resources*, 34, 4, pp. 710-733.

Jacobs, J., [1969], *The economy of Cities,* Vintage, New York.

Jaffe A.B., Trajtenberg M. and Henderson R., [1993], Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics*, 108(3), pp. 577-598.

Mion, G. and Naticchioni P., [2005], Urbanization Externalities, Market Potential and Spatial Sorting of Skills and Firms, mimeo.

Montanino, A., [2001], Formazione aziendale, struttura dell'occupazione e dimensione dell'impresa, *Rivista di Politica Economica*, pp.13-36.

Moretti E., [2004], Workers' Education, Spillovers and Productivity: Evidence from Plant-Level Production Functions, *American Economic Review***,** 94(3),

Mortensen D. and Pissarides C., [1999], New Developments in Models of Search in the Labour Markets, in Ashenfelter, O. and Card, D. (eds), *Handbook of Labor Economics*.

Muehleman, S., and Walter, S., [2006], Regional effects on employer – provided training: evidence from Switzerland, *Cesifo Working Paper*s, n.1665.

Neal, D., [1995], Industry-Specific Human Capital: Evidence from Displaced Workers , *Journal of Labor Economics*, 13(4), pp. 653-77.

OECD, [2003], Upgrading Workers' Skills and Competencies, chapter 5, *Employment Outlook,* Paris.

Pissarides C., [2000], *Equilibrium Unemployment Theory*, second edition, Cambridge, MA: MIT Press.

Podolny J. M. and Baron J. N., [1997], Resources and Relationships: Social Networks and Mobility in the Workplace, *American Sociological Review*, 62, pp. 673-93.

Rosenthal, S. and Strange W., [2004], Evidence on the Nature and Sources of Agglomeration Economies, in J. V. Henderson and J.F. Thisse, eds., *Handbook of Urban and Regional Economics*.

Signorini, L. F., [2000], *Lo sviluppo locale. Un'indagine della Banca d'Italia sui distretti industriali*, Roma, Donzelli Editore.

Soskice, D., [1994], Reconciling markets and institutions: the German apprenticeship system, in Lynch, L., 1994, *Training and the Private Sector*, Chicago University Press.

Staiger, D. and Stock J., [1997], Instrumental Variable Regression With Weak Instruments, *Econometrica*, , 65, 5, pp.557-586.

Stevens, M., [1996], Transferable training and poaching externalities, in Booth and Snower (eds), *Acquiring Skills*, Cambridge University Press, Cambridge, pp. 19-40.

Wooldridge J., [2002], *Econometric Analysis of Cross Section and Panel Data,* MIT Press.

Table 1. Descriptive statistics

	Firms reporting training	*Firms not reporting training*
% of trained employees	12.2	-
% of firms participating in networks	10.0	7.7
% firms established after 1990	21.1	21.1
Share of R&D expenditure over total sales in 1999	0.9	0.5
% of employees with upper secondary education or higher	40.3	35.4
% of white collars	28.7	24.2
% of firms which received public subsidies	47.9	34.4
Number of employees	66.11 [87.45]	34.59 [45.19]
% of firms belonging to an industrial district	41.3	36.5
Local industrial density (industrial employment/KM)	74.52 [86.75]	78.05 [91.92]
Local density (total employment / KM)	220.42 [285.08])	238.73 [304.21]
Number of firms	1124	2393

Note: standard errors within brackets.

Table 2. IV estimates. All firms. With and without correction for selectivity. Log density measured as the log of the ratio between industrial employment and squared kilometers of the area. Dependent variable: ln(T/(1-T)).

	(1)	*(2)*	*(3)*
Log industrial density	-0.075	-0.162	-0.167
	[.050]	[.078]**	[.078]**
Networks	0.183	0.348	0.352
	[.112]*	[.137]***	[.138]***
% sec. educ. or college	0.504	0.631	0.634
	[.128]***	[.165]***	[.164]***
Percentage white collars	0.450	0.887	0.886
	[.185]***	[.278]**	[.274]***
Expenditure in RD	3.855	4.942	4.944
	[1.489]**	[1.969]**	[1.977]**
Incentives	0.045	0.237	0.239
	[.072]	[.076]***	[.076]***
Firm established after 1990	0.115	0.158	0.157
	[.067]*	[.089]*	[.090]*
Number of employees *100	-0.309	-0.063	-0.062
	[.031]***	[.070]	[.070]
Industrial district dummy	0.111	0.153	0.154
	[.054]*	[.080]*	[.080]*
Inverse Mills ratio		1.093	1.095
		[.234]***	[.232]***
Specialization index			.162
			[.272]
Hansen J test – p value	0.665	0.639	0.636
Observations	1124	1124	1124
R-squared	0.14	0.15	0.15
F test first stage	146.94		
	(.000)		

Note: bootstrapped standard errors in brackets, p-values within parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. Each regression includes a constant, industry and regional dummies. Instruments of log density: log population density in 1871 and 1911.

Table 3. Probit Equation. Dependent variable: dummy equal to 1 if the firm reports training and to 0 otherwise. Firms with 10 to 500 employees.

Networks	0.209
	[.082]***
% upper secondary or higher ed.	0.205
	[.112]*
% white collars	0.524
	[.166]***
Expenditure in RD	1.693
	[1.202]
Incentives	0.215
	[.049]***
Number of employees	0.266
	[.044]***
Industrial district dummy	0.058
	[.054]
Log population density in 1871	0.196
	[.175]
Log population density in 1911	-0.229
	[.158]
Observations	3517
Pseudo R Squared	0.103

Note: robust standard errors within brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. The regression includes a constant, industry, regional dummies, and interactions of regional dummies with a dummy equal to 1 for firms with less than 30 employees and 0 otherwise.

Table 4. IV estimates. All firms. With and without correction for selectivity. Log density measured as the log of the ratio between total employment and squared kilometers of the area. Dependent variable: ln(T/(1-T)).

	(1)	(2)	(3)
Log total density	-0.069 [.046]	-0.148 [.070]**	-0.152 [.071]**
Networks	0.184 [.111]*	0.349 [.137]***	0.350 [.137]***
% sec. educ. or college	0.504 [.128]***	0.629 [.165]***	0.630 [.164]***
Percentage white collars	0.447 [.185]***	0.876 [.275]**	0.874 [.275]***
Expenditure in RD	3.884 [1.487]**	4.993 [1.960]**	4.995 [1.967]**
Incentives	0.043 [.072]	0.232 [.075]***	0.234 [.075]***
Firm established after 1990	0.115 [.067]*	0.158 [.089]*	0.157 [.089]*
Number of employees *100	-0.309 [.031)]**	-0.064 [.071]	-0.064 [.070]
Industrial district dummy	0.107 [.055]**	0.137 [.082]*	0.137 [.082]*
Inverse Mills ratio		1.082 [.234]***	1.083 [.232]***
Specialization index			.196 [.272]
Hansen J test – p value	0.639	0.712	0.600
Observations	1124	1124	1124
R-squared	0.14	0.15	0.15
F test first stage	164.68 (.000)		

Note: bootstrapped standard errors in brackets, p-values within parentheses; * significant at 10%; ** significant at 5%; *** significant at 1%. Each regression includes a constant, industry and regional dummies. Instruments of density: log population density in 1871 and 1911.

Table 5. Probit estimates of voluntary turnover and temporary job contracts. The European Community Household Panel, 2001.

	(1)	**(2)**
	Voluntary turnover	Temporary jobs
High school	0.075 [0.41]*	-0.105 [0.08]
College	0.009 [0.03]	0.270 [0.15]*
Gender	-0.356 [0.06]***	0.260 [0.03]***
Age	0.124 [0.03]***	-0.124 [0.02]***
Age squared	-0.001 [0.00]***	0.001 [0.00]***
Log density	0.211 [0.06]***	-0.204 [0.07]***
Observations	4350	4350

Note: robust standard errors within brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. The regression includes a constant.

Table 6. IV estimates. All firms. With interactions of log density with firm size. Dependent variable: ln(T/(1-T)).

	(1)	*(2)*
Log density * (dummy =1 if employment <30)	-0.170 [.076]***	
Log density * (dummy =0 if employment >=30)	-0.133 [-.092]	
Log density		-.204 [.084]**
Log density * (dummy =1 if firms belongs to industrial district)		.240 [.129]*
Networks	0.321 [.136]***	0.338 [.137]***
% sec. educ. or college	0.583 [.169]***	0.631 [.163]***
Percentage white collars	0.797 [.293]***	0.846 [.274]***
Expenditure in RD	4.605 [1.983]**	4.854 [1.955]**
Incentives	0.205 [.086]***	0.222 [.075]***
Firm established after 1990	0.139 [.089]	0.155 [.088]*
Number of employees *100	-0.093 [.076]	-0.074 [.069]
Industrial district dummy	0.144 [.078]*	0.879 [.417]*
Inverse Mills ratio	0.826 [.412)***	1.037 [.229]***
Hansen J test – p value	0.371	0.963
Observations	1124	1124
R-squared	0.15	0.15

Note: bootstrapped standard errors in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%. Each regression includes a constant, industry and regional dummies. Instruments of density: log population density in 1871 and 1911.

Appendix to Section 1.6

Define for simplicity $\frac{L}{K}=g$ and $s=g(1-u)$. Total differentiation of equations [19] and [20] yields

$$A_t \partial t + A_q \partial q = -y_s (1-u)\partial g$$

$$B_t \partial t + B_q \partial q = -y_{st} (1-u)\partial g$$

where:

$$A_t = y_t (t,s) - c_t (t)[1-p(q)][2(r+q+j)+f(q)] > 0$$

$$B_t = y_{tt}(t) - [2(r+q+j)+f(q)]c_t (t) < 0$$

$$A_q = -f_q (q)\{d + c(1-p(q)\} + [2(r+q+j)+f(q)]\left[c(t)p_q + \frac{d}{h^2}h_q (q)\right] + \frac{y_s g f_q (q)(q+j)}{[q+j+f(q)]^2}$$

$$B_q = -c_t (t) f_q (q) + \frac{y_s g f_q (q)(q+j)}{[q+j+f(q)]^2}$$

Using Cramer's rule:

$$\frac{\partial t}{\partial g (1-u)} = \frac{y_{st} A_q - y_s B_q}{A_t B_q - B_t A_q}$$

$$\frac{\partial q}{\partial g (1-u)} = \frac{y_s B_t - y_{st} A_t}{A_t B_q - B_t A_q}$$

As shown in section 1.5, if we impose that the equilibrium is locally stable, then $A_t B_q - B_t A_q < 0$. This implies that the denominator in the comparative statics is negative. With this in mind, it is easy to check that the sign of the derivatives does not change with respect to the baseline case.

Next notice that we are interested in how τ and θ vary when density *g(1-u)* varies. Therefore we are interested in the sign of $\frac{\partial t}{\partial g (1-u)}$ and $\frac{\partial q}{\partial g (1-u)}$. In the former case this derivative is equal to

$$\frac{\partial t}{\partial g}\left[\frac{\partial g(1-u)}{\partial g}\right]^{-1}=\frac{\partial t}{\partial g}\left[(1-u)-gu\frac{\partial u}{\partial q}\frac{\partial q}{\partial g}\right]$$

Since the term within brackets is positive, the sign of $\frac{\partial t}{\partial g}$ and $\frac{\partial t}{\partial g(1-u)}$ is the same.

The same argument can be applied to $\frac{\partial q}{\partial g(1-u)}$.