

IZA DP No. 1173

**Training and the Density of Economic Activity:
Evidence from Italy**

Giorgio Brunello
Maria de Paola

June 2004

Training and the Density of Economic Activity: Evidence from Italy

Giorgio Brunello

*University of Padova, CESifo
and IZA Bonn*

Maria de Paola

University of Calabria

Discussion Paper No. 1173
June 2004

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available on the IZA website (www.iza.org) or directly from the author.

ABSTRACT

Training and the Density of Economic Activity: Evidence from Italy*

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 with local density; on the other hand, the negative influence of higher wages in denser areas reduces training. Overall the relationship can be either positive or negative, depending on the relative strength of the 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. This empirical result confirms previous evidence by Brunello and Gamberotto (2004) based on UK data.

JEL Classification: J24, R12

Keywords: training, local labor markets, Italy

Corresponding author:

Giorgio Brunello
Department of Economics
University of Padova
Via del Santo 33
35100 Padova
Italy
Email: Giorgio.brunello@unipd.it

* We are grateful to Francesca Gamberotto for several discussions on the topics covered by this paper.

Introduction

Does training vary with local economic conditions? One possible answer is that firms operating in imperfectly competitive labor markets with a relatively high unemployment rate have higher monopsonistic power, which generates wage compression (see Acemoglu and Pischke, 1999) and encourages the provision of training. Based on this view, training should be higher in a depressed local labor market, also because the poaching threat is low. An alternative answer relies on the fact that high local unemployment, by increasing the availability of skilled employees in the local labor market, reduces the costs of locating an adequate match and consequently represents an opportunity to save training costs, which reduces the incentives to train (see Brunello and Medio 2001, Barron, Black, and Loewenstein 1989, Stevens 1994).

Yet another answer is based on the observation that active and competitive local labor markets generate positive spillovers for firms. One of the advantages identified by Marshall for a firm locating near other producers in the same industry is the effect that geographic proximity produces on the spread of information and on the ability of firms to exchange ideas, discuss solutions to problems, and be cognizant of other important information (Feldman, 1993). In Marshall 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 economic system. The geographic range of knowledge spillovers is strictly related to the distinction between written and unwritten knowledge. While the first can be formalized and easily transferred across distant geographic areas, the second is difficult to codify and is best transmitted via face-to-face interactions and frequent and repeated contacts taking place only inside a limited geographical area (von Hippel, 1994, Jacobs 1969, Storper and Venables, 2004).

The ability to adapt new technology and ideas developed by other firms is strictly related to the skills of the local labor force. It follows that training can have a relevant role to play in helping firms benefit from positive externalities deriving from dense labor markets. Knowledge acquisition and the ability to apply tacit knowledge in order to improve production are enhanced when workers possess the skills and capabilities, which help them understand and adapt the information they receive.

As recently argued by Brunello and Gambarotto, 2004, however, firms located in dense labor markets need to balance in their training decisions these positive external effects with the negative effects deriving from the higher risk of poaching. This implies that the relationship between local agglomeration effects, measured by local employment density, and employer-provided training, can take either sign. Their empirical investigation on UK data shows that the balance of positive and negative effects is tilted in favor of the latter, because training incidence is lower, *ceteris paribus*, in denser economic areas.

In this paper we follow the same thread of thought and develop a model of employer-provided training in a search equilibrium setup *a la* Pissarides, 2000, to show that the complementarity between local knowledge spillovers and training is not sufficient to deliver a positive relationship between local economic density and training incidence. The model provides guidance for the empirical investigation, which is the core of the paper. We use data drawn from the 8th survey on Italian manufacturing firms (*Indagine sulle Imprese Manifatturiere*) conducted by Mediocredito Centrale to estimate the relationship between training incidence, measured as the annual percentage of trained employees in each sample firm, and local labor market density, where the local labor market is identified with the province. Our empirical findings confirm the key results by Brunello and Gambarotto, 2004, in that the estimated relationship is negative and statistically significant. An important difference between this paper and theirs, however, is that while they find that training incidence is higher in areas with higher industrial variety, we have evidence that training increases with industrial specialization. We speculate that this difference may be due to differences in the industrial structure of Italy and the UK.

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 estimated model is specified. The main results are discussed in section 4, while section 5 presents some robustness exercises. Conclusions follow.

1. The Model

We consider an economy populated by a large and constant number of identical risk neutral firms and workers who are infinitely lived and discount the future at the common rate r . Workers and firms match in a market characterized by frictions (Pissarides, 2000; Mortensen and Pissarides, 1999; Blanchard and Diamond, 1989). After the initial match has taken place, each firm decides once and for all whether and how much to train the employee. An assumption that greatly simplifies the model below is that training occurs only in the initial period. In the case of separation at the end of each period the firm and the worker can re-match but no other training occurs.

Training is general and increases the worker's productivity both in the incumbent firm and in other firms. Since 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. Training increases individual productivity for two reasons. First, the employee increases her skill 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 in higher productivity on the job (Jaffe, Trajtenberg and Henderson, 1993; Anselin, Varga and Acs, 1997). 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). This view suggests that the positive externalities of agglomeration are better exploited by skilled labor. Following Ciccone and Hall, 1996, we posit that the intensity of these externalities increases with the density of economic activity in a local labor market.

Training costs are sunk during the initial period. With a perfect capital market and conditional on training, the asset value of employment in each period F_E satisfies the following Bellman equation

$$[1] \quad rF_E = y(\tau) + \psi\tau\sigma - w(\tau) + q(F_V - F_E)$$

where y is productivity, τ is training, w are wages and F_V is the asset value of a vacant job slot, σ is a measure of the intensity of knowledge spillovers in the local labor market, ψ is a parameter and q is the rate of exogenous job separation.

The filled job yields a net income flow equal to productivity minus the wage paid to the worker. Training increases productivity both directly and indirectly because of the complementarity of skills and knowledge spillovers, which we capture with the term $\psi\tau\sigma$. Naturally, the wage also increases with training. With exogenous probability q the pair separates and a job vacancy is created. The associated change of state yields a return equal to $F_V - F_E$. Since firms are identical and the wage rate does not vary across firms, there is no endogenous turnover.

The asset value of a vacant job satisfies the condition

$$[2] \quad rF_V = -d + h(F_E - F_V)$$

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 will find a new match. The change of state yields a return equal to $F_E - F_V$.

Assuming that all profit opportunities for new jobs are exploited, the free-entry condition is $F_V = 0$. Substituting this condition in (2) yields

$$[3] \quad F_V = \frac{hF_E - d}{r + h}$$

or by substitution

$$[4] \quad F_E = \frac{d}{h}$$

Therefore, the expected profits from a job are equal to the expected cost of finding a new match (see Pissarides, 2000).

1.1 Wage Determination

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

$$[5] \quad rE = w(\tau) + q(U - E)$$

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 a separation $(U - E)$, which occurs at rate q . There is no job-to-job mobility and each separation is an inflow in the unemployment pool. Letting unemployment benefits be equal to zero, the asset value of unemployment is

$$[6] \quad rU = f(E - U)$$

where f is the hazard rate from unemployment. The (trained) worker gets nothing with probability $(1-f)$ and the surplus $E-U$ with probability f . Using expression [5] and [6], the surplus is equal to

$$[7] \quad E - U = \frac{w(\tau)}{r + q + f}$$

On the other hand, we can use the zero profit condition to determine the surplus obtained by a firm with a filled job

$$[8] \quad F^E - F^V = \frac{y(\tau) + \psi\tau\sigma - w(\tau)}{r + q}$$

The Nash bargaining solution, with the hypothesis of equal bargaining power, implies that the wage is chosen to split equally the total surplus. Therefore, $E - U = F^E - F_M^V$ and

$$[9] \quad w(\tau) = \frac{[y(\tau) + \psi\tau\sigma](r + q + f)}{2(r + q) + f}$$

1.2 Training

Training is decided once and for all at time 0 and each firm chooses τ to maximize expected profits by taking as given the training level set by other firms, $\bar{\tau}$. Since firms are homogeneous, $\tau = \bar{\tau}$ in equilibrium. Therefore, in the initial period each firm maximizes

$$[10] \quad F^E(\tau) - c(\tau)$$

where c is the training cost. Substitution of (9) into (1) allows us to rewrite (10) as

$$[11] \quad F^E(\tau) - c(\tau) = \frac{y(\tau) + \psi\tau\sigma}{2(r+q) + f} - c(\tau)$$

The optimal level of training τ^* satisfies the following first order condition

$$[12] \quad \frac{y'(\tau^*) + \psi\sigma}{2(r+q) + f} - c'(\tau^*) = 0$$

Where the prime is for the first derivative. Training is positive when $2(r+q) + f < \infty$. In the following sub-section we investigate how training is affected by the size of knowledge spillovers σ .

1.3 Equilibrium

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

$$[13] \quad x(u, v) = mu^{1-\alpha}v^\alpha$$

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 α is a parameter smaller than one. With this technology we have:

$$[14] \quad h = \frac{m}{v} = m \left[\frac{v}{u} \right]^{\alpha-1} = m\theta^{\alpha-1}$$

$$[15] \quad f = \frac{m}{u} = m \left[\frac{v}{u} \right]^\alpha = m\theta^\alpha$$

where θ is the ratio of vacancies to unemployment. In equilibrium the inflow into the unemployment pool of trained workers must equal the outflow. Therefore

$$\dot{u} = (1-u)qL - m\theta^\alpha uL = 0$$

and

$$[16] \quad u = \frac{q}{q + m\theta^\alpha} \quad N = \left(1 - \frac{q}{q + m\theta^\alpha}\right)L$$

where N is employment. The equilibrium value of the vacancy-unemployment ratio θ is obtained by using (4) in (1), which yields

$$[17] \quad w(\tau) = y(\tau) + \psi\tau\sigma - (r+q)\frac{d}{h(\theta)}$$

In equilibrium (17) must be equal to the wage equation (9). Using (14) and (15), we can write this equality in implicit form as

$$[18] \quad A(\tau, \theta, \sigma) = y(\tau) + \psi\tau\sigma - [2(r+q) + f(\theta)]\frac{d\theta}{f(\theta)} = 0$$

where A is the labor market tightness equation in implicit form. Similarly, condition (12) can be written as

$$[19] \quad B(\tau, \theta, \sigma) = y'(\tau) + \psi\sigma - [2(r+q) + f(\theta)]c'(\tau) = 0$$

where B is the training equation in implicit form. Equations (18) and (19) determine the equilibrium values of τ and θ as functions of σ and other exogenous parameters. Differentiation of [18] and [19] yields

$$\begin{aligned} [20] \quad & A_\tau \partial \tau + A_\theta \partial \theta = -\tau \psi \partial \sigma \\ & B_\tau \partial \tau + B_\theta \partial \theta = -\psi \partial \sigma \end{aligned}$$

where

$$\begin{aligned} A_\tau &= y'(\tau) + \psi \sigma > 0 \\ B_\tau &= y''(\tau) - [2(r+q) + f(\theta)] c''(\tau) < 0 \\ A_\theta &= -d - \frac{2(r+q)d[f(\theta) - \theta f'(\theta)]}{[f(\theta)]^2} < 0 \\ B_\theta &= -c'(\tau) f'(\theta) < 0 \end{aligned}$$

in which the two primes are for the second derivative. Using Cramer's rule we get the key comparative statics

$$[21] \quad \frac{\partial \tau}{\partial \sigma} = \frac{\psi(A_\theta - B_\theta \tau)}{A_\tau B_\theta - B_\tau A_\theta}$$

$$[22] \quad \frac{\partial \theta}{\partial \sigma} = \frac{\psi(B_\tau \tau - A_\tau)}{A_\tau B_\theta - B_\tau A_\theta} > 0$$

An increase of the intensity of local knowledge spillovers σ leads to an increase in θ , since both the determinant $\Delta = A_\tau B_\theta - B_\tau A_\theta$ and the numerator are negative. The positive externalities associated to a higher value of σ make job creation more profitable and lead to a higher equilibrium ratio of jobs to workers. Conversely, the effect of local knowledge spillovers on training cannot be signed: on the one hand, an

increase of σ generates an increase of productivity and, as a consequence, encourages firms to invest in training; on the other hand, a higher value of σ leads to a higher θ , which increases the hazard rate from unemployment, and reduces the incentive to train, because wages are higher. Figure 1 shows the equilibrium in the training - tightness space. Notice that the slope of function A is positive: a higher level of training makes job creation more profitable, which increases the equilibrium ratio of jobs to workers.

When training is high the value of θ is also high because $\frac{\partial \tau}{\partial \theta} = -\frac{A_\theta}{A_\tau} > 0$. Moreover, since,

$\frac{\partial^2 \tau}{\partial \theta^2} < 0$ ¹, the curve is concave. On the other hand, the slope of the curve B is negative,

because $\frac{\partial \tau}{\partial \theta} = -\frac{B_\theta}{B_\tau} < 0$. When θ goes to zero the firm provides a level of training equal to

$c^{-1} \left[\frac{y'(\tau) + k\sigma}{2(r+q)} \right]$, while for θ tending to infinity training tends to zero. Moreover, since

$\frac{\partial^2 \tau}{\partial \theta^2} > 0$, the curve is convex. The equilibrium pair (τ^*, θ^*) lies at the intersection of the

two curves and, if it exists, is unique.

A higher value of σ shifts both the A and the B curve to the right. The overall effect on training can be both positive and negative, depending on the relative shifts of the two curves. On the other hand, the effect on the index of labor market tightness is always positive.

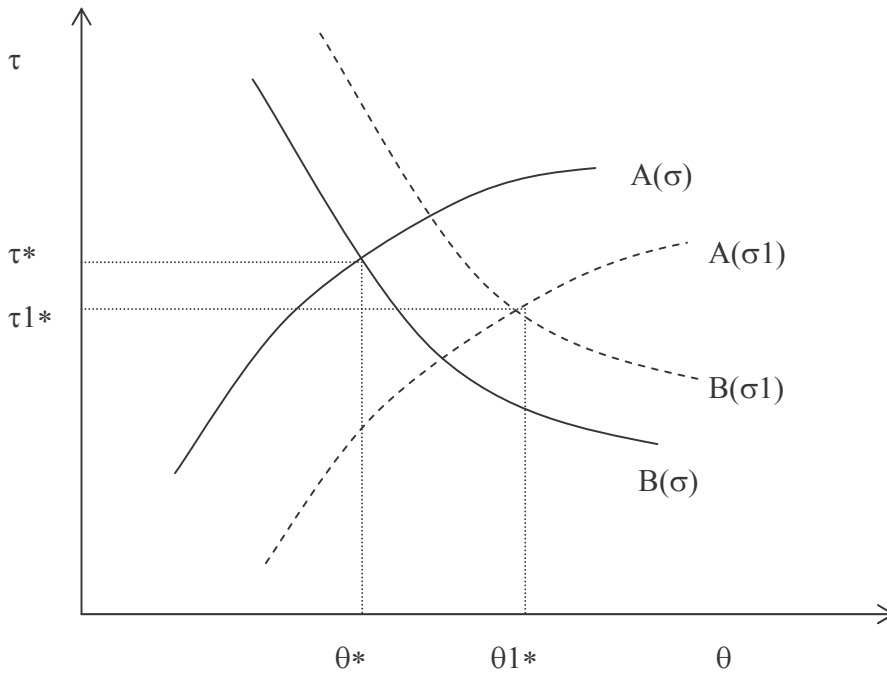
An additional link between labor market tightness and training is that turnover is higher in tighter labor markets. Following Pissarides, 2000, an increase in θ with on-the-job search reduces job duration. A higher θ implies faster job destruction because the worker is more likely to find another job 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 θ . Nothing of substance changes in the model, except that training is discouraged both by higher wages and higher labor market turnover.

¹ See the Appendix.

1.4 Endogenous Local Economic Density

In the previous sections we have defined σ as the size of local knowledge spillovers. In the relevant literature, these spillovers are captured by the local density of economic activity. According to Hall and Ciccone, 1996, the denser a local labor market the higher the opportunities of interactions and the faster the diffusion of knowledge. They measure local density as the number of employees per squared kilometer. Letting $\gamma = \frac{L}{K}$, where K is the size of the local labor market, local density is given by $\sigma = \gamma(1-u)$, which is an endogenous variable in our model because the unemployment rate depends on local 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 specific case.

Figure 1. The equilibrium in the training - tightness space



1.5 Endogenous Local Labor Force

In the previous sections we have considered a single labor market. With free entry, allowing for many different labor markets does not change the decisions of firms. Independently of where they localize, firms are bound by the zero profit condition for vacancies and therefore are indifferent between any geographical areas. On the other hand, workers can choose to move across geographical areas 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:

$$[23] \quad U_1 = U_2$$

Using (6) and (7) in (23), the arbitrage condition is satisfied when:

$$[24] \quad \frac{f_1(\theta_1)[y(\tau_1) + \tau_1\gamma_1(1 - u_1)]}{2(r + q) + f_1(\theta_1)} = \frac{f_2(\theta_2)[y(\tau_2) + \tau_2\gamma_2(1 - u_2)]}{2(r + q) + f_2(\theta_2)}$$

Since the total population is given and equal to $L = L_1 + L_2$, it follows that

$$[25] \quad \frac{L}{K} = \frac{K_1}{K} \frac{\gamma_1}{1 - u_1} + \frac{K_2}{K} \frac{\gamma_2}{1 - u_2}$$

These two equations determine the values of γ_1 and γ_2 .

2. The Data

In the empirical analysis we use firm level data drawn from the 8th survey “Indagine sulle Imprese Manifatturiere”, conducted by Mediocredito Centrale in 2001 with a questionnaire administered to a representative sample of Italian manufacturing firms². The sample is stratified for firms with up to 500 employees and covers the universe of firms with bigger employment size. Strata are based on the geographical area, the industry and the firm size. The survey collects background data on firms over a three-years period (1998, 1999, 2000), including firm location at the level of the municipality, the industry in which each firm operates, the year of birth, the total number of employees in each year (with a breakdown between white and blue-collar employees, the share of college and high school graduates³), 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.

The key question on training in the questionnaire is formulated as follows: “How many employees - during each of the years considered in the survey - have participated to training courses run by public or private organizations specialized in the provision of training?” 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-the-job and informal training. 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 learning, perhaps because by so doing they can reduce turnover. In this case the negative effect of density on off-the-job training needs 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 80% of the 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

² Mediocredito Centrale is an Italian investment bank. Detailed information about the surveys can be found in the Mediocredito Centrale’s web site, www.mcc.it

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 organizations⁴. 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 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 in the firm premises and via informal activity based on the mentoring by experienced workers.

Our sample consists of 3525 manufacturing firms. Of these, only 1272 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 10.56 for the year 2000. While comparisons are obviously difficult, this percentage is not very different from that reported by employee surveys 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 table also reports descriptive statistics for the firm-specific and area-specific variables used in the empirical analysis. The vast majority of the firms in our sample are small: about 76 percent has between 10 and 50 employees, 16% has 51 to 250 employees and the rest is bigger. About 56 percent of these firms are localized in the industrialized Northern part of the country, and only 11 percent are located in the South. The composition of the labor force by occupation is heavily concentrated in blue-collar jobs, which cover 66 percent of the total. Moreover, only 37 percent of the

³ The number of workers with temporary and part-time contracts, even if available, has not been used in our estimates because of the large number of missing values

⁴ Evidence on this is provided for Italy by Montanino (2000). According to the second Continued Vocational Training Survey, carried out 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.

employees have a high school diploma or higher education, and more than 8 percent of the firms in the sample belong to employer associations and networks (*consorzi*), which have been established for financial reasons and to promote exports and innovation. Finally, close to 40 percent of the sampled firms have received public subsidies, national and local, in the year 2000.

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). An alternative definition, corresponding to the statistical LLMA (local labor markets or travel to work areas), is in our view less appropriate both because local economic areas defined in such a way may result too small to take into account all relevant knowledge spillover⁵ and because of endogeneity problems, since LLMA have been defined according to economic factors, while provinces have been established mainly on the basis of political and bureaucratic reasons. Since LLMA is an interesting economic concept, however, in the robustness section of this paper we experiment with this alternative definition of local labor market.

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. These and other ISTAT data (Demographic Indicators 2000, Population Census 1991) are used to compute area-specific variables such as the percentage of college graduates, the average age of the population, the percentage of male employees in local industry and the local unemployment rate. It turns out that the average percentage of college graduates in the population (aged 6 years or older) is low, at 3.93 percent. On the other hand, unemployment is high, at 9.75 percent. Finally, the average age of employees in each area is slightly above 42 years.

As shown in Table 1, average total employment in the Nuts 3 areas in 2000 was 575 thousand employees, with a minimum of 32.7 thousand (Isernia) and a maximum of 2008 thousand (Milan). Average employment density was 236.4 thousand employees per squared kilometer, with a maximum value of 1014 and a minimum value of 12.8. In the

⁵ Paci and Usai (2000) discuss for Italy the presence of technological externalities across LLMA borders that decrease and die out as the distance from the local area increases.

same year industrial employment was on average 156.5 thousand employees with a minimum of 6.5 (Crotone) and a maximum of 620 (Milan). Average industrial employment density was 30.41 employees per squared kilometer, ranging between 313 and 2.5. 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 202 thousand employees per squared kilometre, with a range from 21.8 to 1014 thousand.

3. The Empirical Specification

Our measure of training T is the percentage of employees in each firm who received training during the year 2000. Since percentages are bounded between zero and one, we use the logistic transformation $\Sigma = \ln \frac{T}{1-T}$ (see Greene 1990). A feature of our data is that slightly less than one firm out of three provides information on T . We have little reason to expect that reporting training information is random. Therefore we use Heckman's correction to take into account endogenous selectivity. More precisely, let

$$[26] \quad \Sigma = \Gamma\beta + u_1$$

be the logistic equation, where Γ is a vector of controls to be specified later, and assume that Σ is observed if

$$[27] \quad Z\gamma + u_2 > 0$$

where $u_1 \approx N(0, \sigma^2)$, $u_2 \approx N(0, 1)$, $\text{corr}(u_1, u_2) = \rho$ and Z is a vector of controls, which includes at least one variable excluded from Σ . We estimate the training equation (26) and the selection equation (27) jointly by maximum likelihood. The vector Γ includes

both firm-specific and area-specific variables. Among the former we include a measure of average educational attainment (the percentage of employees with at least upper secondary education), the percentage of blue 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 workforce⁶. Since training and innovation are complements (see Acemoglu 1997), 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. Our survey includes two indicators of the degree of labor and product market competition in the area, the participation in networks and the presence of local competitors, and we define a dummy indicating whether the firm has competitors located in the same area and another dummy indicating whether the firm belongs to a production or financial network or employer association (*consorzio*). 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 set of area-specific variables includes log industrial employment density, our key explanatory variable, and a set of regional dummies to control for region-specific shocks. Since our measure of density is 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 selection equation includes all the firm and area-specific variables listed above plus the real value of sales in 2000, both alone and interacted with the dummy S , equal to 1 if the firm operates in the South of Italy and to zero otherwise. We also include the interaction of the dummy S with the dummy indicating the presence of public incentives.

⁶We do not have information on the composition of the workforce by gender. Since manufacturing sectors have different shares of female labor in total employment, industry dummies should capture part of the gender effect.

Since density varies only across provinces, it is important that we try to control as well as possible for other province-specific effects, such as the slackness of the local labor market – measured by the unemployment rate in the province. If a province has higher unobserved productivity, it is likely to attract better educated workers. Since education and training are complements, labor mobility can generate a positive spurious correlation between training and density. For this reason, we include the percentage of college graduates in the province among the regressors. We also control for the composition of the local labor force with the average age and the percentage of male employees in local industry.

Equally dense areas can vary significantly in the degree of industrial specialization. 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. In the literature this is an open issue, and it is not clear whether MAR (Marshall- Arrow-Romer) externalities, driven by industrial specialization, prevail over Jacobs externalities, originated by the diversity of the production structure (see Glaeser et al, 1992). Therefore, we include among the regressors the following index of industrial specialization (see Ciccone and Cingano, 2000)

$$[28] \quad S_j = \sum_k \left(\frac{E_{kj}}{E_j} \right)^2$$

where j is the province and k is the industry. This index is higher the more specialized the local industrial structure. Finally, we add to the list of explanatory variables a dummy equal to one if the firm belongs to an industrial district and to zero otherwise.

The presence of industrial districts, defined as clusters of firms involved, at various stages and in various patterns, in the production of homogeneous goods is an interesting feature of Italian medium and small firms (see Signorini, 2000, and Gambarotto, 2002). Industrial districts 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 (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.

Since the province-specific variables are at a different level of aggregation than the firm-specific variables, we adjust the standard errors of the estimated parameters for the lack of independence of errors within clusters. In spite of the inclusion of province-specific controls, we cannot rule out the possibility that the uncovered correlation between training and density in a cross section of firms be spurious because of omitted factors, which affect both variables. We try to take the potential endogeneity of density into account by using instrumental variables. Even assuming that local economic density is (weakly) exogenous in the current setup, IV estimation is desirable because density can be measured with error. 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 provinces⁷ 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 firms⁸. By so doing we reduce measurement error. Since this error cannot be wholly eliminated, however, we also use IV estimates.

Our preferred measure of density is the log of the ratio between industrial employment and the size of the province expressed in squared kilometers. The focus on

⁷ We have no information on whether a firm in our sample has more than one plant and on the location of these plants,

⁸ In the Appendix we also present robustness exercises using all firms and the narrower subset of firms with at most 250 employees.

industrial density is appropriate because the firms in our sample belong to the manufacturing sector. We also experiment, however, both with a wider definition, which uses total employment as the numerator, and with an industry-specific measure, where the numerator is employment in the industry to which the firm belongs.

4. The Results

Our main findings are illustrated in Table 2, where we report the estimated coefficients of the training equation and of the selection equation. A Wald test of independent equations clearly rejects the null of independency in all the selected specifications. The first column in the table includes regional dummies among the regressors, but excludes all province-specific controls a part from log density. The following three columns add the local rate of unemployment in 2000, the percentage of college graduates in the provincial population, the average age of employees, the percentage of male workers in industry and the index of industrial specialization. The last two columns add the industrial district dummy (column V) and this dummy plus the interaction between log density and the index of specialization (column VI).

We find that training incidence is higher when the percentage of blue-collar workers in the firm is lower and the share of employees with at least upper secondary education is higher. As expected, training is higher in younger firms, established since 1990, even though the relevant coefficient is not always precisely estimated.

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. Perhaps more interestingly, we also find that training is higher among firms which belong to formal networks and associations of firms, which specialize in financial assistance, export promotion, research and development and else. This might depend on the fact that these associations make member firms more productive

and able to capture profit opportunities and knowledge spillovers (see Hirshman, 1977; Podolny and Baron, 1997)⁹. It is important, however, not to confound these associations with industrial districts, and we find no evidence that being located in such a district has affected, *ceteris paribus*, the percentage of trained employees.

The key result of the table is that training is higher in provinces with lower employment density. While the relationship between density and training is statistically significant, the size of the effect is rather small: when evaluated at the sample average percentage of trained employees, we find that a 10 percent increase in local density reduces the percentage of trained employees in the range of 0.8 to 1.2 percent, depending on the specification¹⁰. The negative correlation between local density and off-the-job training is affected neither by the inclusion of the local unemployment rate – a negative but often not statistically significant coefficient – nor by the inclusion of the percentage of college graduates – which attracts a positive and statistically significant coefficient. Interestingly, we find evidence that training is more likely the higher the index of industrial specialization. On the other hand, and conditional on provincial variables, we find no evidence that the percentage of training employees varies significantly with firm size.

Firm size matters instead in the selection equation, as larger firms are more likely to report the percentage of trained employees than smaller units. The additional variables not included in the training equation are log real sales both alone and interacted with a South dummy and the interaction of the incentives dummy with the South dummy. Firms, which receive incentives, are more likely to report training – perhaps because they need to justify at least part of the public money received. This incentive is smaller, however, for firms located in the South. Conditional on firm size, higher real sales increases the likelihood of reporting training figures. This increase is relatively larger for firms located in the South.

Notice that the probability of reporting training figures is lower in denser areas, even though the estimated coefficient is not statistically significant. Finally, the

⁹ Several sociological studies point out that networks, by promoting social interaction, favor the transfer of tacit knowledge.

estimated correlation between the error terms in the training and in the selection equation is very high – close to 0.98 – which suggests that endogenous selection is a relevant problem in our data.

As shown in the Appendix, our findings are robust both to a larger sample, which includes all firms, and to the sub-sample of firms with less than 250 employees. They are also robust to a definition of density, which includes all employment, rather than only industrial employment in the numerator, as shown in Table 3, where we only report the estimates of the training equation to save space.

The uncovered negative correlation between training and local density confirms previous results based on a different dataset and a different country. Brunello and Gambiarotto, 2004, study this relationship using individual rather than firm data and find similar results for the UK. They also find, however, that training incidence is lower, conditional on density, when the specialization of economic production is higher. In this paper we find evidence to the contrary, with training being higher in areas with a higher specialization of industrial production.

The model in Section 2 of the paper provides a framework for the interpretation of our results. To this purpose, it is useful to think of Figure 1 as the description of equilibrium configurations in two local labor markets, which differ only in the density of economic activity. In the figure, the less dense labor market is characterized by continuous lines and the denser market by broken lines. The denser labor market influence training incidence for two reasons. First, the more intense knowledge spillovers increase productivity and the marginal benefits of training, which are higher for any value of θ . Second, higher positive externalities in the denser labor market increases job creation and vacancies, which increase wages and turnover. For any value of training, while the former effect increases training in equilibrium, the latter effect reduces it. Our results suggest that, in the Italian institutional environment, it is the latter effect which prevails.

Conditional on density, the effects of industrial specialization on training are ambiguous: higher specialization increases the importance of MAR externalities and

¹⁰ The elasticity is obtained by multiplying the estimated coefficient of log density reported in the table by $(1-T)$, where T

raises at the same time the probability of turnover. The balance in this dataset is in favor of the positive effects. As shown in the next section, this result is robust to changes in the definition of economic density and is confirmed by the positive sign attracted by the coefficient of local competition, which captures whether the firm has competitors located in the same area.

Previous results based on a different dataset go, however, in the opposite direction – more specialization, less training (see Brunello and Gambarotto, 2004) - and do not allow us to generalize our findings. These diverging results could depend on the differences in the composition of manufacturing in each country. The Italian industrial system is especially characterized by traditional manufacturing, with industries such as Machinery and Fabricated Metal Products accounting for an important share of the total. Conversely, the UK industrial structure is characterized by a broader presence of innovative industries, such as Chemicals and Aircraft.¹¹ A number of studies have shown that the agglomeration of firms belonging to the same industry plays a decisive role in the creation of competitive advantages in the mature technological sector; high-tech sectors, however, which are characterized by faster technological change, take advantage mainly from the variety of local activities (Henderson et al., 1995). This difference might explain why industrial specialization produces a positive effect on training in Italy, and a negative effect in the UK: only in the former case are MAR externalities sufficiently high to compensate for the higher risk of turnover.

5. Robustness

In this section we examine the robustness of our results to alternative definitions of employment density and local labor markets. We also consider whether our estimates are affected when we take into account the potential endogeneity of log density in the training and selection equations.

is evaluated at its sample mean. We exclude the last column from the computation of the elasticities.

¹¹ If we look to specialization ratios, which compare for a given country the importance of a particular NACE group in total manufacturing value added with the same ratio for the EU as a whole, we notice that Italy shows the highest ratios in Ceramic Tiles and Motorcycles-Bicycles and UK records the highest ratios in Aircraft-Spacecraft and Pesticides and other Agro-Chemical products (EUROSTAT, European business - Facts and figures - Data 2001).

In our framework density measures the importance of economic spillovers associated to agglomeration. If these spillovers occur mainly, if not exclusively, within an industry, however, the relevant measure of density for a firm operating in a given industry is employment in that industry per squared kilometer. Therefore we decompose density in the province into two components: a) the log density of employment in the industry to which each firm in the sample belongs; b) the log density of employment in the remaining industries located in the same province. Conditional on the first component, the second captures in a gross way the variety of industrial production in the area. The combined effect of a) and b) should capture both specialization and diversification effects. For this reason, we omit the index of specialization from the estimates presented in Table 4¹².

The results in the table show that both measures of density attract a negative coefficient. Nevertheless, only the coefficient associated to the density of employment of the other industries is statistically significant. These results suggest that variety of industrial production in a local area discourages training, and confirm the positive relationship between industrial specialization and training found in the previous table.

Next, we turn to consider the potential endogeneity of log density in the training and selection equations. This endogeneity is driven by the mobility of individuals among local areas: the higher density of some geographical areas depends on the fact that these areas are more productive and attract talented individuals, who in turn are more likely to be trained. To the extent that unobserved ability and productivity is not captured by province-specific controls, such as the percentage of graduates and the local unemployment rate, estimates are likely to be biased. We expect the sign of the bias to be positive, because areas with higher unobserved (to the econometrician) productivity attract skilled individuals whose probability to be involved in training is higher.

Following Ciccone, 2001, we instrument density with the size of the area, measured in squared kilometers. Since provinces in Italy are defined mainly as administrative units, their size is unlikely to reflect the average ability or productivity of its working population. We also use as instruments the interactions of size with

¹² We use the standard measure of density in the selection equation.

regional dummies and run a first stage regression with log density as the dependent variable and the set of exogenous variables and additional instruments as the explanatory variables. The F-test on the additional instruments is equal to 82.40 and clearly rejects the null of no joint significance¹³. We use the estimates of the first stage regression to predict log density and replace actual density in the training and selection equations with its predicted value. Table 5 presents the results of the two stages least squares estimate of the baseline model reported in Table 2. These findings confirm the presence of a negative and statistically significant relationship between the percentage of trained employees and local density. If anything, the size of the estimated coefficients is larger in absolute value, as expected. Based on the IV estimates, the expected reduction in the percentage of trained employees when local density increases by 10 percent is approximately equal to 1.6 percent.

Finally, we use an additional and potentially interesting definition of the local area: the LLMA (local labor market), a self-contained geographical area, capable of offering employment to the majority of its resident population. The degree of self-containment is measured by commuting behavior: a geographical area is classified as a LLMA if the ratio between the number of persons who work and live there and the number of persons who work but do not live in the area is above 70-75 percent. There are 784 such LLMA in Italy, and clearly they are much more detailed than the close to 100 Italian provinces.¹⁴

Table 6 replicates Table 5 using log employment density at the LLMA level¹⁵ and the size of the local area in squared kilometers as an instrument of density. Notice that this instrument is potentially weaker than the size of the province, because LLMA are defined mainly according to mobility patterns. This explains our preference for NUTS 3 areas. The 2SLS results in the table confirm the basic thrust of our results in that the estimated coefficient associated to log density is always negative and often statistically significant. We conclude that our results are robust to changes in the definition of the local area.

¹³ The results of the first stage regression are available from the authors upon request.

¹⁴ The average total employment in LLMA is about 227 thousand employees, with a minimum value of 1 thousand employees and a maximum value 1 million and 260 thousand employees (Rome).

7. Conclusions

Economic density, according to a growing body of research, encourages shared learning among individuals and firms and generate positive knowledge spillovers. Since the ability to translate information in real economic advantages is strictly related to the skills of the labour 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 off-the-job training is higher in provinces with lower employment density. This result confirms the evidence presented by Brunello and Gambiarotto, 2004, for the UK.

An interesting difference emerging from results obtained in this study and those obtained by Brunello and Gambiarotto for the UK is that, while they show a negative effect of industrial specialization on training, we find a positive effect, which suggests that higher MAR externalities prevail over the higher turnover risk associated to a specialized industrial structure. We reconcile these diverging results for Italy and the UK by arguing that they are related to differences in the industrial composition of each country. Italy, being especially characterized by traditional manufacturing, might experiment stronger MAR externalities compared to UK, where the number of innovative industries is higher. As argued by Henderson et al (1995), while high-tech industries benefit from industrial variety and Jacobs externalities, in traditional manufacturing the agglomeration of firms from the same industry plays the decisive role in the creation of competitive advantage.

¹⁵ We use total employment rather than individual employment.

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

References

- Acemoglu, D. and Pischke, J. [1999], The structure of wages and investment in general training, *Journal of Political Economy*, 107, 539-572.
- Acemoglu, D., [2000], Technical Change, Inequality and the Labor Market, *NBER Working Paper* n. 7800
- Anselin, L., Varga, A., and Acs, Z. [1997], Local Geographic Spillovers between University Research and High Technology Innovations, *Journal of Urban Economics*, 42, pp. 422-48.
- Audretsch, D. B., and Feldman, M.P. [1996], “R&D Spillovers and the Geography of Innovation and Production,” *American Economic Review*, 86, pp. 630-40.
- Bala V. and S. Goyal [1998], Learning from Neighbors, *Review of Economic Studies*, 65, pp. 595-621.
- Bala V. and S. Goyal [2000], A Noncooperative Model of Network Formation, *Econometrica*, 68, pp. 1181-1229.
- Barron, J., Berger, M., and Black, D., [1999], Replacing General Training with Specific Training: Why Restricting Alternatives Makes Sense, *Research in Labour Economics*, 18,
- Barron, J., Black, D. and Loewenstein, M., [1989], “Job Matching and On-the-job Training”, *Journal of Labor Economics* 7 (1), 1-19.
- Blanchard, O. and Diamond, P. [1989], “The Beveridge Curve,” *Brooking Papers on Economic Activity*, vol. 1, pp. 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*, Vol. 16, pp. 19-87.
- Breschi S., and Lissoni F., [2001], Knowledge spillovers and local systems of innovation: A critical survey, *Industrial and Corporate Change*, 10(4), 975-1005.
- Brunello, G. and Medio, A., [2001], An explanation of international differences in education and workplace training, *The European Economic Review*, pp.307-322.

Brunello, G. and Gambarotto, F. [2004], Agglomeration Effects on Employer-Provided Training: Evidence from the UK, IZA DP n.1055.

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

Ciccone, A. and Cingano F. [2002], Skills and clusters, Bank of Italy, mimeo.

Cingano, F. and Schivardi, F., [2003], Identifying the sources of local productivity growth, Bank of Italy, mimeo.

Combes, P. and Duranton G., [2001], Labor pooling, labor poaching and spatial clustering, LSE Discussion Paper 2975.

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, 451-470.

Feldman M.P. [1994], *The geography of innovation*, Kluwer Academic, Boston.

Gambarotto, F., [2002], Le relazioni economico – istituzionali nello spazio delle economie locali: un'ipotesi teorica, in Camagni, R. and Fiorentini, R., eds, Scritti in memoria di Eugenio Benedetti, Cedam, Padova

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

Greene, WA, [2000], *Econometric Analysis*, 4th edition, Prentice Hall.

Hirschman, A.O., [1958], *The Strategy of Economic Development*. Yale University Press, New Haven, CT.

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

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

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

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

- Marshall, A., [1890], *Principles of Economics*, London: Macmillian.
- Mortensen, D. and Pissarides, C. [1999], New Developments in Models of Search in the Labour Market, in Ashenfelter, O. and Card, D., (eds.), *Handbook of Labor Economics*.
- Montanino, A. , [2001], Formazione aziendale, struttura dell'occupazione e dimensione dell'impresa, *Rivista di Politica Economica*,
- Neal, D., [1995], Industry-Specific Human Capital: Evidence from Displaced Workers Journal of Labor Economics, 13(4): 653-77
- Paci, R. and Usai S., [2000], Externalities, Knowledge Spillovers And The Spatial Distribution Of Innovation, *European Regional Science Association* conference papers n.00, p104.
- 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, 673-93.
- Sedgley, N., and Elmslie, B., [2001], The Geographic Concentration of Knowledge: Scale, Agglomeration and Congestion in Innovation across U.S. States, unpublished manuscript.
- Signorini, L. F. [2000], *Lo sviluppo locale. Un'indagine della Banca d'Italia sui distretti industriali*, Roma, Donzelli.
- Stevens, M. [1995], Transferable training and poaching externalities, in Booth and Snower, *Acquiring Skills*, Cambridge University Press, Cambridge, pp. 19-40.
- Storper, M. and Venables, T., [2004], Buzz: the economic force of the city, mimeo
- Rosenthal, S. and Strange, W., [2003], Geography, industrial organization and agglomeration, *The Review of Economics and Statistics*, 85, 2, pp. 377-393.
- Von Hippel E. [1994], Sticky Information and the Locus of Problem Solving: Implications for Innovation *Management Science*, 40, pp. 429-439.

Table 1. Descriptive statistics

	Mean	St. Dev
Percentage of trained employees	.105	.154
Percentage of firms established after 1990	.214	.410
Percentage of firms participating in networks	.088	.283
Share of R&D expenditure over total sales in 1999	.006	.019
Percentage of employees with upper secondary education or higher	.373	.249
Percentage of blue collars	.661	.169
Percentage of firm which received public subsidies	.398	.489
Dummy for the presence of local competition	.112	.315
Percentage of firms with 10-20 employees	.397	.489
Percentage of firms with 21-50 employees	.371	.483
Percentage of firms with 51-250 employees	.164	.361
Percentage of firms with 251-500 employees	.039	.194
Percentage of firms with >500 employees	.026	.161
Percentage of firms located in the North of the country	.556	.496
Average age in the province	42.508	2.135
Percentage of firms belonging to an industrial district	.377	.484
Percentage of males employed in industry and area	.049	.092
Local unemployment rate	.097	.077
Percentage of college graduates in the local population	.039	.012
Total employment in the area	571.533	621.466
Local employment density (total employment/KM)	236.40	303.83
Local industrial density (industrial employment/KM)	78.01	9.19
Region specific index of specialization in the Nuts 3 area	.160	.106

Table 2. Training regression. Baseline model for firms with 11-500 employees. Dependent variable: $\ln(T/(1-T))$. With Heckman correction for selectivity. Log density measured as the log of the ratio between industrial employment and squared kilometers of the area

	(1)	(2)	(3)	(4)	(5)	(6)
Training equation						
expenditure in RD in 1999	4.7968* (2.2910)	4.6641* (2.3313)	4.5285 (2.4788)	4.8822 (2.5077)	4.8719 (2.4970)	4.8460 (2.4782)
networks	0.4356** (0.1255)	0.4312** (0.1261)	0.4928** (0.1294)	0.5155** (0.1320)	0.5149** (0.1317)	0.5142** (0.1314)
% with upper secondary ed. or higher	0.6501** (0.2063)	0.6463** (0.2047)	0.7024** (0.2083)	0.7171** (0.2131)	0.7160** (0.2135)	0.7200** (0.2129)
incentives	0.3838** (0.0699)	0.3827** (0.0694)	0.4087** (0.0739)	0.3828** (0.0739)	0.3817** (0.0729)	0.3812** (0.0728)
Established after 1990	0.1492 (0.1037)	0.1519 (0.1041)	0.1662 (0.1106)	0.2010 (0.1061)	0.2101* (0.1060)	0.2109* (0.1055)
local competition	0.0475 (0.1108)	0.0426 (0.1111)	0.0187 (0.1156)	0.0071 (0.1162)	0.0056 (0.1152)	0.0029 (0.1150)
percentage blue collars	-1.1564** (0.2976)	-1.1646** (0.2972)	-1.0958** (0.3085)	-1.1321** (0.3155)	-1.1319** (0.3156)	-1.1208** (0.3186)
size 21-50	0.0455 (0.0941)	0.0442 (0.0936)	0.0811 (0.0852)	0.0744 (0.0844)	0.0734 (0.0848)	0.0718 (0.0843)
size 51-250	-0.0142 (0.0950)	-0.0157 (0.0943)	-0.0114 (0.0932)	0.0095 (0.0902)	0.0087 (0.090)	0.0062 (0.0901)
Size 251-500	0.2862 (0.1879)	0.2875 (0.1865)	0.3612 (0.1982)	0.2631 (0.1990)	0.2621 (0.1996)	0.2577 (0.1986)
log density	-0.1129* (0.0477)	-0.0967* (0.0450)	-0.1342* (0.0585)	-0.1234* (0.0563)	-0.1246* (0.0554)	-0.2330 (0.1358)
local rate of unemployment		-2.0146 (1.8289)	-2.7448 (1.4441)	-3.3326* (1.3557)	-3.2725* (1.3590)	-3.2744* (1.3498)
% graduates in the area			6.1298* (2.4123)	6.6657** (2.4474)	6.8731** (2.3829)	7.3248** (2.6113)
average age			-0.0084 (0.0101)	-0.0107 (0.0105)	-0.0106 (0.0106)	-0.0093 (0.0109)
percentage males			0.7268* (0.3258)	0.8457** (0.3005)	0.8431** (0.3024)	0.8157** (0.2812)
specialization index				0.9412 (0.6114)	0.9071 (0.6485)	4.0585 (3.9801)
industrial district					0.0162 (0.0567)	0.0155 (0.0575)
log density * specialization						0.8725 (1.0778)

Selection equation

Established after 1990	0.0880 (0.0577)	0.0880 (0.0575)	0.0946 (0.0594)	0.1176 (0.0588)	0.1178 (0.0587)	0.1189 (0.0586)
networks	0.2045** (0.0703)	0.2067** (0.0705)	0.2045** (0.0695)	0.2244** (0.0694)	0.2245** (0.0694)	0.2237** (0.0694)
local competition	-0.0032 (.0704)	-0.0043 (0.0700)	-0.0134 (0.0690)	-0.0196 (0.0681)	-0.0196 (0.0681)	-0.0188 (0.0681)
log sales	0.0530* (0.0255)	0.0542* (0.0255)	0.0666** (0.0226)	0.0646** (0.0230)	0.0649** (0.0231)	0.0653** (0.0233)
log sales * South	0.0731 (0.0545)	0.0681* (0.0545)	0.0485 (0.0514)	0.0720 (0.0473)	0.0730 (0.0478)	0.0695 (0.0486)
incentives	0.2309** (0.0376)	0.2292** (0.0371)	0.2257** (0.0385)	0.2165** (0.0382)	0.2168** (0.0384)	0.2158** (0.0387)
incentives * South	-0.2438* (0.0941)	-0.2288* (0.0939)	-0.1836* (0.0908)	-0.0958 (0.0868)	-0.0990 (0.0883)	-0.0949 (0.0899)
expenditure in RD in 1999	2.3147 (1.4353)	2.3743 (1.4371)	2.1955 (1.4606)	2.5573 (1.5031)	2.5609 (1.5072)	2.5654 (1.5115)
percentage blue collars	-0.4355** (0.1605)	-0.4335** (0.1603)	-0.4112** (0.1617)	-0.4242** (0.1652)	-0.4240** (0.1647)	-0.4233** (0.1647)
% with upper secondary ed. or higher	0.3124** (0.1151)	0.3096** (0.1146)	0.3422** (0.1154)	0.3525** (0.1193)	0.3522** (0.1195)	0.3521** (0.1195)
size 21-50	0.3118** (0.0490)	0.3116** (0.0490)	0.3030** (0.0462)	0.3012** (0.0461)	0.3009** (0.0459)	0.3010** (0.0462)
size 51-250	0.7089** (0.0649)	0.7080** (0.0649)	0.6905** (0.0619)	0.7211** (0.0587)	0.7207** (0.0589)	0.7199** (0.0593)
Size 251-500	1.3219** (0.1404)	1.3192** (0.1391)	1.3063** (0.1387)	1.2693** (0.1347)	1.2680** (0.1364)	1.2694** (0.1366)
log density	-0.0482 (0.0319)	-0.0479 (0.0318)	-0.0411 (0.0299)	-0.0469 (0.0286)	-0.0469 (0.0286)	-0.0467 (0.0288)
rho	0.9847** (0.0069)	0.9847** (0.0065)	0.9868** (0.0059)	0.9872** (0.0065)	0.9870** (0.0067)	0.9867** (0.0068)
sigma	1.7015** (0.08479)	1.7002** (0.0810)	1.7400** (0.0826)	1.7232** (0.0877)	1.7212** (0.0888)	1.7180** (0.0886)
# obs :	3494	3494	3392	3357	3357	3357

=====

Note: robust and cluster adjusted standard errors in parentheses with p<0.05 = *, p<0.01 = **.

A constant plus industry and regional dummies are included in both equations.

Table 3. Training regression. Baseline model for firms with 11-500 employees. Dependent variable: $\ln(T/(1-T))$. With Heckman correction for selectivity. Log density measured as the log of the ratio between total employment and squared kilometers of the area.

	(1)	(2)	(3)	(4)	(5)	(6)
expenditure in R&D	4.8682* (2.2857)	4.7223* (2.3271)	4.5758 (2.4798)	4.9405* (2.5063)	4.9317* (2.4963)	4.8703* (2.4705)
networks	0.4404** (0.1251)	0.4348** (0.1260)	0.4941** (0.1296)	0.5164** (0.1320)	0.5158** (0.1317)	0.5151** (0.1314)
% with upper sec or mopre	0.6456** (0.2064)	0.6411** (0.2045)	0.6936** (0.2091)	0.7096** (0.2137)	0.7086** (0.2141)	0.7158** (0.2130)
incentives	0.3802** (0.070)	0.3796** (0.0696)	0.4047** (0.0739)	0.3792** (0.0739)	0.3782** (0.0730)	0.3783** (0.0727)
established after 1990	0.1501 (0.1039)	0.1526 (0.1043)	0.1668 (0.1107)	0.2101* (0.1060)	0.2101* (0.1060)	0.2120* (0.1052)
local competitors	0.0413 (0.1109)	0.0366 (0.1111)	0.0126 (0.1159)	0.0012 (0.1164)	-0.0002 (0.1154)	-0.0041 (0.1148)
% blue collars	-1.1557** (0.2982)	-1.1636** (0.2977)	-1.1031** (0.3084)	-1.1370** (0.3155)	-1.1366** (0.3155)	-1.1187** (0.3183)
size 21-50	0.0458 (0.0943)	0.0444 (0.0937)	0.0817 (0.0852)	0.0737 (0.0845)	0.0728 (0.0848)	0.0708 (0.0842)
size 51-250	-0.0167 (0.0950)	-0.0182 (0.0944)	-0.0120 (0.0933)	0.0076 (0.0903)	0.0069 (0.090)	0.0027 (0.090)
size 251-500	0.2920 (0.1888)	0.2918 (0.1873)	0.3678 (0.1987)	0.2665 (0.1988)	0.2657 (0.1995)	0.2614 (0.1973)
log density	-0.1042* (0.0490)	-0.0853 (0.0460)	-0.1503** (0.0572)	-0.1357* (0.060)	-0.1365* (0.0595)	-0.3142* (0.1376)
local rate of unemployment		-2.0711 (1.8329)	-2.5284 (1.4636)	-3.1004* (1.3919)	-3.0493* (1.3891)	-2.9842* (1.3778)
% graduates in the area			7.5629** (2.7924)	7.7371** (2.8511)	7.9262** (2.7804)	8.7069** (2.9339)
average age			-0.0086 (0.010)	-0.0111 (0.0104)	-0.0109 (0.0104)	-0.0093 (0.0107)
% males			0.7535* (0.3136)	0.8741** (0.2953)	0.8728** (0.2973)	0.8166** (0.2672)
specialization index				0.8413 (0.6453)	0.8105 (0.6788)	4.5212 (2.8143)
industrial district					0.0146 (0.0565)	0.0140 (0.0577)
log density *specialization						1.4108 (1.0195)
# obs :	3497	3497	3395	3360	3360	3360

Note: see Table 2. The selection equation is not reported.

Table 4. Training regression. Dependent variable: $\ln(T/(1-T))$. With Heckman correction for selectivity. Log density measured as the log of the ratio between employment in the specific industry and squared kilometers of the area

	(1)	(2)	(3)	(4)
expenditure in RD	5.1973* (2.1196)	5.0799* (2.1293)	4.9454* (2.2579)	4.9109* (2.2567)
networks	0.4562** (0.1342)	0.4519** (0.1340)	0.5139** (0.1420)	0.5134** (0.1418)
% with upper sec. Or more	0.6909** (0.2025)	0.6852** (0.2010)	0.7194** (0.2071)	0.7168** (0.2072)
incentives	0.3656** (0.0873)	0.3654** (0.0870)	0.3827** (0.0934)	0.3800** (0.0937)
established after 1990	0.1765 (0.0991)	0.1793 (0.0990)	0.2073* (0.1047)	0.2069* (0.1046)
local competition	0.0353 (0.1195)	0.0311 (0.1192)	0.0211 (0.1299)	0.0367 (0.0770)
percentage blue collars	-1.1546** (0.3003)	-1.1630** (0.3000)	-1.1285** (0.3148)	-1.1289** (0.3154)
size 21-50	0.0396 (0.0920)	0.0380 (0.0920)	0.0699 (0.0943)	0.0684 (0.0945)
size 51-250	-0.0118 (0.1143)	-0.0133 (0.1138)	0.0091 (0.1212)	0.0092 (0.1211)
size 251-500	0.1889 (0.2052)	0.1903 (0.2045)	0.2566 (0.2162)	0.2568 (0.2150)
Log density Industry	-0.0605 (0.0350)	-0.0572 (0.0354)	-0.0467 (0.0413)	-0.0539 (0.0455)
log density oth. ind.	-0.1229* (0.050)	-0.1057* (0.0518)	-0.1522** (0.0571)	-0.1531** (0.0570)
local rate of unemployment		-1.7534 (1.4610)	-2.5821 (1.5878)	-2.4459 (1.6237)
% graduates in the area			6.7370* (3.0907)	7.2607** (3.0828)
average age			-0.0138 (0.0122)	-0.0134 (0.0122)
percentage males			0.8668* (0.3613)	0.8468* (0.3656)
Industrial district				0.0375 (0.0646)
# obs :	3452	3452	3355	3355

Note: see Table 2. The selection equation is not reported.

Table 5 Training regression. Baseline model for firms with 11-500 employees. Dependent variable: $\ln(T/(1-T))$. Two stages least squares with Heckman correction for selectivity. Log density measured as the log of the ratio between industrial employment and squared kilometers of the area.

	(1)	(2)	(3)	(4)	(5)	(6)
expenditure in RD	4.6786* (2.2415)	4.4139 (2.2559)	4.4798 (2.3845)	4.8158* (2.4446)	4.8013* (2.4313)	4.7998* (2.4187)
networks	0.4836** (0.1281)	0.4396** (0.1241)	0.5364** (0.1287)	0.5632** (0.1315)	0.5615** (0.1312)	0.5615** (0.1312)
% with upper Sec. or more	0.6736** (0.2159)	0.6196** (0.2086)	0.6566** (0.2188)	0.6855** (0.2236)	0.6836** (0.2243)	0.6836** (0.2243)
incentives	0.3746** (0.0717)	0.3671** (0.0680)	0.3927** (0.0720)	0.3713** (0.0727)	0.3691** (0.0716)	0.3690** (0.0714)
established after 1990	0.1688 (0.1043)	0.1492 (0.1034)	0.1610 (0.1091)	0.2147* (0.1024)	0.2151* (0.1024)	0.2150* (0.1027)
local competition	0.0184 (0.1106)	0.0310 (0.1092)	0.0381 (0.1173)	0.0362 (0.1170)	0.0321 (0.1160)	0.0321 (0.1160)
percentage blue collars	-1.1366** (0.3147)	-1.2007** (0.3078)	-1.1794** (0.3222)	-1.2192** (0.3275)	-1.2178** (0.3277)	-1.2173** (0.3308)
size 21-50	0.0420 (0.0872)	0.0566 (0.0931)	0.0919 (0.0840)	0.0861 (0.0834)	0.0837 (0.0835)	0.0837 (0.0835)
size 51-250	-0.0039 (0.0958)	0.0027 (0.0943)	0.0255 (0.0912)	0.0485 (0.0895)	0.0465 (0.0894)	0.0464 (0.0894)
size 251-500	0.2708 (0.1986)	0.3000 (0.1901)	0.2903 (0.2015)	0.1915 (0.2022)	0.1897 (0.2023)	0.1895 (0.2024)
log density	-0.2149** (0.0718)	-0.1409 (0.0804)	-0.2278** (0.0761)	-0.2216** (0.0845)	-0.2227** (0.0831)	-0.2278 (0.1830)
local rate of unemployment		-1.8589 (2.0336)	-1.8893 (1.5983)	-2.0702 (1.5683)	-1.9285 (1.6013)	-1.9266 (1.6298)
% graduates in the area			8.3110* (3.0066)	9.8846** (2.9878)	10.4147** (3.1644)	10.4229** (3.2161)
average age			-0.0116 (0.0107)	-0.0187 (0.0111)	-0.0183 (0.0109)	-0.0183 (0.0109)
percentage males			0.5792 (0.3291)	0.7112* (0.2910)	0.7032* (0.2932)	0.7035* (0.2934)
specialization index				0.5914 (0.6276)	0.5201 (0.6699)	0.6719 (4.4072)
industrial district					0.0397 (0.0581)	0.0395 (0.0581)
log density *specialization						0.0431 (1.2613)
# obs :	3297	3497	3136	3049	3049	3049

Note: see Table 2. The selection equation is not reported.

Table 6 Training regression. Baseline model for firms with 11-500 employees. Dependent variable: $\ln(T/(1-T))$. Two stages least squares with Heckman correction for selectivity. Log density measured as the log of the ratio between employment and squared kilometers of the local labor market (LLMA).

	(1)	(2)	(3)	(4)	(5)	(6)
expenditure in RD	4.6216* (2.2460)	4.3897 (2.2638)	4.4417 (2.3592)	4.7936* (2.4167)	4.7902* (2.4084)	4.6515 (2.4080)
networks	0.4950** (0.1267)	0.4565** (0.1245)	0.5506** (0.1289)	0.5742** (0.1316)	0.5732** (0.1313)	0.5725** (0.1304)
% with upper Sec. or more	0.6647** (0.2149)	0.6180** (0.2076)	0.6482** (0.2183)	0.6819** (0.2226)	0.6813** (0.2227)	0.6850** (0.2226)
incentives	0.3642** (0.0712)	0.3544** (0.0669)	0.3717** (0.0707)	0.3618** (0.0723)	0.3612** (0.0718)	0.3660** (0.0720)
established after 1990	0.1593 (0.1058)	0.1461 (0.1041)	0.1492 (0.1090)	0.2028* (0.1023)	0.2028* (0.1023)	0.2086* (0.1016)
local competition	0.0131 (0.1114)	0.0289 (0.1091)	0.0357 (0.1176)	0.0288 (0.1173)	0.0273 (0.1163)	0.0265 (0.1166)
percentage blue collars	-1.1595** (0.310)	-1.2397** (0.3069)	-1.2560** (0.3190)	-1.2625** (0.3250)	-1.2597** (0.3243)	-1.2319** (0.3254)
size 21-50	0.0308 (0.0876)	0.0503 (0.0927)	0.0747 (0.0862)	0.0714 (0.0857)	0.0702 (0.0859)	0.0717 (0.0859)
size 51-250	-0.0140 (0.0959)	-0.0015 (0.0948)	0.0102 (0.0941)	0.0323 (0.0925)	0.0312 (0.0919)	0.0308 (0.0913)
size 251-500	0.2440 (0.2044)	0.2865 (0.1928)	0.2655 (0.2057)	0.1779 (0.2042)	0.1772 (0.2044)	0.1786 (0.2013)
log density	-0.1661* (0.0640)	-0.1535* (0.0650)	-0.1967* (0.0789)	-0.1081 (0.0672)	-0.1032 (0.0703)	-0.4733* (0.1990)
local rate of unemployment		-2.0341 (1.8381)	-3.1728 (1.7104)	-4.2169** (1.4557)	-4.2298** (1.4639)	-4.2987** (1.4378)
% graduates in the area			4.6062 (3.7129)	4.4564 (3.5187)	4.4653 (3.5334)	4.9779 (3.3106)
average age			-0.0052 (0.0110)	-0.0060 (0.0109)	-0.0054 (0.0112)	-0.0045 (0.0115)
percentage males			0.8206** (0.3170)	0.9180** (0.2884)	0.9187** (0.2906)	0.8213** (0.2451)
specialization index				1.1536* (0.5296)	1.1325* (0.5588)	0.7254 (0.6251)
industrial district					0.0151 (0.0590)	0.0046 (0.0574)
log density *specialization						2.9709** (1.3949)
# obs :	3307	3507	3146	3059	3059	3059

Note: see Table 2. The selection equation is not reported.

Appendix

1. Second derivatives

In order to show that the curve A is concave, we calculate the second derivative of the implicit function

$$A(\tau, \theta, \sigma) = y(\tau) + \kappa\tau\sigma = [2(r+q) + f(\theta)] \frac{d\theta}{f(\theta)}$$

$$\frac{\partial^2 \tau}{\partial \theta^2} = \frac{2(r+q)d[-\theta f''(\theta)]f(\theta)^2 - 2(r+q)d[f(\theta) - \theta f'(\theta)]2f(\theta)f'(\theta)}{f(\theta)^4 A_\tau}$$

taking into account that $f(\theta) = m\theta^\alpha$

$$\begin{aligned} \frac{\partial^2 \tau}{\partial \theta^2} &= \frac{2(r+q)d[-\theta m\alpha(\alpha-1)\theta^{\alpha-2}]m^2\theta^{2\alpha} - 2(r+q)d[m\theta^\alpha(1-\alpha)]2\alpha m^2\theta^{2\alpha-1}}{(m\theta^\alpha)^4 A_\tau} \\ &= \frac{2(r+q)d[m^3\alpha(1-\alpha)\theta^{3\alpha-1}] - 2(r+q)d[m^3 2\alpha(1-\alpha)\theta^{3\alpha-1}]}{(m\theta^\alpha)^4 A_\tau} \\ &= \frac{2(r+q)dm^3\alpha(1-\alpha)\theta^{3\alpha-1}[1-2]}{(m\theta^\alpha)^4 A_\tau} < 0 \end{aligned}$$

2. Comparative statics

We show that the comparative statics in Section 2.3 hold also for Section 2.4. Total differentiation of equations (18) and (19) yields

$$A_\tau \partial \tau + A_\theta \partial \theta = -\tau \psi \partial \sigma$$

$$B_\tau \partial \tau + B_\theta \partial \theta = -\psi \partial \sigma$$

where:

$$A_\tau = y'(\tau) + \psi\gamma(1-u) > 0$$

$$B_\tau = y''(\tau) - [2(r+q) + f(\theta)]c'(\tau) < 0$$

$$A_\theta = -d - \frac{2(r+q)d[f(\theta) - \theta f'(\theta)]}{[f(\theta)]^2} + \frac{\psi\gamma q f'(\theta)}{[q + f(\theta)]^2}$$

$$B_\theta = -c'(\tau)f'(\theta) + \frac{\psi\lambda f'(\theta)}{[q + f(\theta)]^2}$$

Using the Cramer's rule:

$$\frac{\partial \tau}{\partial \sigma} = \frac{\psi(A_\theta - B_\theta \tau)}{A_\tau B_\theta - B_\tau A_\theta}$$

$$\frac{\partial \theta}{\partial \sigma} = \frac{\psi(B_\tau \tau - B_\theta \tau)}{A_\tau B_\theta - B_\tau A_\theta}$$

The sign of the determinant $\Delta = A_\tau B_\theta - B_\tau A_\theta$ is unclear but can be signed if we impose that the equilibrium is locally stable. A first order Taylor approximation of (18) and (19) yield

$$B(\tau, \theta) \approx B_\tau(\tau - \tau^*) + B_\theta(\theta - \theta^*)$$

$$A(\tau, \theta) \approx A_\tau(\tau - \tau^*) + A_\theta(\theta - \theta^*)$$

The Jacobian of this system, J, evaluated at the equilibrium point τ^*, θ^* is

$$\begin{bmatrix} B_\tau & B_\theta \\ A_\tau & A_\theta \end{bmatrix}$$

The pair τ^*, θ^* is a stable fixed point if $\text{tr } J < 0 < \det J$. Therefore, stability requires that $B_\tau A_\theta - B_\theta A_\tau > 0$, which can be written as $B_\theta A_\tau - B_\tau A_\theta < 0$. 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 $\gamma(1-u)$ varies. Therefore we are interested in the sign of $\frac{\partial \tau}{\partial \gamma(1-u)}$ and $\frac{\partial \theta}{\partial \gamma(1-u)}$. In the former case this derivative is equal to

$$\frac{\partial \tau}{\partial \gamma} \left[\frac{\partial \gamma(1-u)}{\partial \gamma} \right]^{-1} = \frac{\partial \tau}{\partial \gamma} \left[(1-u) - \eta u \frac{\partial u}{\partial \theta} \frac{\partial \theta}{\partial \gamma} \right]$$

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

same argument can be applied to $\frac{\partial \theta}{\partial \gamma(1-u)}$.

3. Tables

Table A1. Training regression. Baseline model for firms with 11-250 employees. Dependent variable: $\ln(T/(1-T))$. With Heckman correction for selectivity.

	(1)	(2)	(3)	(4)	(5)	(6)
expenditure in R&D	4.4352 (2.6614)	4.3230 (2.7057)	4.2107 (2.8664)	4.3869 (2.8936)	4.3721 (2.8811)	4.3383 (2.8584)
networks	0.5021** (0.1220)	0.4989** (0.1225)	0.5680** (0.1250)	0.5846** (0.1276)	0.5836** (0.1273)	0.5821** (0.1272)
% with upper sec. or more	0.6097** (0.2167)	0.6058** (0.2157)	0.6602** (0.2211)	0.6579** (0.2279)	0.6560** (0.2290)	0.6597** (0.2282)
incentives	0.4320** (0.0715)	0.4311** (0.0714)	0.4525** (0.0757)	0.4203** (0.0756)	0.4192** (0.0749)	0.4189** (0.0748)
established after 1990	0.1284 (0.1061)	0.1304 (0.1066)	0.1422 (0.1134)	0.1855 (0.1085)	0.1859 (0.1081)	0.1864 (0.1073)
local competition	0.0495 (0.1097)	0.0470 (0.110)	0.0186 (0.1156)	0.0059 (0.1158)	0.0042 (0.1151)	0.0013 (0.1147)
% blue collars	-1.1785** (0.3011)	-1.1864** (0.3020)	-1.1238** (0.3132)	-1.1710** (0.3179)	-1.1715** (0.3184)	-1.1601** (0.3217)
size 21-50	0.0436 (0.0944)	0.0424 (0.0941)	0.0835 (0.0869)	0.0737 (0.0855)	0.0726 (0.0860)	0.0701 (0.0854)
size 51-250	-0.0190 (0.0998)	-0.0206 (0.0993)	-0.0155 (0.0987)	0.0016 (0.0941)	0.0009 (0.0941)	-0.0027 (0.0944)
log density	-0.0945 (0.0501)	-0.0812 (0.0475)	-0.1145 (0.0597)	-0.1170 (0.0601)	-0.1183* (0.0588)	-0.2317 (0.1351)
local rate of unemployment		-1.5759 (1.7382)	-1.9511 (1.3886)	-2.7994* (1.3323)	-2.7361* (1.3375)	-2.7229* (1.3518)
% graduates in the area			5.6257* (2.5446)	6.4012* (2.5697)	6.6381** (2.5008)	7.1343** (2.7570)
average age			-0.0133 (0.0119)	-0.0148 (0.0124)	-0.0144 (0.0126)	-0.0127 (0.0130)
% males			0.7994* (0.3233)	0.8305** (0.3043)	0.8297** (0.3067)	0.8058** (0.2853)
specialization index				1.0616 (0.6216)	1.0255 (0.6590)	4.3147 (4.0489)
industrial district					0.0183 (0.0596)	0.0186 (0.0602)
log density * specialization						0.9109 (1.1022)
# obs :	3389	3389	3293	3262	3262	3262

Note: robust and cluster adjusted standard errors in parentheses with $p < 0.05 = *$, $p < 0.01 = **$. A constant plus industry and regional dummies are included in both equations

Table A2. Training regression. Baseline model for all firms. Dependent variable: $\ln(T/(1-T))$. With Heckman correction for selectivity

	(1)	(2)	(3)	(4)	(5)	(6)
expenditure in R&D	4.6849* (2.2561)	4.5586* (2.2915)	4.4492 (2.4566)	4.5163 (2.4883)	4.5138 (2.4828)	4.4723 (2.4569)
networks	0.4528** (0.1256)	0.4485** (0.1254)	0.5099** (0.1320)	0.5395** (0.1336)	0.5392** (0.1338)	0.5381** (0.1330)
% with upper sec. or more	0.6240** (0.2122)	0.6220** (0.2101)	0.6662** (0.2194)	0.7049** (0.2254)	0.7044** (0.2251)	0.7159** (0.2238)
incentives	0.3632** (0.0717)	0.3623** (0.0711)	0.3847** (0.0763)	0.3677** (0.0782)	0.3673** (0.0773)	0.3659** (0.0770)
established after 1990	0.1758 (0.1049)	0.1782 (0.1050)	0.1914 (0.1121)	0.2335* (0.1066)	0.2335* (0.1066)	0.2351* (0.1058)
local competitors	0.0522 (0.1133)	0.0477 (0.1131)	0.0234 (0.1192)	0.0257 (0.1198)	0.0253 (0.1192)	0.0204 (0.1188)
% blue collars	-1.0774** (0.3005)	-1.0813** (0.2983)	-1.0291** (0.3116)	-1.0666** (0.3188)	-1.0665** (0.3187)	-1.0423** (0.3189)
size 21-50	0.0604 (0.0932)	0.0588 (0.0927)	0.0950 (0.0836)	0.0955 (0.0838)	0.0950 (0.0841)	0.0910 (0.0831)
size 51-250	0.0128 (0.0959)	0.0110 (0.0950)	0.0112 (0.0950)	0.0202 (0.0942)	0.0199 (0.0940)	0.0138 (0.0940)
size 251-500	0.3342 (0.1849)	0.3348 (0.1834)	0.4174~ (0.1945)	0.3425 (0.1982)	0.3421 (0.1988)	0.3320 (0.1964)
size >500	0.1521 (0.3057)	0.1597 (0.3059)	0.1730 (0.3112)	0.1216 (0.3221)	0.1217 (0.3222)	0.1019 (0.3087)
Log density	-0.1240** (0.0413)	-0.1089** (0.0377)	-0.1586** (0.0513)	-0.1113* (0.0522)	-0.1117* (0.0517)	-0.3123* (0.1574)
local rate of unemployment		-1.9689 (1.6936)	-2.7982 (1.4426)	-3.5198** (1.3157)	-3.4999** (1.3231)	-3.5609** (1.2966)
% graduates in the area			6.6267** (2.5364)	7.5126** (2.5659)	7.5889** (2.5926)	8.3977** (2.8542)
average age			-0.0058 (0.0107)	-0.0085 (0.0112)	-0.0084 (0.0113)	-0.0059 (0.0115)
% males			0.5859 (0.3276)	0.6905* (0.2919)	0.6894* (0.2928)	0.6401* (0.2722)
specialization index				1.3783* (0.6637)	1.3651* (0.6831)	7.1813 (4.4563)
industrial district					0.0060 (0.0577)	0.0061 (0.0587)
log density specialization						1.6123 (1.220)
*# obs :	3552	3552	3448	3411	3411	3411

Note: robust and cluster adjusted standard errors in parentheses with $p < 0.05 = *$, $p < 0.01 = **$. A constant plus industry and regional dummies are included in both equations.