DOCUMENTO DE TRABAJO

AN ECONOMETRIC ANALYSIS OF EDUCATION EXTERNALITIES IN THE MATCHING PROCESS OF UK REGIONS (1992-99)

Documento de Trabajo n.º 0403

Pablo Burriel-Llombart

BANCO DE ESPAÑA

SERVICIO DE ESTUDIOS

AN ECONOMETRIC ANALYSIS OF EDUCATION EXTERNALITIES IN THE MATCHING PROCESS OF UK REGIONS (1992-99) (*)

Documento de Trabajo nº 0403

Pablo Burriel-Llombart
BANCO DE ESPAÑA

(*) I am thankful to my supervisor Professor Christopher Pissarides for his comments and help. I would also like to thank Richard Barwell, Olimpia Bover, Paloma Lopez, Francois Ortalo-Magne, Barbara Petrongolo, Steve Pischke, Merxe Tudela, Jonathan Wadsworth and an anonymous referee. All remaining errors are mine. This work is part of my PhD Thesis at the London School of Economics and was carried out while I was a member of the CEP. The views expressed here are those of the author and do not necessarily reflect those of the Banco de España.

ISSN: 0213-2710 (print)
ISSN: 1579-8666 (online)
Depósito legal: M. 6897-2004
Imprenta del Banco de España

Abstract

This paper studies the existence and the scale of education externalities in the unemployment durations suffered by workers in the UK. First, we develop a theoretical model. Using a matching framework we show that a rise in the average level of education of a labour market will affect unemployment durations in two different ways. It will increase the firms' expected profits per vacancy opened, since firms expect to be matched with a more qualified worker, rising job creation and reducing unemployment durations. We call this the Composition or External effect. But, since more qualified workers are more efficient in the process of job search, it will also rise the competition amongst workers for opened vacancies, increasing unemployment durations. We name this the Competition effect. In the most skilled segments of the labour market the composition effect will dominate the competition effect, while in the least skilled segments the opposite will be true. Then, we test these theoretical results empirically using data from the UK Labour Force Survey for the 17 UK regions over the period 1992Q1-99Q4. We find that a 1% rise in the average level of education reduces unemployment durations of individuals from skilled occupations by 2.9% on average, while it rises the unemployment durations of individuals from unskilled occupations by 1.9% on average. This effect is robust to different measures of education, to controlling for unobserved heterogeneity and to different parameterisations of the hazard function.

1 Introduction

Education externalities have been at the heart of the economic and policy debate for the last two decades. Different theoretical explanations have been developed, and these can be grouped into two main categories: technological externalities (or non-pecuniary) and pecuniary externalities. Both these types of externalities were already mentioned by Alfred Marshall (Marshall (1920)) as reasons explaining the concentration of economic activity. However, they were not further developed until more recently.

The first type of externality was re-discovered by the works of Romer (1986) and Lucas (1988). They showed theoretically that, in an area with a higher average level of education processes like the exchange of ideas, imitation or learning by doing are more likely to occur, in turn fostering technological progress. These type of spill-overs have been explored in great detail by the endogenous growth literature. The second type of externality was re-discovered by the works of Krugman (1991a) and further developed by the new economic geography literature. However, these ideas have also been used in other areas of economic research. Acemoglu (1996) showed that, in a labour market were it is costly for firms and workers to find each other, if the average level of education of workers is high then firms will invest more in physical capital. This generates a pecuniary education externality which does not work through technology, but through improving the search process. A similar type of externality is developed in this paper. Other authors have used these ideas to stress instead the interactions between education attainment and the location of economic activity. Rotemberg and Saloner (2000) study an externality coming from the interaction between regional agglomeration of production and the incentive for workers to invest in industry-specific skills. Amiti and Pissarides (2002) show that the existence of a bigger pool of qualified workers improves the quality of matches and helps agglomeration of activity.

The existence of education externalities means that economically identical workers will tend to earn higher wages and enjoy greater employment rates in human capital rich areas than in human capital poor areas. That is, the social return to education may exceed the private return and, therefore, individuals will tend invest in education below what is socially optimal. This has provided one of the main justifications for the public provision of formal education. Primary and secondary education, as well as an important part of university education, is almost completely subsidized by the state in most countries around the world. Therefore, the magnitude of the social return to education is crucial for assessing the efficiency of public investment in education. In addition, many regional policies of development are justified on the existence of agglomeration externalities often related to education externalities.

However, despite the significant policy implications and the theoretical developments, there had not been any relevant empirical work on this area until fairly recently. Moreover, most of the existing work has concentrated on the estimation of the effect of average education in an area on individual wages, that is, on estimating the social returns to education. Rauch (1993) was the first attempt to estimate human capital externalities. He used data from the United States' 1980 Census to test the effect of average education in the Standard Metropolitan Statistical Areas on individual wages. He found that a one-year increase in average schooling in an area raises individual wages by between 3-5%. Moretti (1999) re-estimates this effect for US cities using instruments for the average level of education to avoid an omitted variable bias problem. He finds that a 1 percentage point increase in college share in a city raises average wages by 1.2%-1.4% above the private returns to education. Other studies have found little evidence of significant external returns. Acemoglu and Angrist (2000) estimated the effect of average schooling in US states on individual wages, using the change in State compulsory attendance laws and child labour laws as instruments. They found modest and statistically insignificant external returns. Ciccone and Peri (2000) used a standard neoclassical growth model to identify external effects and found that these were negligible and insignificant. Instead, they found substantial scale effects.

Another line of empirical research has looked at the effect of education externalities on employment growth in cities. Simon (1998) found that a rise in the supply of high school and degree graduates in a US Standard Metropolitan Area in 1940 increases employment growth in the area and that this effect is persistent, lasting up to 40 years. Simon and Nardelli (1996) looking at English cities found that this effect could last up to a century. Glaeser et al (1992) studied the effect of knowledge spill-overs on employment growth in industries within cities. They found that these effects are more likely to occur between industries within a city and when competition between firms in an industry is strong.

In this paper we take a different approach - we test the existence of education externalities in the matching process. Firstly, a model showing the existence of a pecuniary education externality in the matching process is developed. In this model, human capital externalities arise when there are matching frictions in the labour market because firms have to decide whether to create a job or not before knowing who they will finally employ. Thus, a more educated labour force will increase the expected profits per vacancy opened and increase job creation in that area, at the same time increasing the worker's probability of finding a job.

We call this external effect. But, a more educated labor force will also be more efficient searching for jobs. This will increase the competition amongst workers for opened vacancies and reduce the average probability of finding a job. The net effect on the probability of finding a job will depend on the relative importance of these two forces. Then, we hypothesize that for the more skilled segments of the labour market, where the worker's qualifications are a fundamental determinant of the firm's expected profits, the external effect dominates the composition effect. Meanwhile, in the unskilled segments of the labour market, firms benefit only from workers achieving a certain minimum level of education. Therefore, job creation reacts less to a more qualified workforce and the competition effect dominates the external effect. This means that if externalities exist in the matching process, one should find that unemployed individuals belonging to the skilled (unskilled) segments of a labour market where the labour force is better educated have a higher (lower) probability of moving from unemployment into employment than otherwise similar unemployed individuals in labour markets where the labour force is less educated. In addition, for a given average, the more equal (unequal) the distribution of education the higher the probability of transition.

We test this theory by a maximum-likelihood estimation of a model of the duration of unemployment which assumes a discrete-time semi-parametric hazard function and allows the covariates to vary within each unemployment spell. The estimation shows that the average level of education affects positively the probability of transition from unemployment to employment for individuals from skilled occupations, while the variance of education affects it negatively. The opposite is found for individuals from unskilled occupations. Both of these effects are statistically significant at the 10% level. The magnitude of this effect is in line with the findings in the literature. A 1% increase in the average education of the labour market rises (reduces) on average the probability of employment of individuals from skilled (unskilled) occupations by 2.9% (-1.9%).

Finally, an important issue in this literature is whether these effects are due to education externalities or to complementarities in skills. Moretti (1999) indicated that the fact that average education affects wages does not necessarily imply the existence of education externalities. This result could be due to complementarities between high and low educated workers. However, he argued that if different skills are perfect substitutes, the effect of an increase in the supply of educated workers on their own wage had to be an external effect. By doing this he found that a 1 percentage point increase in the labour force share of college graduates increases wages of college graduates by 1.2%, and therefore concluded that education externalities are important in US cities. Ciccone and Peri (2000) argued instead that if skills are imperfect substitutes, one cannot separately identify the external effect from the effect of the complementarity of skills using a regression of individual wages on average wages. They used a standard neoclassical growth model to identify external effects and found that these were negligible and insignificant. But, their methodology is completely dependent on the theoretical model used. It is difficult to apply directly their work to the model of this paper since they implicitly assume a competitive labour market with no unemployment. However, what their work shows is that there is a significant effect of an area's skill composition on wages paid, but this is not due to an externality that works through improvements in productivity. Then, using their model they conclude that their result has to be due to complementarity of skills. Alternatively, one could interpret this result as suggesting that it might be other types of externalities not working through productivity which are important. One such type could be the one studied in this research, which affects wages through improvements in the matching process.

This paper will be organised as follows. In section 2 the theoretical model is developed. The dataset used as well as some descriptive analysis of the distribution of education and its relation with labour market performance is described in section 3. The econometric methodology is explained in section 4. Section 5 outlines the estimation results and some robustness analysis. In section 6 we try to confirm the relevance of the results by estimating directly the effect of education externalities on job creation. Finally, section 7 concludes.

2 Theoretical model

Workers[1]

In the economy there are L individuals, each one born with a different ability (a_i). When young they attend full-time education and then enter the labor market with the human capital obtained. Individual human capital h_i is a function of innate ability and assumed to be given, $h_i = h(a_i)$. In the labor market individuals and firms engage in a search process, which produces a number of matches. Individuals search with different intensity s_i depending on their characteristics, mainly the level of education. This can be expressed as:[2]

$$s_i = s(h_i) = h_i^{\delta} \tag{1}$$

[1] This theoretical model is based on Acemoglu (1996) and Burriel-Llombart (2002).

[2] The individual's decisions of education and search efficiency are assumed exogenous to simplify the model, but can be easily endogenized. See Burriel-Llombart (2002).

The more intensively an individual searches (the more units of search he supplies), the higher his probability of finding a job. However, in equilibrium, the individual probability of employment also depends on how intensively the other individuals in that labour market search, that is, on the aggregate supply of units of search S.

$$S = \int s_i f(i)di = \int h_i^\delta f(i)di = LE(h_i^\delta) \tag{2}$$

Those individuals who find a job, produce and earn a wage, while the rest remain unemployed and earn a subsidy.

Notice that we can define $\frac{\int h(i)^\delta f(i)di}{L}$ as $E[h^\delta]$, which from now on will be called "average" education. However, this term depends on the whole distribution.

Firms

The number of firms active in the economy is variable. When a firm decides to enter the labor market opens a vacancy and starts looking for a worker. The cost of opening the vacancy is sunk, so the firm will only open one when expected profits are non-negative. Once a firm and a worker meet, the firm buys the appropriate technology for the worker's human capital level and the worker brings one unit of labor and his human capital. The result of the match is the production of y_i units of product using the following technology:

$$y_i = \begin{cases} Ah_i^\alpha & if \quad h_i < \overline{h} \\ A\overline{h}^\alpha & otherwise \end{cases} \tag{3}$$

where $A > 0$ is a constant representing the technological level and $1 > \alpha > \delta > 0$[3].[4] According to this technology the worker's human capital contributes to production up to a certain level, $\overline{h}$. Any human capital above this level does not improve the worker's productivity and therefore does not benefit the firm. This upper-bound will be higher the more skilled the occupation or segment of the labour market in which the firm operates. This assumption precludes the counterfactual result that a more qualified worker will always be more productive independently of the type of job for which he is employed. That is, for example the productivity of an operator will improve significantly after obtaining the compulsory level of education or a vocational qualification but it will not increase any further after obtaining

[3] This is found to be always true when the individual decisions are endogeneized. It only requires the cost of search to be concave.

[4] Physical capital is not introduced in this version of the model to simplify and emphasize the main mechanism studied.

a degree.

When a match is realized, the occupied job will yield a return that is at least as high as the sum of the expected returns of a searching firm and a searching worker. Wages are set to share this economic rent according to the Nash Solution to a bargaining problem, as in Pissarides (2000). This means that the worker receives a share β of output, while the firm receives the rest of the output.[5]

$$\begin{aligned} w_i &= \beta y_i \\ \pi_i &= (1-\beta) y_i \end{aligned} \tag{4}$$

Labor Market Equilibrium

The labor market is composed of L workers supplying S units of search and V vacancies who engage in a search process by which N matches are created. Since each worker searches for a job with different search intensity, s_i, - i.e. each worker supplies a different number of units of search - the number of matches depends on the aggregate supply of search, S, instead of the number of people searching L, and on the number of vacancies:

$$N = m(V, S)$$

where $m(.,.)$ represents a matching function with standard properties.[6] The level of employment (N) should also be equal to the number of units of search supplied by individuals looking for work times the probability that a worker meets a firm (q), $N = qS$. This implies that on average,

$$q = \frac{N}{S} = m(\frac{V}{S}, 1) = m(\theta, 1). \tag{5}$$

where $\theta = \frac{V}{S}$ is the labour market tightness. It is assumed that q is also the average transition probability. Since in this model each worker searches with different intensity for jobs, each worker will face a different probability of employment. In particular, the individual probability of employment will be equal to the number of search units supplied by that individual times the average probability of employment.

$$q_i = s_i q \tag{6}$$

The average probability of filling a vacancy (p) has to be equal to the level of employment,

[5] The parameter β might be interpreted as the worker's relative bargaining power.

[6] Increasing in both arguments, concave and homogeneous of degree one.

N, over the number of vacancies opened, V.

$$p = \frac{N}{V} = m(1, \frac{\theta}{1}) = \frac{q}{\theta} \tag{7}$$

The expected profit of a firm from a vacancy ($E(\pi)$) will be equal to the probability of filling a vacancy (p) with a worker, times the profit obtained from employing that worker. Since the firm does not know which worker will arrive we have to integrate over all possible individuals.

$$E(\pi) = p\frac{\int [y(i) - w(i)]\, f(i)di}{S}$$

There is a fixed cost of opening a job equal to k, which is independent of the type of worker recruited. In equilibrium no firm can open a job and make a positive profit since there are no barriers to entry, therefore $E(\pi) = k$. Substituting the equations determining p, y_i, w_i and S ((equations (7), (4) and (2)) into this equation, and solving for θ we obtain the labor market tightness as a function of the employment rate and the distribution of education:

$$\theta = q\left[\frac{(1-\beta)A}{k}\right]\left[\frac{\int_0^{\bar{i}} h(i)^{\alpha} f(i)di + \overline{h}^{\,\alpha} \int_{\bar{i}}^{\infty} f(i)di}{\int h(i)^{\delta} f(i)di}\right] \tag{9}$$

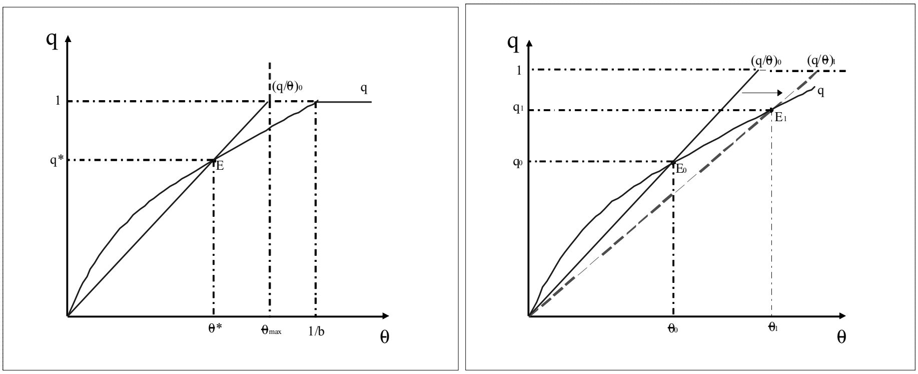
where $\bar{i}$ represents the marginal individual with education equal to $\overline{h}$. According to this expression market tightness depends positively on the employment rate, as well as on the worker's average product per unit of search. This is in turn a function of the distribution of education in that labour market segment. Thus, a rise in the average product per unit of search increases vacancy creation and labour market tightness.

For simplicity, we define $\underline{L} = \int_0^{\bar{i}} f(i)di$ the labour force with education below $\overline{h}$, $\overline{L} = \int_{\bar{i}}^{\infty} f(i)di$ the labour force with education above $\overline{h}$ and $E(\underline{h}^{\alpha}) = \int {}_0^{\bar{i}} h^{\alpha} f(i)di$ is the "average" level of education for the labour force with education below $\overline{h}$. Then, we can re-write the worker's average product per unit of search as follows:

$$\frac{\int_0^{\bar{i}} h(i)^{\alpha} f(i)di + \overline{h}^{\,\alpha} \int_{\bar{i}}^{\infty} f(i)di}{\int h(i)^{\delta} f(i)di} = \frac{\underline{L}\, E(\underline{h}^{\alpha}) + \overline{L}\, \overline{h}^{\,\alpha}}{LE(h_i^{\delta})} \tag{10}$$

Finally, assuming the matching function in the equation determining the average probability of employment (equation (5)) takes a Cobb-Douglas form, $m\,(V, S) = (bV)^{\phi}\,(S)^{1-\phi}$, and substituting *theta* into this equation using equation (9) we obtain the average probability of employment that equilibrates the labour market. This equilibrium is also shown in figure 1a), where market tightness (equation (9)) is represented by the straight line named q/θ and

Figure 1: a) Labour market equilibrium. b) Rise in $E[h^{\alpha}]$.



the average probability of employment (equation (5)) is represented by the curve named q:[7]

$$q = \left(\frac{b(1-\beta)A}{k}\right)^{\frac{\phi}{1-\phi}} \left[\frac{\underline{L}\, E(\underline{h}^{\alpha}) + \overline{L}\,\overline{h}^{\,\alpha}}{LE(h_i^{\delta})}\right]^{\frac{\phi}{1-\phi}}$$

Then, substituting this expression into equation (6) we obtain the individual probability of employment in equilibrium:

$$q_i = h_i^{\delta} \left(\frac{b(1-\beta)A}{k}\right)^{\frac{\phi}{1-\phi}} \left[\frac{\underline{L}\, E(\underline{h}^{\alpha}) + \overline{L}\,\overline{h}^{\,\alpha}}{LE(h_i^{\delta})}\right]^{\frac{\phi}{1-\phi}} \tag{11}$$

This expression is the basis for the empirical estimation of the model developed. It shows that, as one would expect, the individual probability of finding a job depends positively on his level of education, the average product per unit of search, the firm's share of output, the technology and the efficiency of search, while it depends negatively on the cost of opening vacancies. However, a rise in the average level of education has an ambiguous effect on this probability, since it affects both the denominator and numerator of this expression. We will

[7] An alternative equilibrium with full employment (and high market tightness) may prevail when the "average" level of human capital is very high, or the cost of opening a vacancy is very low. This equilibrium will exist when the two equations in figure 1a) cross above the upper bound of full employment. The reason is that, when the "average" human capital in the economy is very large, the firms' expected profits are so big that they will keep on creating new jobs beyond the point in which market tightness is large enough to achieve full employment. This will happen until expected profits disappear and labour market equilibrium is achieved. We will not discuss this alternative equilibrium since the focus of the paper is on the effects of education externalities on the employment rate.

study this issue in the next section.

2.1 Education Externalities: Composition vs. Competition effects:

A rise in the average level of education generates two offsetting effects on the average product per unit of search and thus in the equilibrium values of the average probability of employment and market tightness (see equations (11) and (9)):

Composition effect: A rise in the average level of education rises the expected productivity of workers, $\underline{L}\, E(\underline{h}^{\alpha}) + \overline{L}\, \overline{h}^{\alpha}$, increasing the profit per vacancy opened and therefore rising vacancy creation and the probability of employment. This is what we mean in this paper by the external effect of education.

Competition effect: But a rise in the average level of education also rises the aggregate supply of search units ($E[h^{\delta}]$), increasing the competition amongst workers for existing vacancies (the market is more crowded for workers), reducing market tightness and the average probability of employment per unit of search. This is the congestion effect typical of all matching models.

Depending on which of these two effects dominates we may find three different scenarios:

a) Positive external effect: When all workers have a level of education below $\overline{h}$, the composition effect dominates the competition effect. The external effect becomes $\frac{E(h_i^{\alpha})}{E(h_i^{\delta})}$ and, since $\alpha > \delta$, a rise in the average level of education always has a positive effect on the equilibrium employment rate. This is represented in figure 1b).

b) Negative external effect: When all workers have a level of education above $\overline{h}$, the composition effect disappears and the competition effect is the only one present. Then, a rise in the average level of education will only increase competition amongst workers for the existing vacancies and reduce the average employment rate and market tightness. This is represented by a external effect equal to $\left[\frac{\overline{h}^{\alpha}}{E(h_i^{\delta})}\right]$

c) Indeterminate sign of external effect: Finally, when some workers have education above $\overline{h}$ and some below, it is unclear which effect dominates and what will be the sign of the external effect. In this case the external effect takes the form represented in equation (10).

Which of these scenarios prevails will depend on the labour market segment of analysis. The more skilled the labour market segment, the higher is the upper bound on productivity ($\overline{h}$) and therefore the more likely is a positive external effect. Therefore, one would expect to find a positive external effect for skilled occupations a negative external effect for unskilled occupations and an indeterminate effect for the semiskilled ones.

Human capital externalities arise in this model because firms have to decide whether to create a job or not before knowing who they will finally employ and workers have to decided how many units of search to supply before knowing who they will work for. A more educated labour force will increase the expected profits per vacancy opened and increase job creation. But it will also rise the competition amongst workers for the opened vacancies, which might partially or totally offset the first effect. However, if the labour market was perfectly competitive and there were no matching frictions, firms would be matched with workers until the worker's human capital made the firm's profits equal to zero. In this case, every firm and worker knows who it will be matched with and therefore the job creation decision depends only on the individual's human capital. That is, in the competitive case aggregate employment is determined by the position of the marginal worker in the education distribution. With matching frictions it is determined by the whole distribution of education. This external effect has been named a pecuniary externality since it is generated in the matching process and is independent of the existence of increasing returns in the production function.

Technological Education Externalities:

As was mentioned in the introduction, education externalities may also arise through the exchange of ideas, imitation or learning by doing (Romer (1986) and Lucas (1988)). These external effects have been called technological or non-pecuniary externalities because they are generated in the process of production. They can be captured in this model by allowing the aggregate productivity term A to depend on aggregate human capital in the following way: $A = E[h_i^\eta]$. If we substitute this into equation (11) we have:

$$q_i = h_i^\delta \left(\frac{b(1-\beta)}{k} \right)^{\frac{\phi}{1-\phi}} E[h_i^\eta]^{\frac{\phi}{1-\phi}} \left[\frac{\underline{L}\, E(\underline{h}^\alpha) + \overline{L}\, \overline{h}^{\,\alpha}}{LE(h_i^\delta)} \right]^{\frac{\phi}{1-\phi}} \tag{12}$$

In general, this type of external effect will have a positive (or zero) effect on the employment rate, because it increases the productivity of all workers. That is, $\eta \geq 0$. However, in a model with skill-biased technological externalities, like Acemoglu (1998) a rise in the number of skilled people incentivates the development of new technologies that increase the productivity of skilled workers and reduce the productivity of unskilled workers. In this case, the external effect will be positive for skilled workers ($\eta > 0$) and negative for unskilled workers ($\eta < 0$), like in the model developed here. Therefore, as equation (12) shows, pecuniary and non-pecuniary externalities skill-biased externalities cannot be separately identified empirically using this model. Instead, in the empirical estimation we will control for the the level of technology in the segment of the labour market by including the industrial structure of employment by occupation and region (using a 4-industry classification and 4 occupations).

Finally, we will use the specification of equation 12 to test whether the probability of finding a job given that you are unemployed for t periods (or unemployment duration) is positively related with the average level of education in the labour market after controlling for individual education and other individual and local characteristics.

3 The data

The data used in this paper comes from the longitudinal Labour Force Survey (LFS). The LFS is designed to be representative of the total population in GB[8], containing very detailed information on the labour force status of individuals as well as on family and individual characteristics. In addition, we use the non-longitudinal LFS to obtain aggregate variables reflecting the evolution of the British regional labour market over time.

The longitudinal LFS is conducted every quarter on all members of around 60.000 households. One fifth of the sample is renewed quarterly and hence we can observe any individual for a maximum of five quarters. It started in the first quarter of 1992 (march-may) and we use all waves up to the fourth quarter of 1999 (November-January)[9]. This period of nine years covers more than a whole cycle of the British Economy. The sample is constructed using only the unemployment spells taking place during the five quarters each individual is in the sample, to avoid a stock sampling bias problem. That is, spells which start not earlier than the quarter of the first interview. This means that the longest spell will be 14 months. Spells will be measured in months. The resulting sample consists of 15,974 unemployment spells with an average duration of 3.4 months. Out of these durations 40.8% finished with a transition into employment, 15.7% finished with a transition into inactivity

[8] Northern Ireland is excluded from the study since the quarterly LFS was not introduced in this area until the winter of 1994-95.

[9] After 1999 regions are only reported using the new classification of regions (GOR). In addition, the county indicator is also dropped from the LFS at this moment which makes it impossible to construct comparable regions.

and the remaining 43.5% did not conclude before the individual left the sample (see table 12).

Since the aim of this study is the estimation of the effect of education matching externalities on the transitions from unemployment to employment, the most important variables are the ones measuring the distribution of education in the local labour market. We assume that the distribution of education is perfectly described by its mean and variance. The theoretical model predicts that what is important is the specific segment of the labour market the individual is participating in. The problem is how to define the relevant segment of the labour market in the data. In this paper we use occupation groups in each region. Using this definition, we calculate the average level of education as the mean education across all the individuals belonging to the same occupation group in a region at a moment in time. Every individual is attributed the average level of education of his own labour market segment - e.g. a manager is attributed the average level of education of the managers in his area in that quarter. We are only looking at direct spill-over effects, that is, within occupation spill-over. We do similarly for the variance of education.

The education and occupation variables used are explained in table 1.[10] The individual education variable has 9 levels going from low to high education. This classification distinguishes between academic and vocational qualifications. In addition, we have aggregated this variable into 4 education groups, where academic and equivalent vocational qualifications belong to the same category. The Occupation variable follows the 9 Major Occupation Groups defined by the new Standard Occupational Classification (SOC) introduced in 1991. This classification was designed so that the occupational groupings brought together jobs with similar requirements in terms of qualifications, training and experience. The ranking of these nine major categories from 1-9 was meant to reflect the progression of the occupations from those requiring a higher level of qualifications, training and/or experience down to those requiring a lower level of skill or experience. This is particularly relevant for this study since we are using occupation groups to identify segments of the labour market which are fairly homogeneous in terms of the education level (ability) of its workers. We have also constructed two more aggregated occupation variables with 2 and 4 groups, keeping the hierarchical structure of the SOC. The regions considered are based on the Standard Statistical Regions classification, split into metropolitan and non-metropolitan areas whenever possible. This divides GB into 19 regions.[11]

[10] A more detailed definition of the education variable can be found in appendix A, table 10

[11] A detailed list of the regions and the counties included in each region can be found in Appendix B, table 11.

Table 1: Categories of Education and Occupation variables

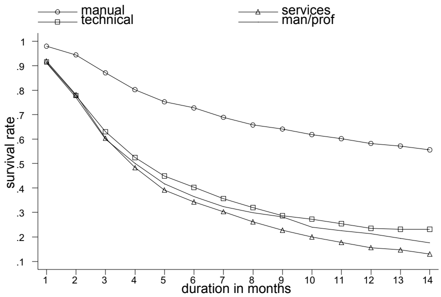
	Education		Occupation		
	9 groups	**4 groups**	**9 groups**	**4 groups**	**2 groups**
9	Degree	Degree or equiv.	Managers	High Skilled	
8	High Voc.	$(8-9)$	Professionals	$(8-9)$	Skilled
7	A Level	A Level or equiv.	Technicians	High Semi-skilled	$(6-9)$
6	Mid. Voc.	$(6-7)$	Craft	$(6-7)$	
5	O Level	O Level or equiv.	Clerical	Low Semi-skilled	
4	Low Voc.	$(4-5)$	Personnel	$(4-5)$	
3	Other Acad.	Other	Sales	Low Skilled	Unskilled
2	Other Voc.	& no qual	Operators	$(1-3)$	$(1-5)$
1	No qual	$(1-3)$	Others		

The other explanatory variables used in this study can be divided into three groups: personal, household and regional characteristics. The personal characteristics include: age, sex, education, last job's occupational group, being white, being married, having migrated in the last year, and whether receiving unemployment benefit or financial help from relatives. The household variables are: region of residence, whether receiving housing benefit, number of dependent children under 6 and between 6 and 16, number of people working in the household and whether it is a one-person-household or a two-person-household. The variables reflecting regional characteristics are obtained from the non-longitudinal LFS and include: average level and variance of education within each occupation group in the region, unemployment rate by region, inactivity rate by region, vacancy rate by occupation and region, ratio of the flow of immigrants to the flow of emigrants and industry's share of employment in the region (10 industries). The migration data is derived from the National Health Service Central Register, provided by NOMIS. Finally, we also include time and region dummies. All the regional variables are included in the estimation in logarithms. The household and regional variables are allowed to vary within each unemployment spell, except for the region of residence[12], while the personal characteristics remain unchanged[13].

[12] The LFS is a survey of non-movers.

[13] Changes in some of the personal characteristics, like age (in years) or education, may occur during an unemployment spell. However, since the maximum spell is 14 months, the effects of these changes are likely to be small.

Figure 2: Survival by 4 occupations



3.1 Descriptive Statistics: Education, Survival rates and Vacancies.

The theoretical model predicts that regions with a relatively higher level of education will have higher job creation and lower unemployment durations. We can use the Kaplan-Meier empirical survival in unemployment to have an initial idea about the differences in unemployment durations across the UK local labour markets during the 1990s. The empirical survival is the fraction of unemployment spells ongoing at the start of a month which do not end during that month[14]. It represents the probability of remaining unemployed given that you have been unemployed for x months.

In general, the most qualified occupations have the lowest probability of remaining in unemployment for all durations (figure 2). An exception are the Service occupations which have the lowest survival for durations longer than two months. If we now look across regions, we observe that the empirical survival in unemployment is lower in the region with the highest education level (Scotland) than in the one with the lowest level (Metropolitan West Midlands) (figure 3).[15] We can also look at different regions by occupation. In figure 4 we compare the empirical survival in unemployment state of the top and bottom regions in terms of qualifications by four occupation groups. It shows that regions with a more

[14]The empirical survivor for month t is equal to the number of spells which do not end during month t, divided by the size of the risk set at the beginning of month t. The size of the risk set at the beginning of month t is the number of people whose spells have not ended or been censored at the beginning of month t.

[15]This is not true for all regions. For example, Greater London has one of the highest survival rates for all durations of unemployment although it has one of the most qualified workforce, while Rest of Northern region has a medium survival rate but one of the lowest levels of qualifications.

Figure 3: Survival of top & bottom regions

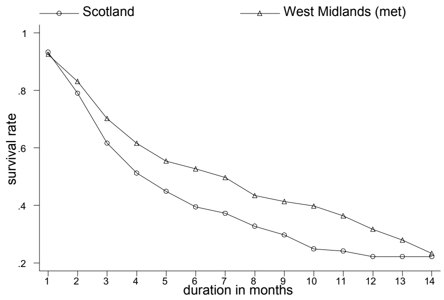


Figure 4: Survival in state of unemployment by occupation for top & bottom regions

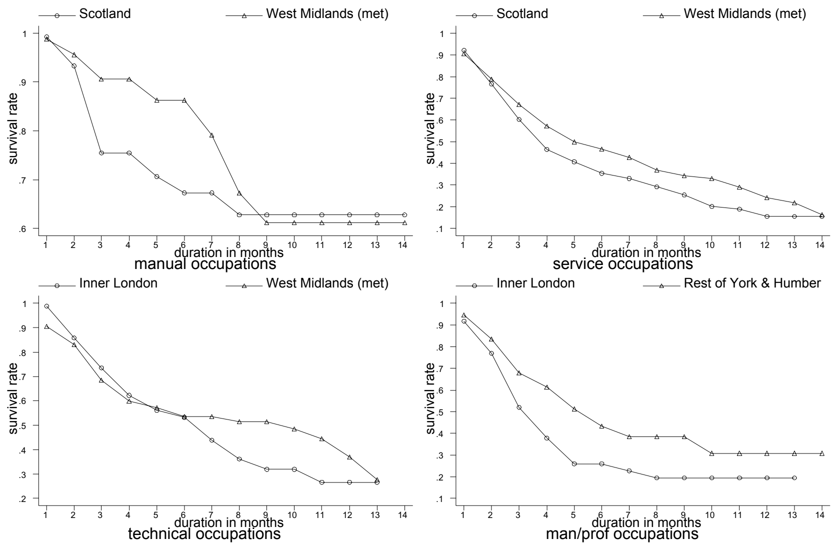


Figure 5: Market tightness by occupation

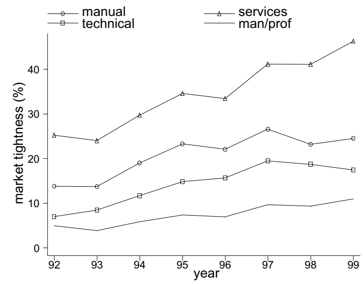
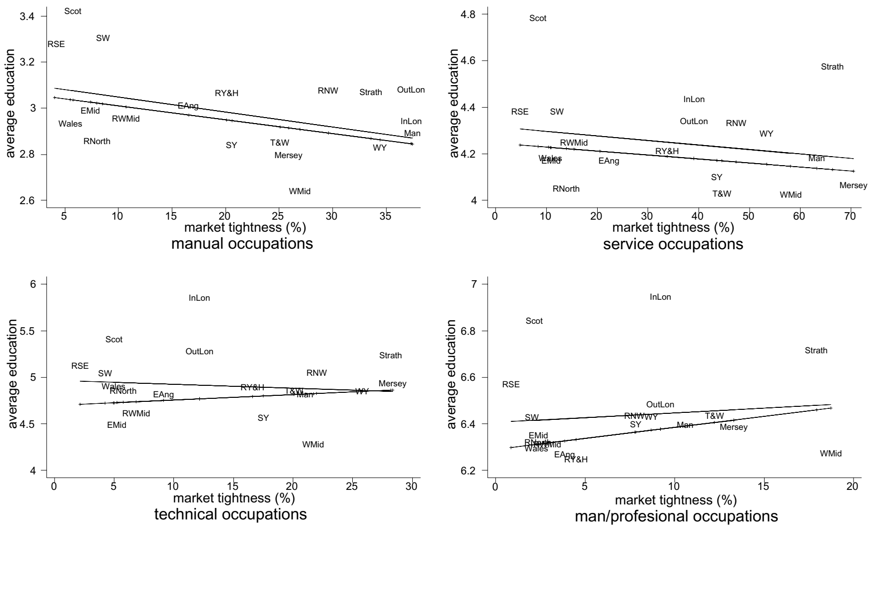


Figure 6: Average education vs labour market tightness by occupation (line with crosses excludes 3 top & bottom regions)



qualified workforce enjoy lower probabilities of remaining unemployed.[16]

Another important variable determining the performance of a labour market is labour market tightness, measured as the ratio of opened vacancies to the number of unemployed. Contrary to what was expected, figure 5 suggests that the most skilled occupations have the lowest market tightness. However, this is not longer true when we look at the relation between education and labour market tightness by occupational group and region. In figure 6, we see that there is a positive correlation between these variables for the more skilled occupations, but a negative one for the least skilled ones. This result is strongest when we drop the top and bottom regions in terms of education [17].

In conclusion, all the descriptive evidence points towards a positive relationship between regional education by occupation and labour market performance, confirming the conclusions of the theoretical model. Of course, it could be possible that the regions with the most qualified workforce have labour markets performing better simply because more qualified people face lower unemployment rates and shorter unemployment durations and not because of an external effect. That is why we now move to test this hypothesis using econometric techniques which allow us to control for the individual and regional characteristics which could be driving this result.

4 Econometric specification.

In order to study the determinants of the transitions from unemployment, we apply econometric duration models to the duration of unemployment spells. The time to exit of unemployment can be thought of as a random variable, T. This variable represents the duration of stay in the state of unemployment. The probability distribution of the duration of unemployment - the probability that the random variable T is less than some value t - can be specified by the distribution function $F(t) = \Pr(T < t)$. Two other functions which are particularly relevant in studying duration data are the Hazard and the Survivor functions. The survivor function, $S(t)$, represents the probability of remaining in a specific state, i.e., the probability that the random variable T will equal or exceed the value t. It can be defined as $S(t) = 1 - F(t) = \Pr(T \geq t)$. The hazard function, $\lambda(t)$, is the rate at which unemployment spells will be completed at duration t, given that they last until t. It is defined as

[16] A similar picture is obtained by looking at unemployment durations. The most skilled occupations have the lowest durations, while regional education by occupation is negatively correlated with unemployment duration.

[17] Scotland, South East, metropolitan West Midlands and rest of the Northern region

$\lambda(t) = f(t)/S(t)$. Both of these functions depend on a vector of explanatory variables $x(t)$ and some unknown coefficients β and λ_0.

We consider a proportional hazard model (Cox (1972)). This model separates the hazard function, into two parts. The first part, $\lambda_0(t)$, is called the "baseline" hazard and represents a functional form for the dependence of the hazard on duration t. The second part, $\phi[.]$, describes the way in which the hazard shifts between individuals endowed with different $x(t)$ at a given duration t. The effect of explanatory variables is to multiply the hazard λ_0 by a factor ϕ which does not depend on duration t. A convenient specification is $\phi[x(t), \beta] = \exp[x(t)'\beta]$ since it ensures the non-negativity of $\lambda[.]$ without constraining the parameter space for β. In addition, with this specification we can interpret the coefficient β as the proportional effect of $x(t)$ on the conditional probability of completing a spell.

In our dataset we only observe the transitions out of unemployment on a monthly basis, so we have to use a discrete-time hazard function, $h_i(t)$. This function denotes the conditional probability that an unemployment spell lasts until time $t+1$, given that it has lasted until t. We will use a complementary log-log specification, which has been shown to be the discrete-time counterpart of an underlying continuous-time proportional hazard model (Prentice and Gloeckler (1978)).

$$\begin{aligned} h_i(t) &= \Pr[T_i = t+1 \mid T_i \geq t, x_i(t)] \\ &= 1 - \exp\left\{-\int_t^{t+1} \lambda_i(s)ds\right\} \\ &= 1 - \exp\left\{-\exp[x_i(t)'\beta] \cdot \int_t^{t+1} \lambda_0(s)ds\right\} \end{aligned} \tag{13}$$

given that $x_i(t)$ is constant between t and $t+1$. Equation (13) can be rewritten as

$$h_i(t) = 1 - \exp\left\{-\exp[x_i(t)'\beta + \gamma(t)]\right\} \tag{14}$$

where

$$\gamma(t) = \int_t^{t+1} \lambda_0(s)ds$$

is called the integrated "baseline" hazard.

Initially, we will not assume a specific functional form for $\gamma(t)$ and estimate the model semiparametrically. Then, we check the robustness of our results, by estimating the model parametrically assuming $\gamma(t)$ takes a Weibull form (see Kiefer (1988)) like $\gamma(t) = \alpha_0 t^{\alpha_1}$.

The contribution to the log-likelihood of the ith individual with a spell of length t_i is given by

$$\ln L_i = c_i \ln f(t_i) + (1 - c_i) \ln S(t_i)$$

where right-censored observations, $c_i = 0$, contribute to the likelihood only with the survivor function since in that case all we know is that the spell of unemployment has lasted until moment t_i. Substituting the definition of the discrete-time hazard function (equation 14) and the survivor function, we get the likelihood function that will be estimated

$$\begin{aligned} \ln L_i &= c_i \{\ln [1 - \exp \{-\exp[x_i(t_i)'\beta + \gamma(t_i)]\}] - \exp[x_i(t_i)'\beta + \gamma(t_i)]\} \\ &\quad - \sum_{t=1}^{t_i} \exp[x_i(t)'\beta + \gamma(t)] \end{aligned} \tag{15}$$

So far, we have wrongly considered that there is only one possible transition out of unemployment. An unemployment spell can terminate when the individual finds a job, but also when he gives up searching and becomes inactive. Given that we are interested in the first type of transition, we need to estimate a competing risk model of duration that distinguishes exit into employment from exit into inactivity. Narendranathan and Stewart (1993) show that the parameters of the hazard into employment can be estimated by treating durations finishing for other reasons as censored at the time of exit. Having done this, the proportional hazard specification used for the single-risk model can be applied to the job-finding hazard.

Using this methodology one can also control for unobserved heterogeneity by conditioning the hazard rate on an individual's unobserved characteristics, summarized in the variable v (Lancaster (1990), chapter 4). This is a random variable taking on positive values, with the mean normalized to one (for identification reasons) and finite variance σ^2. Then, the conditional hazard function (in continuous time) can be re-written as:

$$\lambda[t, x(t), \beta, \lambda_0] = \lambda_0(t) \exp[x(t)'\beta + v_i]$$

with v_i independent of x_i and t. Since each individual v_i is unobserved, we have to specify a distribution for v, so that we can write the unconditional hazard and the survivor function in terms of parameters that can be estimated and of the observable regressors included. This is known as "integrating out" the unobserved effect. In the case of the discrete time proportional hazard model, the Gamma distribution has been the most popular choice in the empirical literature. This takes the form $f(v) \propto v^{\sigma^{-2}-1} \exp(-\sigma^{-2}v)$. The resulting proportional hazard specification identifies three sources of variation among individual hazard rates: the duration of unemployment (t), the observable differences among individuals ($x(t)$) and the unobservable ones (v). In a competing risk framework like this one, we also have to impose the independence of these disturbance terms across the cause specific hazards. Under these assumptions, the log likelihood described in equation 15 becomes the following:

$$\begin{aligned} \ln L_i &= \ln[(1+\sigma^2\sum_{t=1}^{t_{i-1}}\exp\{x_i'\beta+\gamma(t)\})^{-1/\sigma^2} \\ &\quad -c_i(1+\sigma^2\sum_{t=1}^{t_i}\exp\{x_i'\beta+\gamma(t)\})^{-1/\sigma^2}] \end{aligned} \tag{16}$$

Finally, since the variable of interest, the mean and variance of education, varies only across time, regions and occupation groups, when calculating the standard errors we have to allow for correlation of the errors between individuals belonging to the same cluster (see Moulton (1986) for a detailed analysis of this problem for the OLS case).

5 Results

We are now in a position to study the effect of education externalities in the matching process on the conditional probability of leaving unemployment. The theoretical model predicts that education matching externalities will affect positively the hazard of employment of individuals belonging to skilled occupations and negatively the one of individuals from unskilled occupations. The effect on individuals belonging to intermediate occupations will have an indeterminate sign. The results of the estimation are reported in table 2. Each of the specifications is estimated first including all individuals in the sample and then splitting the sample into skilled and unskilled occupations. Finally, in table 3 the sample is divided into 4 groups: high skilled, high semi-skilled, low semi-skilled and low skilled occupations.

The estimated coefficients of the average education and the variance of education have the signs predicted by the theoretical model, confirming the existence of a positive and significant effect of education externalities in the matching process of the skilled segments of the UK local labour markets and a negative and significant effect in the one of the unskilled segments (table 2).

Table 2: Semiparametric estimates by 2 occupations

	no unobserved heterogeneity			unobserved heterogeneity		
Variables	all	skilled	unskilled	all	skilled	unskilled
Baseline estimation (using levels of education)						
Average education	0.705	3.375***	−3.342**	0.682	2.890*	−3.386**
var. of education	−0.409*	−1.306**	1.044*	−0.297	−1.024	1.456**
Using Years of education						
Ave. years of edu.	0.420	2.514***	−1.467*	0.314	2.168**	−1.199*
var. years of edu.	−0.437	−2.184***	0.914	−0.279	−1.948**	0.812**
Using Shares of Education (4 levels)						
Share degree	0.004	0.483	−0.060	0.026	0.280	−0.057
Share A Level	−0.096	0.132	−0.197	−0.076	0.099	−0.193
Share O Level	0.399**	−0.414*	−1.193***	−0.372**	−0.529	−1.189**

NOTES: a) *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.
b) The robust standard errors are clustered by region, occupation and time period.

Table 3: Semiparametric estimates by 4 occupations

	no unobserved heterogeneity				unobserved heterogeneity			
Variables	High	H-Semi	L-Semi	Low	High	H-Semi	L-Semi	Low
Baseline estimation (using levels of education)								
Ave edu	5.321*	2.560	−2.449*	−3.377**	5.142*	2.544	−1.746	−3.658**
Var edu	−1.385	−0.478	1.473*	1.062**	−0.438	−0.366	1.314*	1.183**
Using number of years of education								
Ave edu	2.663*	2.116*	−1.713**	−0.968	2.582*	2.048*	−1.398*	−0.864
Var edu	−3.135	−1.695	1.605**	0.651	−1.347	−1.538	1.343**	0.604
Using shares of education (4 levels)								
Degree	3.003**	0.255	0.238	0.003	2.265*	0.095	0.197	−0.009
A Level	0.569	0.209	−0.580**	0.091	0.136	0.304	−0.480**	0.123
O Level	0.349	−0.388	−0.915	−1.061**	0.306	−0.275	−0.841	−1.238**

NOTES: a) *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.
b) The robust standard errors are clustered by region, occupation and time period.

We estimate first the model using the education variables measured in levels (first row of table 2). When we estimate the model for all the occupations together (column 1) we obtain a positive coefficient for the average education and a negative coefficient for the variance of education, but both of them are not significant at the 10% significance level. However, when we estimate instead the model separately for skilled (column 2) and unskilled occupations (column 3)[18] we realize that this aggregate result is the consequence of two effects of opposed sign, as predicted by the theoretical model. The estimated coefficient for the average education is positive and significant at the 1% level of significance for the skilled occupations, while it is negative and significant at the 5% level for the unskilled occupations. We find a similar picture but with opposite signs for the variance of education. These results are confirmed when we disaggregate further the occupations and estimate the model separately by four occupational groups. The average level of education has the largest positive effect on the hazard rate of the most skilled occupations (managerial & professional occupations), with a level of significance of 10%. This effect is still positive but of smaller size and not significant at the 10% level for the "high semi-skilled" occupations. While it becomes negative and significant, for the "low semi-skilled occupations". Finally, it is more negative and significant at the 5% level for the "unskilled" occupations.

Re-estimating the model controlling for unobserved heterogeneity shows that although heterogeneity is important, the results are qualitatively unchanged (see right hand side panel of table 2). The likelihood ratio test indicates that we cannot reject the model with unobserved heterogeneity and the estimated gamma variance is significantly different from zero (see last part of table 14 in appendix D). The average level of education and the variance of education still have a significant effect when we run the estimations separately for high skilled and low skilled occupations, although of smaller magnitude - a coefficient of 2.9 and -3.4 instead of 3.4 and -3.3 for high and low skilled occupations, respectively. The same is true for the more dissaggregated estimations (see RHS of table 3). This indicates that part of the effect originally attributed to the education externality is actually due to unobserved characteristics. However, the fact that the effect is still statistically significant confirms the predictions of the theoretical model and the existence of education externalities in the matching process taking place in the UK local labour markets.

These results are robust to different to different measures of the distribution of education and to different parameterizations of the hazard function. Firstly, we re-estimate the model

[18] Skilled occupations correspond to occupations 5-9 of the SOC classification, while unskilled occupations correspond to occupations 1-4.

using two alternative measures of the distribution of education: the average and the variance of the number of years in post-compulsory education[19] and the share of the labour force by four education levels. Table 2 (row 2) shows that using the number of years of education provides qualitatively similar results to using an ordinal classification of education. The main difference being that the negative effect on less skilled occupations is less significant. And in particular is insignificant for the least skilled group of the four occupations used. This is due to the fact that this education variable gives a lower weight to the lower education categories which are the predominant ones in the less skilled occupations. The third row of table 2 shows that using the labour force shares by education level we also find a qualitatively similar picture, specially when looking at the four occupations separately (table 3). A larger proportion of people with degree increases the transitions into employment of the most skilled occupations, while a larger share with A Level reduces the transitions into employment of the semi-skilled occupations and a larger share with O Level reduces the transitions of unskilled occupations. Secondly, re-estimating the model using a fully parametric approach assuming a Weibull hazard function (table 13) results in coefficients for the average education and the variance of education of similar sign and statistical significance, but of marginally larger magnitude (in absolute terms). These results are also robust to controlling for unobserved heterogeneity (RHS of table 13).

Considering all these results together we can conclude that the effect of the distribution of education on the transitions into employment takes a u-shaped form across occupations. The re-employment probability is higher (lower) for individuals belonging to a skilled (unskilled) segment of the local labour market where the average level of education is higher and where the variance is lower. The reason for this differential effect, as explained in the theoretical section, is that an increase in the average education level of a labour market generates two opposing forces affecting the matching process: a external or composition effect and a competition effect. The composition effect appears because a more qualified workforce increases expected profits by firms raising job creation and reducing unemployment duration. However, when average education is higher, the competition between unemployed individuals for the available vacancies becomes more intense, increasing unemployment duration. In the case of skilled jobs, where the worker's qualifications are very important for firms, the external effect dominates the competition effect. In the case of unskilled jobs, firms consider education as a minimum requirement but not as a fundamental determinant of the expected profits from the job. This means that job creation reacts less to a more qualified workforce

[19] Instead of using an ordinal classification of education like in the rest of the paper, we attribute the number of years of post-compulsory education required to acquire that qualification - 0 years for other, 1 year for O Level, 3 years for A Level and 5 years for degree.

Figure 7: Baseline hazard all individuals

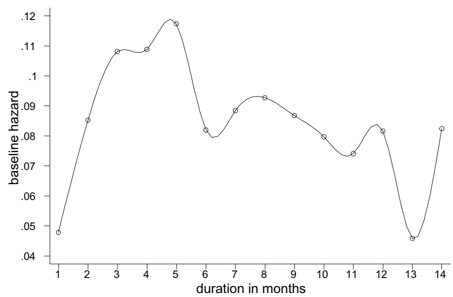
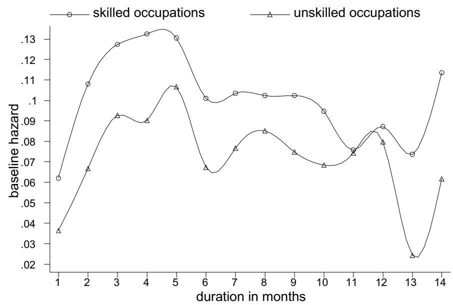


Figure 8: Baseline hazard by occupation group



and the competition effect dominates the external effect.

Figure 7 shows the estimated baseline hazard of the representative individual[20] for the estimation including all individuals from all occupations. The hazard of re-employment is increasing for durations up to 4-5 months and then decreasing, with some small peaks, which is consistent with the literature on unemployment duration (see Narendranathan and Stewart (1993) and Boheim and Taylor (2000)). The estimated baseline hazard for skilled and unskilled occupations shows a very similar picture (figure 8). The skilled occupations have a higher hazard than the unskilled ones for all durations. That is, individuals from the

[20]The representative individual has the following characteristics: male, 25-49 years of age, white, married, non-migrant, head of household, education of O Level or equivalent, personal or clerical occupation and living in the region with the median national average level of education (non-metropolitan areas of the West midlands region). We use the average value for all the other variables.

skilled occupations have a higher probability of leaving unemployment independently of how long they have been unemployed for.

In order to have an idea of the magnitude of this effect on the probability of re-employment we look at the shift in the estimated baseline hazard of a representative person after a change in the average level of education. A 1% rise in average education shifts the hazard upwards of the representative skilled individual by 2.9% on average, while a 1% decrease in the variance of education shifts the hazard upwards by 0.2% on average (table 4). The opposite pattern is obtained for the representative unskilled individual, - 1.9% and -0.2% for a 1% change in the variables. These numbers are comparable to those obtained by other studies in the literature. Moretti (1999) finds that a 1% rise in the share of college graduates increases wages of graduates by 1.2%, while Rauch (1993) finds that a one-year increase in average schooling raises individual wages between 3 and 5%.

Table 4: Average % change in baseline hazard

	All	Skilled Occup.	Unskilled Occup.
1% $\uparrow$ average education	0.40	2.88	-1.92
1% $\downarrow$ variance of education	0.03	0.24	-0.19
average region to Scotland	2.59	12.02	-5.27
average region to West Mid (met)	-3.40	-34.84	31.84

One can also use this method to have an idea of the magnitude of the effect of regional differences in the distribution of education. A representative individual of the skilled occupations would experience an increase in the probability of finding a job of 12% if the regional level of education increased from the national mean to the level of the top region (Scotland). That same individual would experience a reduction in his probability of leaving unemployment of 35% if the level of education decreased from the mean to the level of the worst region (West Midlands Met.). That is, regional differences in education could imply a difference in the probability of leaving unemployment of up to 47%. This number goes down to 37% when we consider the unskilled occupations.

The model has also been estimated separately by sex and occupational group and by age group and occupational group (table 5).[21]

[21] The age groups are defined as follows: young = 16-34 & old = 35-59 if female and 35-64 if male. We run

Table 5: Maximum likelihood estimates of re-employment probabilities by gender and age group

Variables	Semiparametric no unobserved heterogeneity all	skilled	unskilled	Semiparametric unobserved heterogeneity all	skilled	unskilled
Male						
Average education	0.787	3.289	−5.626***	1.097	1.938	−5.167**
var. of education	−0.251	−1.783	2.050***	−0.181	−1.653	2.123***
Female						
Average education	0.698	3.542**	−0.794	0.352	1.388**	−0.277
var. of education	−0.486*	−1.141*	0.257	−0.321	−0.443	0.325
Young (16 − 34)						
Average education	0.623	6.079***	−5.105***	0.410	1.982***	−2.184**
var. of education	−0.211	−2.114***	1.721***	−0.165	−1.139*	1.826**
Old (35 − 59/65)						
Average education	1.025	0.928	−0.613	0.560	0.736	0.048
var. of education	−0.755**	−0.653	0.332	−0.509*	−0.193	0.266

NOTES: a) *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.
b) The robust standard errors are clustered by region, occupation and time period.

The effect of the education externality in the matching process shows a similar pattern to the estimation for all the individuals, with a stronger positive effect for the most skilled occupations and a stronger negative effect for the least skilled occupations. However, the magnitude of the effect and the levels of significance are different across gender and age group. By age group, the effect is larger in magnitude and statistically more significant for younger workers independently of the level of aggregation of the occupations, while it is always insignificant for older workers. By gender, the average level of education has a stronger and more significant negative effect on the transitions into employment of male workers from unskilled occupations, while it has a stronger and more significant positive effect on the transitions into employment of female workers from skilled occupations.

Finally, we need to evaluate the relative importance of matching education externalities with

these estimations separately, not because we believe the externality should affect these groups differently, but because for other reasons, like female participation or youngsters' lack of job experience, these might be completely different labour markets.

respect to skill-biased technological externalities. As mentioned earlier, the theoretical model does not provide a way of distinguishing between these two types of externalities. In this study we have tried to do this by including variables in the estimation which act as a proxy for the technological level of the occupation. These variables are the share of employment of each industry in each occupation and region over time.[22]. In the first row of table 7 we have the results of estimating the standard econometric model excluding the variables measuring the technology in the occupation[23]. These estimation results are very similar to those of the standard estimation (including the regional technological level) although moderately lower. This suggests that technological externalities do not reduce the probability of leaving unemployment. Therefore, if one accepts this way of controlling for technology as correct, one should conclude that the most relevant type of externality in the matching process are the pecuniary externalities.

5.1 Regional labour market variables and individual controls

The effect of regional labour market variables on unemployment duration in the standard model estimated are shown in table 6. Firstly, the estimated coefficient of the regional unemployment rate has a negative sign and is significant for all the specifications, while the one for the inactivity rate has a positive sign and is significant for the skilled occupations (and for all the occupations together). The reason for this is that the bigger the number of people unemployed and/or active the greater the number of people looking for work and therefore the lower the market tightness, which reduces the probability of finding a job. This is an standard theoretical result of matching models (see Pissarides (2000) and Petrongolo and Pissarides (2001) for a survey of the literature) and is in accordance with the model developed in this paper. However, it is more difficult to explain why does a higher inactivity rate benefit only the skilled individuals. Anyway, this effect becomes insignificant when controlling for unobserved heterogeneity.

Secondly, The vacancy rate has a positive sign, as expected from the theoretical model (see also Petrongolo and Pissarides (2001)). The larger the number of vacancies opened, the easier it is for unemployed individuals to find a job and therefore the shorter the average unemployment spell. However, it is not significant at the 10% level. This is most probably

[22] They are calculated using 4 industries, 4 occupations and 19 regions.

[23] We have included instead variables measuring the regional technological level, i.e., the employment share of each industry in each region. The idea being that we still want to avoid the bias coming from technological externalities which are neutral across skills. But excluding these aswell does not change the results qualitatively

due to the fact that vacancies registered at job centers are known to under-represent the total number of vacancies.[24]

Table 6: Regional variables in standard estimation using all individuals in the sample (Table 2)

	Semiparametric no unobserved heterogeneity			Semiparametric unobserved heterogeneity		
Variables	all	skilled	unskilled	all	skilled	unskilled
Regional labour market variables						
Unemp rate	−0.630***	−0.687***	−0.841***	−0.646***	−0.880***	−1.033***
Vacancy rate	0.038	−0.040	0.046	0.054	0.096	0.057
Inactivity rate	0.561	1.283*	0.205	−0.188**	1.088	0.375
Regional migration						
Migration ratio	0.080**	0.087*	0.080**	0.035	0.111	0.075

Finally, the estimated coefficients of the individual and household variables have in general the expected signs, which are consistent with the existing literature. For the baseline model these are shown in table 14 in appendix D.

5.2 Robustness of Results

In this section we have tried to take into account some of the standard econometric problems one may encounter when estimating external effects. In particular, we have worried about the endogeneity of independent variables and an omitted variable bias (see Brock and Durlauf (2000) and Dietz (2001) for two excellent surveys on this issue).[25]

[24] Alternatively, if the education externality works through vacancy creation, the vacancy rate might not have a significant effect because it is already proxied by the average level of education. However, this seems unlikely, since dropping the regional average and variance of education from the estimation increases the significance of this variable only marginally.

[25] Another related problem when estimating external effects is that of identification or reflection problem (Manski (1993)). This problem arises when a researcher observing the distribution of behaviour in a population tries to infer whether the average behaviour in some group influences the behaviour of the individuals that comprise the group. However, Manski (1993) and Brock and Durlauf (2000) proved that identification is eased in non-linear models and in particular in duration models. In addition, they showed that this problem is further eased if there is within-group heterogeneity. Since our study fulfills both of these properties we understand that identification of the external effects should not be a problem here.

Table 7: Robustness of results by two occupation groups using all individuals in the sample

Variables	Semiparametric no unobserved heterogeneity all	skilled	unskilled	Semiparametric unobserved heterogeneity all	skilled	unskilled
Dropping industrial share						
Average education	0.951	2.616**	−3.139**	0.804	2.193**	−3.324**
var. of education	−0.489**	−1.577***	0.910*	−0.466**	−1.255***	1.270**
Dropping top (Scot, SE) & bottom (WMid (met), RestN) regions in education						
Average education	2.468**	5.058***	−1.527	2.083**	4.272**	−1.203
var. of education	−0.747**	−1.303*	0.628	−0.614*	−1.062*	0.590

NOTES: a) *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.
b) The robust standard errors are clustered by region, occupation and time period.

The endogeneity of independent variables is a common problem to all studies of external and neighbourhood effects. If location and occupational group were predetermined and fixed for all individuals there would not be a problem. However, we know that individuals are able to choose to a certain extent the area and group they belong to. In addition, we know that firms are also able to choose the area where they search for workers and the definition of occupational groups. This could generate a sorting process by which individuals with similar characteristics live together and belong to similar occupational groups. This would then be problematic for the estimation of the external effects because the characteristics determining the group of interest are the same as those determining the problem we are studying, therefore generating biased and inconsistent estimators (Greene (1993)).The question is whether the group characteristics used to measure these effects are exogenous variables or not with respect to the formation of the group. Therefore, the empirical definition of the market segment is very important. As explained in section 3, we have assumed the market segment is the occupation (using 4 occupational groups) within each region. One can think of two main problems with this definition. The first problem is that given the focus of this work is on education externalities, one would want to have segments whose boundaries do not change with the rise in education. But we know that in reality in parallel to the rise in education attainment there has been an increase in the education requirements of vacant jobs and therefore the boundaries of the occupational groups have moved upwards in the education scale. As a consequence, even if the working population had not increased their education level, we would observe a rise in the average education level of all the occupations.[26] This phenomenon has been named the problem of overeducation or degree's inflation, implying that education attainment is used as a signaling device. However, this issue will not bias the empirical results of this paper. To do so, it would require that the change in the boundaries of the occupational groups generates both, a rise in the average education of the segment and in the transition rates into employment of its members, for a given number of vacancies (otherwise it would be due to education externalities). While the former is always true, the latter is not. By definition, if the rise in the average education of each occupational group is only due to signalling reasons, it will not increase workers' productivity and, therefore, it will not increase the transition rates into employment. The second problem with the definition of labour market segment used is how to classify workers who do not report an occupation, either because they have never worked before or because they have been unemployed or inactive for a long period of time. In this work we have chosen to include these workers in the least skilled occupation. This is not very problematic for the case of previously inactive or long-term unemployed workers, since it is consistent with assuming that skills depreciate very quickly when out of work, a common assumption in the literature. However, it might be more problematic for the case of new market entrants. In this case, the estimated effect of education externalities could be biased downwards, since one would expect these workers to be very qualified and therefore improve the average level of education of the least skilled occupation, instead of the one of the occupation to which they should have been allocated.

The omitted variable bias is another potential problem. If there are omitted variables correlated with both the dependent variable and the regressors measuring the distribution of education, the estimated coefficients could be biased and wrongly significant.

Most studies in this literature solve both of these problems by instrumenting for the variables measuring the distribution of education. The problem is that instrumental variables is not a technique that can be used with duration models.[27] Therefore, the relevant question that needs to be answered here is how important is the endogeneity due to the sorting process and omitted variables bias in this particular study.

[26]This would be true even for the least skilled occupational group since its upper boundary would have risen aswell and it would englobe now some workers with higher education than before.

[27]But even if we could, we would require a variable that is correlated with the distribution of education but is not correlated with the sorting process or any omitted variable and varies both across occupation groups within regions and over time and such instrument is not available. Previous studies in this literature which used instruments where cross-sectional and focused on education externalities by geographical area, not by occupational group. This allowed them to use instruments which vary across areas but not across time, like the demographic structure a decade before (Moretti (1999) & Ciccone and Peri (2000))

Table 8: Percentage of the Variation of the education variable explained by region, occupation and time

	2 occupational groups		4 occupational groups			
	Skilled	Unskilled	High	Semi-high	Semi-low	low
time	6.9	12.4	37.6	18.9	30.7	30.6
region	9.5	18.1	44.4	74.9	57.1	58.9
occupation	81.6	65.5	−−	−−	−−	−−
region*time	0.1	0.6	18.0	6.2	12.2	10.5
region*occupation	2.0	1.5	−−	−−	−−	−−
occupation*time	0.0	0.0	−−	−−	−−	−−
region*occupation*time	−0.1	1.9	−−	−−	−−	−−

First of all, one would like to know what are the sources of exogenous variation in average education between segments, and of individual education within segments. Both types of exogenous variation are needed to be able to distinguish effects due to differences in aggregates from effects due to differences in individual endowments. As table 8 shows there is exogenous variation in average education between segments, the largest source of this variation coming from differences across occupational groups (82% for skilled and 66% for unskilled occupations), although differences over time and across regions are also important (16% for skilled and 31% for unskilled occupations). When looking at each occupation separately, it is regional differences that account for most of the variation in the education of the segment - between 44% and 75%. The existence of exogenous variation of individual education within segments is clear from the significance of the individual education dummies in table 14 in the appendix.

There are some reasons why the sorting process might not have an important effect on the estimation of external effects in this study. First of all, the areas used in this study - 4 occupations group within each of 19 regions - are very large and heterogeneous. Although migration of skilled workers will generate a higher concentration of qualified individuals in the most skilled occupations of the best regions, there will still be a large number of workers from all types in each region (eg. because of house ownership or housing benefit) and a perfect sorting process would be impossible. Secondly, even if the sorting process is important, there is no reason to believe that it has has been increasing during the sample period. Thus, it would mainly be a level effect which should be captured by the region, occupation and time dummies.

Nevertheless, we have undertaken some simple tests to check the potential importance of endogeneity. One would expect the effects of the sorting process to be stronger in the areas at the extremes of the distribution. However, when we run the regression excluding the regions that consistently have had the highest or lowest education level for the whole sample [28] we find that the effect of the externality remains very significant and, if anything, is stronger (see second row of table 7), but insignificant for unskilled occupations. Alternatively, one could try to understand the mechanism behind the sorting process. Regional migration would be one of the most important sorting mechanisms. If a large proportion of skilled workers move to the areas with highest wages and lowest unemployment rates, the average education of the workforce there will improve. At the same time, a high immigration rate will increase the demand for housing and raise the costs of living of the recipient area. This might push the least skilled individuals out of these areas, since they cannot afford to live there. Both of these forces increase the average education in the area. In order to take into account this particular sorting mechanism we have controlled in the estimation for people who have migrated into the region during the previous year but also for the aggregate migration flows in the region (number of immigrants over number of emigrants)[29]. Being a migrant has a positive but insignificant effect on the hazard of employment. This is due to the fact that the LFS underestimates the number of migrants because is a survey of non-movers. In terms of the migration flows into the region, the estimated coefficient is positive and significant, but it becomes insignificant when we control for unobserved heterogeneity. Overall, we believe these results indicate that, although there might be some endogeneity in the dependent variable, it is not having an important effect on the estimated coefficients.

It is not so easy to find simple tests to check the importance of a possible omitted variable bias. However, the fact that we include many individual, family and area variables in the estimation should reduce the importance of this effect. Moreover, in order to bias the coefficient measuring the external effect, the omitted variables have to vary both across groups and over time since the region, occupation and time dummies included in the estimation are already capturing any possible effect constant across groups or over time. One way of controlling for omitted variables in the estimation of a duration model is controlling for unobserved heterogeneity. However, this technique only controls for omitted variables which are uncorrelated with both the covariates and time. Nevertheless, one could use it

[28] Scotland, the South East, Metropolitan West Midlands and the Rest of the Northern region

[29] It would be preferable to have the migration ratio by occupational group in each region, but this data is only available from the Labour Force Survey and the sample size is too small to draw any significant conclusions.

as an indicator of the importance of this problem. As was mentioned above, controlling for unobserved heterogeneity (RHS of table 2) does not make the external effect disappear, although it reduces its magnitude. Therefore, even though it seems that this problem is not very important in this study, one should be aware that it could be biasing slightly the results.

6 Vacancy Creation and Education Externalities.

The theoretical model developed in this paper showed that a higher level of education improves the expected profits per vacancy opened, which in turn increases job creation and market tightness (see equation (9)), improving the probability of employment for a given aggregate supply of search units. That is, the positive external effect of education (or composition effect) in the matching process is generated through higher vacancy creation. So far, we have shown that education externalities raise the probability of finding a job for skilled occupations and lower it for unskilled occupations, but we have not analyzed the mechanism through which this effect is working. In this section we will try to do so, by using a quarterly panel on vacancies notified to UK job centers by occupational group (using the SOC 91 classification) and 19 regions for the period 1992-99. This data is obtained from job centers and is provided by NOMIS. We are conscious that the number of vacancies posted at job centers could be significantly lower than the real number of vacancies, especially for the most skilled occupations. This means that any result from this analysis should be considered as a lower bound of the total real effect of education externalities on vacancy creation. The covariates are obtained from the spring quarter of the non-longitudinal LFS.

We estimate a fixed effects panel using the log of market tightness in the local labour market as the dependent variable. This is defined as the number of vacancies notified divided by the number of unemployed by quarter, 9 occupational groups and 19 regions. The covariates are the average education and variance of education of the occupation group within each region and quarter, the unemployment, inactivity rates and the annual immigration and emigration rates by region. The fixed effects used in the estimation are occupation, region and quarter.

The estimation results confirm that in the UK local labour markets during the 1990s (table 9) the effect of education externalities on the matching process is working through higher job creation as indicated by the theoretical model. The effect has a similar pattern to the one on the duration of unemployment. An increase in the average education of the local labour market raises market tightness, while an increase in the variance reduces it. This effect is strongest for the skilled occupations. However, in this model the average education also has a positive effect on the market tightness of the unskilled occupations but of smaller magnitude and less significant, while the variance has a negative effect. The results still hold when we instrument using four lags of the education variables.

Table 9: Fixed Effects panel estimation of the Vacancy rate

Variables	Fixed Effects			I.V. Fixed Effects		
	all	skilled	unskilled	all	skilled	unskilled
Regional Education Externalities						
Av. education	0.895***	1.374***	0.682***	1.903***	2.831***	1.596***
s.d. education	−0.313***	−0.190***	−0.180***	−0.483***	−0.318***	−0.286***
Regional Labour Markets						
Unemp rt	−1.004***	−1.144***	−0.818***	−0.963***	−1.100***	−0.755***
Inactivity rt	1.880***	2.060***	1.683***	1.856***	2.078***	1.527***
Regional Migration by occupation						
Immigration rt	0.041***	0.045***	0.031***	0.042***	0.042***	0.035***
Emigration rt	0.027***	0.033	0.016***	0.018**	0.023	0.005**

NOTES: a) *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.
b) The robust standard errors are clustered by region, occupation and time period.

7 Conclusions

This study examines the effect of education externalities on the matching process taking place in UK local labour markets. First of all, we have shown theoretically that a higher level of education of the workers participating in the labour market raises the expected profits of opened vacancies, since firms expect to be matched with a more qualified worker. This increases job creation and reduces the unemployment rate. This is the external or composition effect. At the same time, a more educated labor force is also be more efficient searching for jobs. This will increase the competition amongst workers for opened vacancies and increase the unemployment rate. The net effect on unemployment depends on the relative importance of these two effects. We hypothesize that for the case of individuals from skilled segments of the labour market the external effect would dominate the competition effect. The reason being that is precisely in these markets where the worker's qualifications are a fundamental determinant of the firm's profits, while in the least skilled segments of the labour market are considered less important. Then, we tested this result empirically by estimating the effect of the education distribution in a labour market (measured by the average and variance of education) on the probability of transition from unemployment to employment.

We find that, for individuals belonging to skilled occupational groups (managers/professionals and technical occupations), a higher average level of education increases the probability of transition from unemployment to employment, while a higher variance, or a more unequal distribution, reduces it. However, the opposite is true for individuals belonging to less skilled occupational groups (clerical, service and manual). The reason behind this is that in the latter case the increase in the competition for vacancies among unemployed workers due to the higher level of education more than offsets the effect of the education externality.

The effect of the education externality on the hazard into employment is not only statistically significant but also relevant in magnitude. A rise of 1% in the average level of education shifts up the baseline hazard of the representative individual by 2.9% on average for the more skilled occupations, while it shifts it downwards by -1.9% for the less skilled occupational groups. Regional differences in the distribution of education also have a very large effect shifting the baseline hazard by around 47% for skilled occupations and by 37% for the nskilled occupations.

The estimated baseline hazard of the representative individual is increasing for durations up to 5 months and then decreasing, with some small peaks. This is true for all individuals and by occupational group. However, skilled occupational groups have a higher baseline hazard for all durations, indicating that they have a higher probability of leaving unemployment independently of how long they have been unemployed for.

Then, we confirme that the positive external effect of education (or composition effect) in the matching process is generated through higher vacancy creation by estimating the effect of the distribution of education on a fixed effects panel of the vacancy rate. We find that the average level of education increases the vacancy creation of skilled occupations, while the variance of education reduces it. However, we do not find the opposite effect for unskilled occupations.

Finally, the estimated results are very robust. They are qualitatively unchanged when we estimate the model controlling for unobserved heterogeneity, using different measures of education and using different parameterizations of the hazard function. In addition, the results are also robust to the standard econometric problems considered in the literature.

A Appendix: Definition of Education Variable.

Table 10: Education Variable

Education level	Composition
Degree or more	Higher degree, First degree or other degree, teaching (all levels)
High Vocational	Nursing, NVQ levels 3-5, HNC, HND, BTEC higher, RSA higher diploma, other higher education qualifications below degree
A Level or equivalent	A Level, Scottish 6th year Certificate, AS Level, SCE higher
Middle Vocational	NVQ level 3, GNVQ advanced, RSA advanced diploma, ONC, OND, BTEC and SCOTVEC national.
O Level or equivalent	O Level, GCSE grade A-C
Low Vocational	NVQ level 2, GNVQ intermediate, RSA diploma, City & Guilds advanced & craft, BTEC/SCOTVEC general diploma and completed apprenticeship
Other academic	CSE below grade 1, GCSE below grade C
Other vocational	NVQ level 1, GNVQ/GSVQ foundation level, BTEC/SCOTVEC general certificate, SCOTVEC modules, RSA other, City & Guild other, YT/YTP certificate, other vocational/professional qualifications
No qualifications	

B Appendix: Classification of regions.

Table 11: Classification of regions

Region	Counties
Rest of Northern Region	Cleveland, Cumbria, Durham, Northumberland
South Yorkshire	South Yorkshire
West Yorkshire	West Yorkshire
Rest of Yorkshire & Humberside	Humberside, North Yorkshire
East Midlands	Derbyshire, Leicestershire, Lincolnshire, Northamptonshire, Nottinghamshire
East Anglia	Cambridgeshire, Norfolk, Suffolk
Inner London	Inner London

continued on next page

Table 11: *continued*

Outer London	Outer London
Rest of South East	Bedfordshire, Berkshire, Buckinghamshire, East sussex, Essex, Hampshire, Hertfordshire, Isle of wight, Kent, Oxfordshire, Surrey, West sussex
South West	Avon, Cornwall, Devon, Dorset, Gloucestershire, Somerset, Wiltshire
West Midlands Metropolitan	West midlands Metropolitan
Rest of West Midlands	Hereford & Worcester, Shropshire, Staffordshire, warwickshire
Greater Manchester	Greater Manchester
Merseyside	Merseyside
Rest of North West	Cheshire, Lancashire
Wales	Clwyd, Dyfed, Gwent, Gwynedd, Mid Glamorgan, Powys, South Glamorgan, West Glamorgan
Strathclyde	Strathclyde
Rest of Scotland	Borders, Central, Dumfries & galloway, Fife, Grampian, Highland, Lothian, Northern & western isles, Tayside

C Appendix: Definition of the variables used in the estimation.

REGIONAL VARIABLES:

Average education: Average value of education variable (defined above) across all individuals of working age by occupational group (4, as defined in table 1), region and quarter.
variance of education: Variance of education variable (defined above) across all individuals of working age by occupational group (4), region and quarter.
Unemp rate: Unemployment rate across all active individuals by region and quarter.
Vacancy rate: Vacancy rate (number of vacancies notified to job centers / number of unemployed) by occupational group (4, as defined in table 1), region and quarter.
Inactivity rate: Inactivity rate across all individuals of working age by region and quarter.
Migration ratio: Migration ratio (number of immigrants / number of emigrants) by region and year.

INDIVIDUAL CHARACTERISTICS:

Sex: 1 male, 0 female.
Age: 16-24, 25-34, 35-49 years of age: 1 age group, 0 otherwise. (50-64 (59 female) reference category.)
Ethnic origin: 1 white, 0 otherwise.
Married: 1 married, 0 otherwise.
Migrant: 1 if migrated from another region since last year, 0 otherwise.
Head of Household: 1 head of household, 0 otherwise.
Individual's Education: Other Voc - Degree: 1 level of highest educational attainment, 0 otherwise.
Last Job's Occupational Group: Operator - Manager: 1 occupational group in last job, 0 otherwise.
no dep child < 6: number of dependent children under the age of six living in the household.
no dep child < 16: number of dependent children under the age of sixteen living in the household.
no working: number of members of the household working at the time of the interview.
one-person house: household made of one person only.
two-person house: household made of a couple.
Housing benefit: receiving housing benefit during unemployment.
Unemp benefit: receiving unemployment benefit during unemployment.
Family benefit: receiving financial help from relatives during unemployment.

Table 12: Sample means of individual characteristics by occupation group

Variables	all occup	skilled occup	unskilled occup
Duration Characteristics			
% exit into employment	40.8	45.0	37.1
% exit into inactivity	15.7	12.5	18.6
% stay unemployed	43.5	42.5	44.3
average duration	3.36	3.31	3.41
Individual Characteristics (%)			
Male	56.1	62.7	50.4
16-24 years of age	35.0	26.2	42.6
25-34 years of age	24.5	25.6	23.5
35-49 years of age	27.7	31.9	24.0

continued on next page

Table 12: *continued*

50-59/64 years of age	12.8	16.3	9.9
Non-white	7.4	6.4	8.3
Married	45.0	51.7	39.0
Migrant	2.2	2.3	2.2
Head of Household	39.3	46.0	33.4
Individual's Education (%)			
Other Voc	9.1	8.0	10.1
Other Acad	7.8	5.8	9.6
Low Voc	16.9	22.1	12.4
O Level	18.9	17.1	20.4
Mid Voc	2.9	3.8	2.0
A Level	6.3	7.2	5.5
High Voc	4.1	6.2	2.3
Degree	6.7	11.9	2.2
Last Job's Occupational Group (%)			
Operator	14.2	0.0	26.6
Sales	11.5	0.0	21.6
Personal	11.4	0.0	21.5
Clerical	15.7	0.0	0.0
Craft	14.6	31.3	0.0
Technical	5.4	11.7	0.0
Professional	3.8	8.2	0.0
Manager	7.1	15.3	0.0
Household Structure			
no dep child < 6	0.29	0.27	0.31
no dep child < 16	0.81	0.70	0.91
no working	1.2	1.2	1.2
% one person house	12.9	15.2	10.9
% two person house	23.7	26.0	21.6
Benefits (quarter before exit (%))			
Housing benefit	7.3	5.7	8.7
Unemp benefit	52.8	59.9	46.6
Family credit	1.3	0.9	1.7
No of cases	$15,974$	$7,459$	$8,515$

NOTES: "high occup" denotes occupations 5-9 of the SOC classification, while "low occup" denotes occupations 1-4 (see table 1.

Figure 9: Baseline hazard by sex

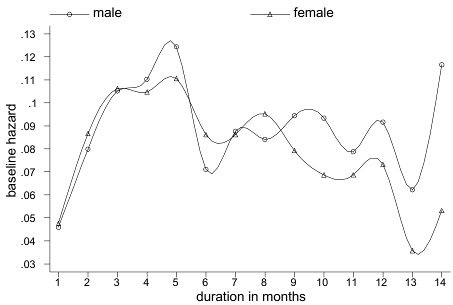
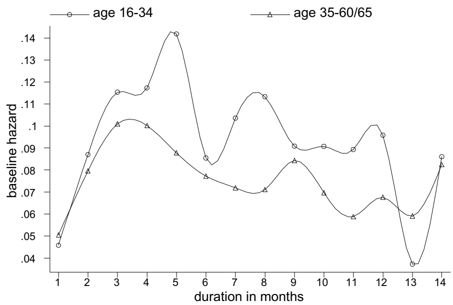


Figure 10: Baseline hazard by age group



D Appendix: Additional tables & figures

Table 13: Estimates by 2 occupation groups

	Weibull no unobserved heterogeneity			Weibull unobserved heterogeneity		
Variables	all	skilled	unskilled	all	skilled	unskilled
Baseline estimation (using levels of education)						
Average education	0.816	3.660***	−3.513**	0.724	2.871*	−3.608*
var. of education	−0.398*	−1.214**	1.247**	−0.304	−1.043	1.613***
Using Years of education						

Ave. years of edu.	0.464	2.650***	−1.557*	0.459	2.175**	−1.507*
var. years of edu.	−0.438	−2.151***	1.062**	−0.383	−1.991**	1.124**
Using Shares of Education (4 levels)						
Share degree	0.038	0.584	−0.011	0.064	0.257	−0.031
Share A Level	−0.118	0.105	−0.207	−0.080	0.055	−0.039
Share O Level	0.468***	−0.478	−1.301***	−0.432**	−0.584	−1.540***

NOTES: a) *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

b) The robust standard errors are clustered by region, occupation and time period.

Table 14: Individual Controls standard estimation using all individuals in the sample by two occupation groups (table 2)

	Semiparametric no unobserved heterogeneity			Semiparametric unobserved heterogeneity		
Variables	all	skilled	unskilled	all	skilled	unskilled
Individual Characteristics						
Sex	−0.034	−0.067	−0.027	−0.032	−0.136*	−0.002
16-24 years of age	0.116**	0.170**	0.081	0.464***	0.314**	0.084
25-34 years of age	0.465***	0.454***	0.486***	0.380***	0.652***	0.516***
35-49 years of age	0.346***	0.393***	0.279***	0.340***	0.581***	0.374***
White	0.235***	0.322***	0.130	0.268***	0.176	0.158*
Married	0.168***	0.124**	0.242***	0.183***	0.180*	0.198***
Migrant	0.057	0.029	0.088	0.022	−0.119	0.058
Head of Household	0.147***	0.191***	0.097**	−0.070**	0.362***	0.146***
Individual's Education						
Other Voc	0.145***	0.176**	0.129**	0.101**	0.156	0.116*
Other Acad	0.019	−0.067	0.050	−0.036	−0.101	0.016
Low Voc	0.139***	0.160***	0.141**	0.196***	0.216**	0.154**
O Level	0.039	0.101	0.017	0.123***	0.035	0.061
Mid Voc	0.061	0.332***	−0.253	0.308***	0.346**	0.095
A Level	0.017	0.089	−0.035***	0.137**	0.218	−0.061
High Voc	0.289***	0.413***	0.105***	0.351***	0.388***	0.186
Degree	0.176***	0.248***	0.103*	0.248***	0.195*	0.181*
Last Job's Occupational Group						
Operator	0.893***		1.292***	1.237***		1.149***
Sales	0.889***		1.313***	1.284***		1.158***
Personal	0.919***		1.353***	1.424***		1.588***

continued on next page

Table 14: *continued*

Clerical	0.776***			1.185***		1.519***
Craft	0.884***	0.350**		1.450***		
Technical	0.804***	−0.071		1.190***	−0.121*	
Professional	0.714*	−0.567**		1.418***	−0.868*	
Manager	0.483	−0.749***		1.149***	−1.013**	
Household Structure						
no dep child < 6	0.078***	0.337	0.087**	−0.173***	0.447***	0.057
no dep child < 16	0.312***	0.048***	0.294***	−0.027	0.053***	0.363***
no working	1.070***	1.072***	1.088***	1.105***	1.256***	1.353***
one person house	1.765***	1.772***	1.748***	−0.204***	1.611***	2.080***
two person house	1.089***	1.055***	1.136***	−0.086**	0.749***	1.313***
Benefits						
Housing benefit	−0.305***	−0.325***	−0.265**	−0.614***	−0.270	−0.325***
Unemp benefit	−0.129	−0.149***	−0.084*	−0.190	−0.123**	−0.147***
Family credit	0.313***	0.119	0.389***	0.217*	0.210	0.462***
Log Likelihood	−16079	−8139	−7868	−18571	−8264	−8032
Gamma Variance				0.439	0.532	0.450
Likelihood ratio st.				224.9	63.4	109.0
No cases	15974	7459	8515	15974	7459	8515

NOTES: a) *** denotes significance at the 1% level, ** at the 5% level and * at the 10% level.

b) The robust standard errors are clustered by region, occupation and time period.

References

[1] Acemoglu, D. (2003), "A Microfoundation for Social Increasing Returns in Human Capital Accumulation", *Quarterly Journal of Economics*, 111(3),779-804.

Acemoglu, D. (1998), "Why Do New Technologies Complement Skills? Direct Technical Change and Wage Inequality", *Quarterly Journal of Economics*, 113(4),1055-1089.

Acemoglu, D. and Angrist, J. (2000), "How Large Are Human Capital Externalities? Evidence from Compulsory Schooling Laws", *NBER Macroannual*, 15, 9-59.

Amiti, M. and Pissarides, C.A. (2002), "Trade and Industrial Location with Heterogenous Labour", CEPR Discussion Paper Series 3366.

Boheim, R. and Taylor, M.P. (2000), "Unemployment Duration and Exit States in Britain", CEPR Discussion Paper Series 2500.

Brock, W.A. and Durlauf, S.N. (2000), "Interactions-Based Models", NBER Technical Working Paper Series 258.

Burriel-Llombart, P. (2002), "Matching, Education Externalities and the Location of Economic Activity", PhD Dissertation, London School of Economics (unpubl.)

Ciccone, A. and Peri, G. (2000), "Human Capital and Externalities in Cities", CEPR Discussion Paper Series 2599.

Cox, D. (1972), "Regression Models and Life-Tables", *Journal of the Royal Statistical Society*, 34, B Series, 187-220.

Dietz, R. (2001), "Estimation of Neighborhood Effects in the Social Sciences: An Interdisciplinary Approach", URAI Working Paper Series 00-03.

Glaeser, E. L., Hallal, H. D., Scheinkman J. A. and Schleifer, A. (1992), "Growth in Cities", *Journal of Political Economy*, 100(6),1126-1152.

Greene, W.H (1993), *Econometric Analysis*, Prentice Hall.

Kiefer, N.M. (1988), "Economic Duration Data and Hazard Functions", *Journal of Economic Literature*, 26(2),646-679.

Krugman, P.A. (1991), "History and Industry Location: The Case of the Manufacturing Belt", *American Economic Review*, 81(2),80-83.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data*, Econometrics Society Monographs. Cambridge University Press.

Lucas, R.E. (1988), "On the Mechanics of Economic Development", *Journal of Monetary Economics*, 22(1), 3-42.

Manski, C. (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies*, 60(3), 531-542.

Marshall, A. (1920), *Principles of Economics*, London, MacMillan.

Moretti, E. (1998), "Social Returns to Education and Human Capital Externalities: Evidence from Cities", Center for Labour Economics, University of California, Berkeley, wp 9.

Moulton, B.R. (1986), "Random group effects and the precision of regression estimates", *Journal of Econometrics*, 32, 385-397.

Narendranathan, W. and Stewart, M.B. (1993), "Modelling the Probability of Leaving Unemployment: Competing Risks Models with Flexible Base-line Hazards", *Applied Statistics*, 42(1), 63-83.

Petrongolo, B. (2001), "Reemployment Probabilities and Returns to Matching", *Journal of Labor Economics*, 19(3), 716-741.

Petrongolo, B. and Pissarides, C. (2001), "Looking into the Black Box: A Survey of the Matching Function", *Journal of Economic Literature*, 39(2), 390-431.

Pissarides, C.A. (2000), *Equilibrium Unemployment Theory*, 2nd edition, Oxford: Basil Blackwell.

Prentice, R. and Gloeckler, L. (1978), "Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data", *Biometrics*, 34, 57-67.

Rauch, J.E. (1993), "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities", *Journal of Urban Economics*, 34, 380-400.

Romer, P.M. (1986), "Increasing Returns and Long-run Growth", *Journal of Political Economy*, 94(5),1002-37.

Rotemberg, J.J. and Saloner, G. (2000), "Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade", *Regional Science and Urban Economics*, 30, 373-404.

Simon, C.J. (1998), "Human Capital and Metropolitan Employment Growth", *Journal of Urban Economics*, 43(2), 223-243.

Simon, C.J. and Nardelli, C. (1996), "The Talk of the Town: Human Capital, Information, and the Growth of English Cities, 1861 to 1961", *Explorations in Economic History*, 33(3), 384-413.