A theory of education, unemployment and job mobility*

Bruno DECREUSE† Pierre GRANIER‡
CNRS-EUREQua GREQAM

January 2002

Abstract: This paper offers a theoretical perspective on the interactions between education and different aspects of mobility on the labour market. We propose a search equilibrium model in which the length of schooling defines the number of occupations a worker can apply for – what we term *job mobility*. Returns to education are twofold: there are *status*-mobility returns, since education raises a person's chance of leaving unemployment; but there are also *wage*-mobility returns, as education raises the speed at which one climbs the wage ladder. Unemployment incidence and duration as well as wage dispersion provide incentives to schooling; in turn, labour demand and wage dispersion respond to education investments. Unlike status-mobility returns, *wage*-mobility returns to schooling are not matched by a social gain, and workers over-educate. A combination of minimum wage and schooling fee can restore social efficiency.

Keywords: Wage posting; schooling duration; over-education; minimum wage

J.E.L. classification: J38, J64

*We thank Pierre Cahuc, Fabien Postel-Vinay and seminar participants at EUREQua and GREQAM for valuable comments. The usual disclaimer applies.

†EUREQua, Université Paris I Panthéon-Sorbonne – Maison des Sciences Economiques, 106-112 boulevard de l'Hôpital, 75647 Paris cedex 13, France. E-mail: decreuse@univ-paris1.fr

‡GREQAM, Université de la Méditerranée – 2, rue de la charité, 13002 Marseille, France. E-mail: granier@ehess.cnrs-mrs.fr

1 Introduction

Most empirical studies suggest that the returns to schooling are deeply related to the different kinds of mobility on the labour market. Education appears to play a key role on transition rates into and out of employment. Unemployment rates decrease with education attainment in almost all the OECD countries[1] (see Drèze and Sneessens, 1997, and Layard and Nickell, 1999). Estimates from duration models generally report that workers with higher levels of formal education move into employment more quickly than workers with less education (see Devine and Kiefer, 1991, for a survey). We shall refer to that kind of job mobility as *status*-mobility.

Status-mobility is only one part of the picture. Better-educated earn higher wages. The positive relationship between education and wages – captured by Mincerian equations – is indeed one of the most certain in Economics (see Card, 1999, for a survey). Besides, a large part of wage growth over one's career is due to occupational mobility (see, e.g. Topel and Ward, 1992, on a sample of young American male workers, and Manning, 2000, on UK data). Taken together, these empirical findings lead to the reasonable thought that most of the wage returns to schooling actually result from improved occupational mobility. This intuition is in accordance with the results obtained by Galor and Sicherman (1990) and Sicherman (1991) who show that given an origin occupation, schooling increases the likelihood of occupational upgrading, either within firms or between firms. We shall refer to that kind of job mobility as *wage*-mobility.

This paper offers a theoretical perspective on the interactions between education and these different aspects of mobility on the labour market. We focus on a set of three questions: First, how do both kind of job mobility affect incentives to schooling – namely, what are the roles played by the scarcity of job opportunities and wage dispersion? Second, how does in return education affect the labour market performance? Lastly, are private returns to schooling equal to, bigger or smaller than the social returns?

We thus propose a general equilibrium model in which schooling length has both status- and wage-mobility returns, and in which the size of those returns depends on

[1] To the noticeable exception of Italy and Israel.

aggregate educational investments. Job mobility is apprehended through the wage posting framework with on-the-job search initially elaborated by Burdett and Mortensen (1998) and extended by Mortensen (2000) to account for labour demand endogeneity; mobility returns to schooling stem from the fact that education has a direct effect on the rate at which one meets a wage offer. Assume that a continuum of equiproductive technologies can be implemented to produce a single final good. Each job, filled or vacant, is endowed with only one of these technologies. When the length of schooling determines the number of technologies a worker can operate, those who have completed the longest studies can apply for a larger proportion of jobs available in the economy[2]: they tend to exit faster from unemployment and benefit from better wage mobility once employed. Therefore, human capital has an horizontal dimension, i.e. the longer the schooling duration, the larger the number of occupations a worker can apply for. A consequence of this assumption is that increasing one's schooling effort creates a congestion externality for the other.

Due to the wage posting structure, there is a non-degenerate wage distribution even though workers are *ex ante* perfectly identical. Allowing workers to search on the job implies heterogenous reservation wages, depending on the worker's position on the wage ladder. Therefore, firms can pay higher wages to attract workers from firms offering lower wages. In this setting, workers do not only invest in education to leave unemployment more rapidly – status-mobility returns to schooling –, but also to upgrade faster in the wage distribution through a process of job shopping – wage returns to schooling. We use this structure to answer the set of questions marked above.

1. Two main determinants of the educational investment are emphasized: unemployment and wage dispersion. First, status-mobility returns to schooling increase when job opportunities are scarce, or when losing one's job is likely. In other words, at the microeconomic level, the schooling length tends to be increasing in the different components of unemployment (unemployment incidence and duration). At the macroeconomic level, those incentives translate into a positive structural relationship between aggregate schooling and the unemployment rate. Second, wage-mobility returns to schooling tend

[2] This view comes close to Thurow (1975), who argues that the higher educated workers are not more productive, but easier to train.

to increase when the wage dispersion becomes more dispersed. Indeed, climbing the wage ladder is not very exciting when the ladder has only a few rungs.

2. Schooling affects labour market performance in two differents ways. More educated/mobile workers are easier to recruit, so that they attract more firms on the search market. However, as workers meet a new offer more quickly, firms willing to keep their workers are obliged to post higher wages. To summarize, education tends to raise labour demand and wage dispersion. The unique equilibrium is thus defined by the intersection of two aggregate relationships between education and the unemployment rate: the upward sloping optimal schooling curve, and the downward sloping labour demand curve. Two main parameters affect the equilibrium mobility returns to schooling and the labour demand: the job destruction rate – downward mobility – and the minimum wage (MW) – which compresses the wage dispersion. On the one hand, education and the unemployment rate are increasing in the job destruction rate, reflecting both the higher rate of entry into unemployment and the increasing need of mobility to find a job more quickly, or to leave the low rungs of the wage ladder. On the other hand, the unemployment rate is strictly increasing in the MW, while the relationship between the MW and the schooling duration is typically U-shaped. Indeed, the MW compresses the wage distribution, which in turn depresses the wage-mobility returns to schooling; but it also increases the unemployment rate, and therefore raises the status-mobility returns. The first effect dominates (is dominated) at low (high) levels of the MW.

3. Turning to our third point of interest, we compare the decentralized outcome to the social commands which would be chosen were the stationary consumption maximised. Social returns to schooling consist in a more efficient matching technology; such returns are very close to the private status-mobility returns. From a social perspective, job-to-job transitions provide no gains (jobs are identical), but create congestion effects on the search market. Therefore, there is no social counterpart to the private wage-mobility returns. Consequently, private returns to schooling are above social returns and workers over-educate. We show that a combination of a minimum wage and a schooling fee per unit of time spent in the schooling system can restore social efficiency.

The closest paper to ours is that by Galor and Sicherman (1990) on career mobility.

Their model is characterized by a variety of occupations that are available for individuals within as well as across firms. Those occupations are ranked according to their skill requirement and to the wage they pay. In this environment, choosing an education investment means choosing a career path. Albeit quite appealing, this paper is only concerned by one component of job mobility – namely, wage mobility – and only offers a partial equilibrium analysis. By contrast, our paper highlights both wage and status components of job mobility, and adopts a general equilibrium perspective. This allows us to separate the impacts of the macroeconomic environment on incentives to schooling, and the effects of aggregate schooling on the performance of the labour market.

Beyond its focus on job mobility, our paper is also related to the large body of literature exploring the connexions between education and the labour market performance within the matching theory. Usually, such models stress two main predictions: because unemployment implies that human capital remains idle over long/frequent spells, it is detrimental to schooling investments (see e.g. Laing, Palivos and Wang, 1995). Second, rent-sharing implies that workers pay the full (marginal) cost of education but only get a share of the total (marginal) reward. Consequently, they under-invest in education (see Acemoglu, 1996). As noted by Moen (1999), both predictions are challenged by the (European) evidence. From that perspective, our model seems to be endowed with quite realistic properties. For instance, different levels of wage rigidity (minimum wage) are associated with different unemployment rates, but may well generate the same schooling duration.

The outline of the paper is as follows. Section 2 describes the structure of a wage posting model with endogenous schooling duration. Section 3 focuses on the individual schooling behaviour, while section 4 is devoted to a general equilibrium analysis. In section 5, we compare private and social returns to schooling and propose a policy to restore efficiency. Section 6 discusses some extensions of the model. Section 7 offers concluding comments.

2 The model

The demographics of the model has two main particularities. First, δ agents enter the market at each instant. This assumption will ensure that the inflow into the education system is constant. Second, agents may die at any moment and δ is also the (age-independent) risk of dying. The population is thus constant and normalized to unity; more importantly, this assumption will ensure that investments in education remain bounded even though the rate of time preference tends to 0. Individuals are risk-neutral and ρ is the pure rate of time preference, as well as the interest rate of the economy.

Production sector. There is a single final good which can be produced by means of a continuum of equiproductive technologies[3]. The set of technologies is normalized to $[0, 1]$. Advertising a vacancy implies posting a wage along with a given a technology. Firms choose the wage level, and pick a technology randomly by a draw from a uniform law on $[0, 1]$. This process is assumed irreversible, so that neither the wage nor the technology can be modified once a job offer is posted. Consequently, only those having acquired the technological knowledge to operate the required technology can apply for this job. Crucially, the number of technologies which can be operated for by a given individual depends on her schooling duration.

Schooling sector. Immediately after birth, the agents enter the schooling system. Education takes time, and its cost only consists in the value of (expected) earnings individuals forego during their schooling period. Education increases workers' mobility since it determines the number of technologies they can operate. A worker who has completed a curriculum of a given length T_i can apply for a job on the $H(T_i)$ submarkets for which he has the required skills, but not on the other $1 - H(T_i)$ submarkets.

The technology H is stricly decreasing and strictly concave. To ensure the existence of a solution to the maximisation problem, $H'(0) = H(0) = 0$, and $H'(\infty) = 0$. Moreover, $H(\infty) = 1$. We shall denote by $\lambda(T) \equiv H'(T)/H(T)$ the *rate of return* of schooling. The concavity of function H implies that the rate of return is decreasing in the schooling duration.

[3] The reason why such technologies coexist is beyond the scope of this paper.

Matching sector. After their schooling period, the agents enter the labour market as unemployed[4], and try to locate a vacancy. As in Burdett and Mortensen (1998), all agents search. The global flow of matches M between job-seekers and vacancies is given by the aggregate matching technology:

$$M \equiv M\left(\overline{H}.N, v\right) \tag{1}$$

where $\overline{H} \equiv E\left(H\left(T_i\right)\right)$ is the number of technologies which can be implemented on average by a "typical" worker, N is the workforce and v is the number of vacancies. The function M has the standard properties of a matching function: it is strictly increasing in each of its arguments, strictly concave and admits constant returns to scale. It satisfies the following boundary conditions: $M\left(x_1, 0\right) = M(0, x_2) = 0$ for all $x_1, x_2 \geq 0$, and the Inada conditions.

Let $m\left(\theta\right) \equiv M\left(1, \theta\right)$. The flow probability for a firm to meet an applicant is $\eta = M/V = m\left(\theta\right)/\theta$, where $\theta \equiv V/\left(\overline{H}N\right)$ is the labour market tightness. By analogy with the standard matching model with search intensity elaborated by Pissarides (1990, chapter 4), $H\left(T_i\right)$ can be interpreted as the individual search intensity[5]. Since education provides workers with a certain amount of search efficiency units, the function $H\left(T_i\right)$ determines the workers' mobility over the space of technologies. Thus, the instantaneous "probability" of locating a vacancy for an individual with a schooling effort T_i is

$$\mu_i = \frac{H\left(T_i\right) m\left(\overline{H}N, V\right)}{\overline{H}N} = H\left(T_i\right) m\left(\theta\right) \tag{2}$$

The individual contact rate is increasing in individual schooling. Mobility returns to schooling will therefore consist of a higher contact rate. At the aggregate level, it is worth noting that the efficiency of the matching process is increasing in workers' education.

Firms' and workers' gains. Let $S\left(T_i\right)$ and $W\left(T_i, w\right)$ be the values of being unemployed and employed respectively for a worker who has completed a period of schooling of length T_i. Let $r \equiv \rho + \delta$, q, w_0 be the global discount rate, job separation rate and MW rate.

[4] We do not allow the agents to search a job during the schooling period.

[5] The main difference with the standard model with search intensity is that search efforts have to be continuously repeated while education decisions are irreversible.

The values $S(T_i)$ and $W(T_i, w)$ satisfy the two arbitrage equations[6]:

$$rS(T_i) = \mu_i \int_{w_0}^{w_{\max}} [W(T_i, w) - S(T_i)]\, dF(w) \tag{3}$$

$$(r+q)\, W(T_i, w) = w + \mu_i \int_{w}^{w_{\max}} [W(T_i, \widetilde{w}) - W(T_i, w)]\, dF(\widetilde{w}) + qS(T_i) \tag{4}$$

where $F(w)$ is the *wage offer* distribution, w_0 is the MW and $w_{\max}$ is the (endogenous) upper-bound of the wage distribution. With (flow) probability μ_i, an unemployed is matched to a vacancy. Once employed, the worker undertakes on-the-job search to get a better wage. The (flow) probability of voluntary quit for a worker earning the wage w is $\mu_i [1 - F(w)]$, the contact rate between firms and workers, times the probability that the offered wage is higher than the current one. Obviously, the probability of finding a better wage offer diminishes when one climbs the wage ladder. Finally, with probability q the match is destroyed and the worker steps into unemployment where he enjoys the felicity level $S(T_i)$.

Let V and $J(T_i, w)$ denote the values of a vacant job and of a filled job. Those values satisfy the two arbitrage equations:

$$\rho V = \max_{w \geq w_0} \langle -\gamma + \eta [E_{T_i} p(w \mid T_i)\, J(T_i, w) - V] \rangle \tag{5}$$

$$(r+q)\, J(T_i, w) = y - w - \mu_i [1 - F(w)] [J(T_i, w) - V] + (q + \delta) V \tag{6}$$

Vacant job slots induce a flow cost of posting a vacancy $\gamma > 0$. With (flow) probability η, a worker is met. The probability that she will accept the job offer is $p(w \mid T_i)$; it depends on the worker's situation on the labour market, and therefore on her schooling duration. Denoting by $u(T_i)$ the unemployment rate of workers whose schooling duration is T_i, we get $p(w \mid T_i) = u(T_i) + (1 - u(T_i))\, G(w \mid T_i)$ – unity if unemployed, and $G(w \mid T_i)$ if employed, where G is the distribution of wages among the employees of schooling duration T_i. Since the education attainment of the incoming worker is *a priori* unknown for the firm, the value of a filled job has to be taken with an expectation operator in equation (5). Filled jobs bring a revenue to the firm of $y - w$. With a flow probability $q + \delta$, jobs are destroyed; with a flow probability $\mu_i [1 - F(w)]$, the worker quits to get a better wage.

[6] Equation (3) implicitly assumes that all wage offers will be accepted by the unemployed. It is indeed true since there are no unemployment benefits, and the contact rate is the same for employed and unemployed workers.

Firms are free to advertise vacancies. Free entry implies the exhaustion of all rents and drives the value of a vacancy down to zero. This yields

$$E_{T_i} p\left(w \mid T_i\right) J\left(T_i, w\right) = \gamma / \eta\left(\theta\right) \tag{7}$$

From equations (3) and (4), we get

$$\begin{aligned} S\left(T_i\right) &= \frac{\mu_i}{r+\mu_i} \int_{w_0}^{w_{\max}} W\left(T_i, w\right) dF\left(w\right) && (8) \\ J\left(T_i, w\right) &= \frac{y-w}{r+q+\mu_i\left[1-F\left(w\right)\right]} && (9) \end{aligned}$$

The next section describes the individual schooling behaviour.

3 Schooling choices

In this section, we focus on the incentives to schooling. We particularly emphasize the role of unemployment and wage dispersion.

We first compute $S\left(T_i\right)$. Integrating by part the right-hand side of equation (8), it comes:

$$\int_{w}^{w_{\max}} W\left(T_i, \widetilde{w}\right) dF\left(\widetilde{w}\right) = W\left(T_i, w_{\max}\right) - F\left(w\right) W\left(T_i, w\right) - \int_{w}^{w_{\max}} \frac{\partial W\left(T_i, \widetilde{w}\right)}{\partial \widetilde{w}} F\left(\widetilde{w}\right) d\widetilde{w} \tag{10}$$

But, from (4),

$$\partial W\left(T_i, w\right) / \partial w = \left(r+q+H_i m\left[1-F\left(w\right)\right]\right)^{-1} \tag{11}$$

Since $W\left(T_i, w_{\max}\right) = \frac{w_{\max}}{r+q} + \frac{q}{r+q} S\left(T_i\right)$,

$$S\left(T_i\right) = \frac{1}{r}\left\{\frac{H_i m}{r+q+H_i m} w_0 + \int_{w_0}^{w_{\max}} \frac{\left(1-F\left(w\right)\right) H_i m}{r+q+H_i m\left[1-F\left(w\right)\right]} dw\right\} \tag{12}$$

Optimal schooling results from

$$\widehat{T}_i = \arg\max_{T_i \geq 0} e^{-rT_i} S\left(T_i\right) \tag{13}$$

The first-order condition writes down $rS\left(\widehat{T}_i\right) = S'\left(\widehat{T}_i\right)$. The marginal cost of schooling consists of the value of (expected) foregone earnings. The latter is equal to the value of search times the effective rate of discount. The marginal benefit can be decomposed as follows:

$$S'(T_i) = \lambda_i \frac{r+q}{r+q+\mu_i} \left\{ \underbrace{S(T_i)}_{\text{status-mobility return}} + \underbrace{\frac{1}{r}\mu_i \int_{w_0}^{w_{\max}} \frac{F(w)(1-F(w))\mu_i}{\left[r+q+\mu_i(1-F(w))\right]^2} dw}_{\text{wage-mobility return}} \right\} \quad (14)$$

There are two returns to schooling: a *status-mobility return*, since education increases the probability of getting a job; a *wage-mobility return* as it also determines the speed at which the worker climbs the job ladder. The analytical expression of those returns are very close to each other because the job-finding rate of an unemployed is assumed to be equal to the wage offer rate of an employee. However, there are two reasons to distinguish them. First, changes in the environment (institutional, or macroeconomic) will affect differently each type of return. Second, job-to-job transitions correspond to a social waste, while status-mobility return will be matched by a social gain. Finally, it is worth noting that, neglecting the wage-mobility return to schooling, the marginal benefit is equal to a given factor $\lambda_i \frac{r+q}{r+q+\mu_i}$ times the value of search. This factor is lower than the rate of return λ_i; this reflects the fact that schooling is less useful once in the job (status-mobility returns vanish).

Proposition 1 *Individual education attainment solves*

$$\begin{array}{l} \frac{w_0}{(r+q+\mu_i)^2} \left\{ \lambda_i - r - \frac{r}{r+q}\mu_i \right\} + \\ \int_{w_0}^{w_{\max}} \frac{(1-F(w))}{[r+q+(1-F(w))\mu_i]^2} \left\{ \lambda_i - r - \frac{r}{r+q}(1-F(w))\mu_i \right\} dw = 0 \end{array} \quad (15)$$

It has the following properties:

(i) $r < \lambda_i < r + \frac{r}{r+q}\mu_i$

(ii) λ_i *tends to* $r + \frac{r}{r+q}\mu_i$ *as* w_0 *tends to* $w_{\max}$

(iii) λ_i *tends to* r *as* $m(\theta)$ *tends to 0, or* q *tends to infinity;* λ_i *tends to infinity as* $m(\theta)$ *tends to infinity*

The three properties we have highlighted concern the magnitude of education investment, the impact of the wage dispersion, and the relationship of the schooling duration with the different components of unemployment (unemployment incidence and duration).

Property (i) shows that despite the apparent complexity of equation (15), the schooling duration remains bounded, and those bounds have a strong economic meaning. They depend on the size of the wage-mobility returns to schooling. Neglecting those returns leads to the upper bound of the optimal rate of return: in that case equation (14) implies that $rS(T_i) = \lambda_i \frac{r+q}{r+q+\mu_i} S(T_i)$, which is equivalent to $\lambda_i = r + \frac{r}{r+q}\mu_i$. Conversely, magnifying those returns leads to the lower bound λ_i. Once employed, the relative importance of improving one's mobility diminishes, as more and more wage offers are refused. Abstracting from such effects provides a supremum to the marginal benefit of schooling: $S'(T_i) < \lambda_i S(T_i)$. Therefore, $\lambda_i < r$.

Property (ii) shows that optimal duration of schooling is minimal when the wage dispersion is nil. The reason is that the wage-mobility returns to schooling vanish in that case. This suggests that in most plausible cases, the individual schooling duration should be increasing in the wage dispersion. However, this cannot be shown in the general case without investigating the relationship between the whole wage distribution and its support; this is done in the next section for the equilibrium wage distribution.

Property (iii) puts the emphasis on the effects of incidence and duration of unemployment on the incentives to schooling. Once again, it is difficult to exhibit analytically a systematic link[7]; however limit properties offer some intuitions. As education investments are aimed at improving workers' mobility, they become useless in a perfectly competitive environment where the contact rate between workers and job offers tend to infinity (or when tightness becomes very high). Conversely, it is maximal when the matching process is so inefficient that is virtually impossible to get a job (when tightness is very low).

4 Schooling and mobility in general equilibrium

In this section, we adopt a general equilibrium perspective and solve simultaneously for a schooling duration, a wage distribution, and an unemployment rate. We only focus on symmetric equilibria. Identical individuals thus choose the same schooling duration, so

[7]Suppose the wage dispersion is nil. It is then clear that schooling is increasing in job destruction, and decreasing in tightness. Other effects seem to be of second-order, but this cannot be established.

that $T_i = T$. Then, the contact rate of individuals with vacancies is $\mu_i = \mu = H(T)\, m(\theta)$.

Population. We first characterize the various (stationary) mass numbers of workers in each of these states: labour force (N), unemployed (U), employed (L). The dynamics of workforce and employment are:

$$dN/dt = \delta e^{-\delta T} - \delta N \tag{16}$$

$$dU/dt = \delta e^{-\delta T} + q(N - U) - (Hm + \delta)\, U \tag{17}$$

Therefore, the measure of the workforce is worth $N = e^{-\delta T}$, the participation rate of the economy. The number of unemployed is $U = uN$, where $u = \frac{\delta + q}{\delta + q + Hm}$ is the unemployment rate. Employment is residually defined as $L = N - U$.

Wage distributions. We briefly present the wage distributions F and G – this does not differ from Burdett and Mortensen (1998). The *wage distribution* $G(w) \equiv G(w \mid T)$ is derived from a flow equilibrium approach. The outflow from the pool of employed workers earning less than w is

$$G(w)(1 - u)\, e^{-\delta T} [\delta + q + \mu (1 - F(w))] \tag{18}$$

This is equal to the proportion of workers paid less than w, times employment, times the rate of gross job destruction $\delta + q$ plus the rate at which those workers obtain wage upgrading. The inflow in the same pool is $\mu u e^{-\delta T} F(w)$: the job-finding rate times unemployment times the proportion of wage offers below w. These two facts together with the Beveridge curve $u = (q + \delta) / (q + \delta + Hm)$ imply that

$$G(w) = \frac{(q + \delta)\, F(w)}{q + \delta + \mu [1 - F(w)]} \tag{19}$$

The probability that a given worker will accept the job offer is thus $p(w) \equiv p(w \mid T) = \frac{q+\delta}{q+\delta+\mu[1-F(w)]}$. The *wage offer distribution* $F(w)$ is then computed from the following free entry condition:

$$\frac{\gamma}{\eta} = p(w)\, J(T, w)\,, \text{ for all } w \in [w_0, w_{\max}] \tag{20}$$

Using (9) and solving for $F(w)$, we get

$$1 - F(w) = \frac{-(2q+r+\delta) + \left\{ (2q+r+\delta)^2 - 4(q+\delta)\frac{\eta}{\gamma}\left[(r+q)\frac{\gamma}{\eta} - y + w\right]\right\}^{1/2}}{2Hm} \tag{21}$$

which defines the equilibrium wage offer distribution for all $w \in [w_0, w_{\max}]$, where the upper-bound of the distribution is

$$w_{\max} = y - (r + q)\,\gamma/\eta \tag{22}$$

by virtue of $F(w_{\max}) = 1$. Finally, the tightness can be derived from equation (20) evaluated in $w = w_0$:

$$\frac{\gamma}{\eta(\theta)} = \frac{q+\delta}{q+\delta+Hm(\theta)}\frac{y-w_0}{r+q+Hm(\theta)} \tag{23}$$

In the remainder, we only focus on the case where the rate of time preference ρ tends to 0. Moreover, we adopt the following minimum wage policy[8]: $w_0 = \beta y$, $\beta \in [0, 1)$. The parameter β will be refered to as the *minimum wage factor*.

Equilibrium. The equilibrium consists of a tightness, which solves the free entry condition (23), a wage distribution, summarized by equations (21) and (22), a schooling duration, which results from (15), and an unemployment rate, defined by the Beveridge curve $u = (q+\delta)/(q+\delta+Hm)$. It is convenient to solve tightness as a function of the unemployment rate and the schooling duration from the Beveridge curve:

$$m(\theta(u,T)) = \frac{q+\delta}{H(T)}\frac{1-u}{u} \tag{24}$$

Manipulating the different equations yields (after some algebra):

$$\gamma/\eta(\theta(u,T)) = \kappa u^2 y/(q+\delta) \tag{FE}$$

$$\left[1+\kappa+2\kappa\frac{u}{1-u}\ln u\right]u\lambda(T) = \delta[1-\kappa u] \tag{OS}$$

where $\kappa \equiv 1-\beta$. Equilibrium is defined by the intersection of two curves: the free entry equation (FE), and the optimal schooling equation (OS) – once taken into account the specific nature of the wage distribution. Both curves relate the unemployment rate to the schooling duration.

Lemma *There exists a unique equilibrium*

[8] This assumption is innocuous; it simply helps saving on notation.

(FE) defines the unemployment rate as a strictly decreasing function of the schooling duration. Indeed, education improves the efficiency of the matching process. The number of vacancies per job-seeker is therefore increasing in aggregate schooling. (OS) defines the schooling duration as a strictly increasing function of the unemployment rate. We have shown in section 3 that education was positively related to the incidence and duration of unemployment. In the particular case where the rate of time preference is nil, those properties imply that schooling directly depends on the unemployment rate. This would not hold in the general case with positive discounting. As the curves (FE) and (OS) have opposite slopes, the equilibrium is unique.

Two parameters are of great economic importance concerning job mobility: the job destruction rate and the MW. The effects of these parameters are analyzed in proposition 2 and depicted by figure 1.

Proposition 2 *There exist* (β_0, β_1)*,* $0 < \beta_0 < \beta_1 < 1$*, such that*

(i) $du^*/dq > 0$ *and* $dT^*/dq > 0$ *for all* $\beta \in [0, 1)$ *and all* $q \geq 0$*;* $\lim_{q\to\infty} T^* = \lambda^{-1}(\delta)$

(ii) $du^*/d\beta > 0$ *and* $dT^*/d\beta < 0$ *whenever* $\beta < \beta_0$ *and* $dT^*/d\beta > 0$ *whenever* $\beta > \beta_1$*;* $\lim_{\beta\to 0} T^* = T_0^* < \lambda^{-1}(\delta)$ *and* $\lim_{\beta\to 1} T^* = \lambda^{-1}(\delta)$

Job destruction only affects the (FE) locus, through its effects on the Beveridge curve, and on the value of a filled job. Both effects imply that the unemployment rate must be higher at given schooling duration, so that the (FE) locus moves up. Due to the positive slope of the (OS) curve, both the unemployment rate and the schooling duration increase. As the job destruction rate becomes arbitrarily large, the unemployment rate tends to 1, and the rate of return to schooling $\lambda(T^*)$ tends to the rate of death δ.

[Figure 1]

The MW deteriorates profit opportunities. Consequently, following an increase in the MW factor, the (FE) locus moves up, which tends to increase both the unemployment rate and the schooling duration. However, the MW also reduces the incentives to schooling by compressing the wage distribution; the (OS) locus moves upward too. There are thus two conflicting effects of the MW on the schooling duration: a negative effect, through falling wage-mobility returns to education, and a positive effect through falling tightness. The negative (positive) effect dominates for low (large) values of the MW factor: the relationship between education and the MW tends to be U-shaped[9]; this is depicted by figure 2.

[Figure 2]

Empirically, proposition 2 implies that two economies characterized by different degrees of wage rigidity (different β) know different unemployment rates, but may well exhibit the same aggregate educational level. In the low MW country, employment is higher and the wage dispersion is relatively more dispersed. Schooling is mainly motivated by wage upgrading through job-to-job mobility. In the high MW country, the wage distribution is more compressed so that job-to-job mobility is relatively less important for education investments; however, the labour demand is weak and workers train so as to compensate (partially) for the low number of vacancies per job-seeker.

We now turn to efficiency considerations.

5 Private and social mobility returns to schooling

This section is mainly devoted to a comparison between social and private returns to schooling. We show that firms tend to post too many vacancies, and workers tend to get too much training. An adequate combination of minimum wage and schooling fee can restore efficiency. The social criterion we choose is standard: a social planner maximises the discounted value of the path of aggregate consumption. When the rate of time preference ρ tends to 0, it is sufficient to maximise stationary consumption. Therefore,

$$\Omega = \max_{T,\theta} \left\{ e^{-\delta T} (1-u) \, y - \gamma v \right\} \tag{25}$$

s.t. $u = \frac{q+\delta}{Hm+q+\delta}$ and $\theta = \frac{v}{He^{-\delta T}}$. We thus assume that the social planner takes as given the fact that all workers search (this assumption is dicussed longer below). The term

[9] All the numerical experiments we have performed under the assumption of a Cobb-Douglas matching technology provide a U-shaped relationship between education and the MW.

$e^{-\delta T}$ is the participation rate of the economy. Note that our assumption of perpetually young agents here makes particular sense: if individuals were infinite-lived, their education duration would tend to infinity.

Let $\alpha \equiv \alpha(\theta) \equiv \theta m'(\theta)/m(\theta)$ be the elasticity of the contact rate per unit of search. The desired rate of unemployment u^p and schooling duration T^p solve the first-order conditions:

$$\frac{\gamma}{\eta(\theta(u,T))} = \alpha u^2 y/(q+\delta) \tag{26}$$

$$(1-\alpha)\, u\lambda = \delta[1-\alpha u] \tag{27}$$

Equations (26) and (27) must be compared to equations (FE) and (OS), *i.e.* the market outcome.

The first point to notice is that at given education attainment, Mortensen's (2000) result still holds: private tightness tends to be too high, and thus the unemployment rate tends to be too low. While posting vacancies, firms increase the rate at which jobs are destroyed, a feature they do not take into account. By lowering firms' profits, the MW can be used to correct this market failure. The optimal MW factor is then $\beta = 1-\alpha$. However, in our framework, the MW will also affect education decisions.

The second point is that over-education takes place. To see this, simply rewrite (27) as follows:

$$\lambda^p = \delta + \frac{1}{1-\alpha}\frac{\delta}{q+\delta} H^p m^p \tag{28}$$

As $\lambda^* < \delta + \frac{\delta}{q+\delta} H^* m$ (see proposition 1), it follows that $\lambda^* < \lambda^p$; consequently, $T^p < T^*$ for a given tightness. The reasons why workers over-educate can be infered from the social marginal benefit and marginal cost of schooling. The marginal benefit MB is worth:

$$MB = e^{-\delta T} u(1-u)\lambda y \tag{29}$$

The marginal benefit consists in a reduction in the unemployment rate through an increase in the search process efficiency. This corresponds to the status-mobility returns to schooling we previously highlighted. But, there are no social wage-mobility returns to schooling. Consequently, the social marginal benefit of schooling is lower than the private benefit.

Consider now the social marginal cost:

$$MC = \delta e^{-\delta T}\left[(1-u)\,y - \gamma\theta H\right] + \gamma\theta H e^{-\delta T}\lambda \tag{30}$$

It is composed of two terms. The first is the decrease in the rate of participation following a marginal increase in the schooling duration. This corresponds to the foregone "earnings" of a student[10]. The second is the increase in recruitment costs induced by the higher search intensity of employed workers. There is no private counterpart to this second cost, so that the social marginal cost is larger than the private one. Over-education thus results from both a too large valuation of the schooling benefits, and too small valuation of the induced costs.

Can the social optimum be decentralized? There are two types of externalities, therefore two policy tools are required: the MW to affect the labour market, and a schooling fee to alter education behaviours. Let $\Sigma \equiv \sigma y$ be the level of the fee per unit of time spent in the education system. Assume also that the global amount of collected taxes is redistributed to newborn individuals. The lump-sum transfer can then be used to finance the tuition fee. Optimal schooling results from:

$$\max_{T}\left\langle e^{-\delta T} S(T) - \int_0^T e^{-\delta z}\sigma y dz\right\rangle \tag{31}$$

Proposition 3 *The decentralized outcome is efficient provided that* $\beta = 1-\alpha^p$ *and* $\sigma = \widehat{\sigma} \equiv 2\frac{\alpha^p}{1-\alpha^p}\left(1-\alpha^p u^p\right)\left(1-u^p\right)\left[1+\frac{u^p}{1-u^p}\ln u^p\right]$.

The fee is all the higher than the targeted unemployment rate u^p is low. The reason is that the share of wage-mobility returns to schooling in total returns to schooling is decreasing in the unemployment rate. Workers must find a job before even thinking to climb the job-ladder. This is simply more obvious when the unemployment rate is very high.

We conclude this section by a discussion about the welfare criterion we have chosen. The fact that all workers search is indeed highly inefficient. If the social planner could prevent job-to-job mobility, then the number of job-seekers would be equal to the number

[10] The rate at which participation is reduced is equal to the private rate at which foregone earnings are evaluated as long as the population is constant.

of unemployed. Tightness would then write down as $\theta = \frac{v}{Hue^{-\delta T}}$; the welfare criterion would then be:

$$\Omega' = \max_{T,\theta} e^{-\delta T} \left\{ \frac{Hm}{Hm+q+\delta} y - \gamma \frac{\theta H}{Hm+q+\delta} \right\} \tag{32}$$

We shall refer to the social commands corresponding to this criterion as defining the first best optimum, while the previous commands will be refered to as the second best optimum. First-order conditions yield:

$$\frac{\gamma}{\eta^o} = \alpha \frac{y}{q+\delta+(1-\alpha) H^o m^o} \tag{33}$$

$$(1-\alpha) \frac{q+\delta}{q+\delta+(1-\alpha) H^o m^o} \lambda^o = \delta \left[1 - \alpha \frac{q+\delta}{q+\delta+(1-\alpha) H^o m^o} \right] \tag{34}$$

A close look at equations (33) and (34) immediately shows that for a given schooling duration, the tightness is higher in the first best than in second best case. The reason is obvious: eliminating socially useless job-to-job transitions makes job creations socially more desirable. The comparison between equations (26) and (27) is a little bit less trivial. However, the term $\frac{q+\delta}{q+\delta+(1-\alpha)H^o m^o} > u$ for given tightness and schooling duration. Therefore, for a given tightness, schooling is longer in first best than in second best. Due to the decreasing relationship between education and tightness, we cannot conclude in all generality. Nevertheless, on-the-job search severely limits both job creation and education opportunities.

6 Discussions

In this section, we discuss some extensions regarding the impact of education on the productivity of the match.

6.1 Education and productivity

Suppose that the productivity of a match is increasing in aggregate schooling, i.e. $y \equiv y(T)$, thanks to an (unmodelized) externality. Three main implications can be analyzed. First, firms cannot design contracts to internalize the relationship between education and productivity. Therefore, this new assumption does not change the individual schooling investment equation (15). Consequently, workers only aquire skills in order to improve their mobility, even though wage perspectives are increasing in match productivity. In the absence of incentives to acquire mobility – such as a high unemployment rate, or a very dispersed wage distribution – education investments collapse and productivity is kept minimal. From this perspective, search frictions are a key factor underlying macroeconomic performance.

Second, the (OS) curve is also unchanged (partly because the MW is indexed on productivity), while the (FE) curve moves upward. Indeed, education not only improves the efficiency of the search process, but it also raises productivity and thus firms' profits. A consequence is that an increase in the job destruction rate, which makes the schooling period longer, can also increase the wage dispersion through its positive effect on the upper-bound of the wage distribution.

Finally, as education investments do not depend on the relationship between education and productivity, under-education becomes a plausible output. More particularly, the schooling length chosen by the social planner is worth:

$$\lambda_y + (1-\alpha)\, u\lambda_H = \delta\, [1 - \alpha u] \tag{35}$$

where $\lambda_y \equiv y'/y$ and $\lambda_H \equiv H'/H$. The fact that workers train too much or too little depends on the relative effects of education on productivity λ_y and mobility λ_H. On the one hand, if λ_H is much larger than λ_y at all schooling durations, then the previous analysis still holds and workers over-educate as indicated above. On the other hand, if λ_H is much lower than λ_y then the schooling duration chosen by the social planner tends to the solution of $\lambda_y^p = \delta\, [1 - \alpha u^p] < \delta < \lambda_H^*$. Therefore $T^p > T^*$ and under-education takes place.

6.2 On-the-job training

In our paper, the wage growth over one's career is only due to job mobility. The literature offers a competing explanation according to which there are returns to seniority: workers receive specific training, which improves the productivity of the match and the wage the worker gets. As educated workers are easier to train, they receive more training and earn better wages. To discuss this view, we incorporate costly on-the-job training in our model and assume the training cost is decreasing in the worker's education. The main results are that (i) leaving aside their lower training cost, higher educated receive more training because they are better paid on average, and (ii) the amount of training may decrease with education at low wages.

Let e represent match-specific training, and let the value of a worker productivity be $y \equiv y(e)$, where y is strictly increasing and strictly concave, with $y'_e(0) = 0$. Firms carry the full cost of training, but as in Thurow's analysis (1975), the training cost is decreasing in education. We assume that $C(e, T) = c(T)e$, with $c'(T) < 0$. Optimal training results from

$$\widehat{e} = \max_e \langle J(T, e, w) - c(T)e \rangle \tag{36}$$

Under free entry ($V = 0$), it comes

$$y'(\widehat{e}) = c(T)\{r + q + H(T)m(\theta)[1 - F(w)]\} \tag{37}$$

As in Mortensen (2000), the amount of training provided is an increasing function of the wage: high wage workers are less likely to quit. *Ceteris paribus*, this implies that training tends to increase with education simply because higher educated workers climb the wage ladder more rapidly. For a given wage, education has ambiguous effects on the amount of training; differentiating with respect to T we get that $d\widehat{e}/dT$ has the sign of

$$\nu\left[1 + \frac{r+q}{\mu[1 - F(w)]}\right] - \lambda \tag{38}$$

where $\nu \equiv -c'(T)/c(T) > 0$. The first term is positive: as education reduces the training cost, education favourizes training. The second term is negative: a more educated worker is more likely to quit, which discourages training. One of the following situations prevails: either training is strictly increasing in education, or there exists a wage $\overline{w} \equiv \overline{w}(T)$, such that training is decreasing in education for wages below $\overline{w}$, and increasing above.

7 Concluding comments

In an imperfect labour market, workers may be willing to invest in education to improve their employment prospects Conversely, macroeconomic conditions are shaped by the size of educational investments. We hence consider a search equilibrium model in which the schooling duration determines the fraction of jobs that a worker can apply for – what we term *job mobility*. Returns to education are twofold: there are *status*-mobility returns, since education raises a person's chance of leaving unemployment; but there are also *wage*-mobility returns, as education raises the speed ay which one climbs the wage ladder. Both determinants of unemployment (unemployment incidence and duration) as well as the wage dispersion provide incentives to schooling; in turn, the labour demand and the wage dispersion respond to education investments.

In general equilibrium, job destruction increases both the unemployment rate and the schooling duration by raising status-mobility returns. The relationship between education and the minimum wage is typically U-shaped: on the one hand, the minimum wage reduces wage-mobility returns to schooling by lowering the wage dispersion; on the other hand, it raises status-mobility returns to schooling through its adverse effects on the labour demand, which translates into longer unemployment spells on average. The former effect dominates (is dominated) at high (low) levels of the minimum wage. As job-to-job transitions correspond to a social waste, wage-mobility returns to schooling are not matched by a social gain, and workers over-educate. A combination of minimum wage and schooling fee can restore social efficiency.

APPENDIX

A Proofs

proposition 1 Let $L(x) = e^{-rx}S(x)$. The function L is continuously differentiable and such that $L(x) \geq 0$ for all $x \geq 0$, $L(0) = 0$, $L(\infty) = 0$. The first-order condition is thus necessary. Manipulating it yields equation (15). Points (i) to (iii) follow from the continuity of the optimal schooling duration with respect to its different components. ■

Lemma Let $\theta(u,T)$ be the solution of $u = \frac{q+\delta}{q+\delta+H(T)m(\theta)}$ and consider the functions ϕ_1, ϕ_2 such that

$$\phi_1(u,T) \equiv \frac{\gamma}{\eta(\theta(u,T))} - \kappa u^2 \frac{y}{q+\delta} \tag{39}$$

$$\phi_2(u,T) \equiv [1+\kappa+2\kappa\ln u]\,u\lambda - \delta[1-\kappa u] \tag{40}$$

An *equilibrium* is a pair (u^*,T^*) which solves $\phi_1(u^*,T^*) = \phi_2(u^*,T^*) = 0$.

Step 1. *Let* $\alpha \equiv \alpha(\theta) \equiv \theta m'(\theta)/m(\theta)$. *Then,*

$$\frac{d\theta}{dT} = \frac{\theta}{\alpha}\frac{1}{q+\delta},\ \frac{d\theta}{du} = -\frac{\theta}{\alpha}\lambda,\text{ and } \frac{d\theta}{dq} = -\frac{\theta}{\alpha}\frac{1}{u(1-u)}$$

Moreover, $\lim_{T\to 0}\theta = \infty$, $\lim_{T\to\infty}\theta = \theta_0 \equiv m^{-1}\left((q+\delta)(1-u)u^{-1}\right)$ *and* $\lim_{q\to\infty}\theta = \infty$.

From the Beveridge curve, $m(\theta) = (u^{-1}-1)(q+\delta)/H(T)$. Noting that $-\theta\eta'(\theta)/\eta(\theta) = 1-\alpha$, the above derivatives follow by direct calculation. The limits are then induced by the properties of the matching technology.

Step 2. *Let* u_0 *be such that* $\phi_1(u_0,\infty) = 0$; *let also* $T_1(u)$ *be the solution of* $\phi_1(u,T_1(u)) = 0$. *Then* $T_1'(u) < 0$ *and*

$$\lim_{u\to u_0} T_1(u) = +\infty,\ \lim_{u\to 1} T_1(u) = 0$$

Taking the derivative of ϕ_1 with respect to T, we get

$$\partial\phi_1/\partial T = -\gamma\frac{\eta'}{\eta^2}\frac{d\theta}{dT} > 0 \tag{41}$$

from step 1. As $\phi_1(u,0) = \infty$ and $\phi_1(u,\infty) < 0$ for all $u < u_0$, there exists a unique $T_1 \equiv T_1(u)$ such that $\phi_1(u,T_1) = 0$. Applying the implicit function theorem and using step 1 yields the claim.

Step 3. *Let $T_2(u)$ be the solution of $\phi_2(u,T_2(u)) = 0$. Then, $T_2'(u) > 0$ and*

$$\lim_{u\to 0} T_2(u) = 0,\ \lim_{u\to 1} T_2(u) = \lambda^{-1}(\delta)$$

Taking the derivative of ϕ_2 with respect to T, we get

$$[1+\kappa+2\kappa\ln u]\,u\lambda' < 0 \tag{42}$$

since H is strictly concave. As $\phi_2(u,0) = +\infty$ and $\phi_2(u,\infty) = -\delta[1-\kappa u] < 0$, there exists a unique $T_2 \equiv T_2(u)$ such that $\phi_2(u,T_2) = 0$. The result follows from the implicit function theorem.

Step 4. *Conclusion.* The fact there exists a unique equilibrium follows directly from step 2 and step 3. Moreover, $T_1'(u^*) < 0 < T_2'(u^*)$. ■

proposition 2 From the Lemma, there exists a unique equilibrium to which the implicit function theorem applies. Let $\mathbf{J}$ denote the Jacobian matrix of function $\Phi \equiv (\phi_1,\phi_2)$ evaluated in equilibrium.

$$\mathbf{J} = \begin{bmatrix} \partial\phi_1/\partial u & \partial\phi_1/\partial T \\ \partial\phi_2/\partial u & \partial\phi_2/\partial T \end{bmatrix} \tag{43}$$

where partial derivatives are computed by means of step 1 of the Lemma:

$$\partial\phi_1/\partial u \equiv -(1+\alpha-2\alpha u)\frac{\kappa}{\alpha}\frac{u}{1-u}\frac{y}{q+\delta} \tag{44}$$

$$\partial\phi_1/\partial T \equiv -\frac{1-\alpha}{\alpha}\kappa u^2\frac{y}{q+\delta}\lambda \tag{45}$$

$$\partial\phi_2/\partial u \equiv \frac{\delta}{u} + 2\kappa u\lambda\left[\frac{1}{1-u}\ln u + \frac{u}{(1-u)^2}\ln u + \frac{1}{1-u}\right] \tag{46}$$

$$\partial\phi_2/\partial T \equiv \delta[1-\kappa u]\frac{\lambda'}{\lambda} \tag{47}$$

The implicit function theorem then implies that for any parameter $x = \kappa$ or $x = q$,

$$\begin{bmatrix} du^*/dx \\ dT^*/dx \end{bmatrix} = -\mathbf{J}^{-1}\begin{bmatrix} \partial\phi_1/\partial x \\ \partial\phi_2/\partial x \end{bmatrix} \tag{48}$$

where

$$\begin{aligned}
\partial\phi_1/\partial\kappa &\equiv -u^2\frac{y}{q+\delta} &(49)\\
\partial\phi_1/\partial q &\equiv -\frac{\kappa}{\alpha}u^2\frac{y}{(q+\delta)^2} &(50)\\
\partial\phi_2/\partial\kappa &\equiv \left[1+2\frac{u}{1-u}\ln u\right]u\lambda+\delta u &(51)\\
\partial\phi_2/\partial q &\equiv 0 &(52)
\end{aligned}$$

and

$$\mathbf{J}^{-1}=\frac{1}{\det\mathbf{J}}\begin{bmatrix}\partial\phi_2/\partial T & -\partial\phi_1/\partial T\\ -\partial\phi_2/\partial u & \partial\phi_1/\partial u\end{bmatrix} \quad (53)$$

with $\det\mathbf{J}>0$ since $T_1'(u^*)<0<T_2'(u^*)$.

The facts that $du^*/dq>0$, and $dT^*/dq>0$ are then obvious. Moreover, u tends to 1 as q tends to infinity for all $T>0$ (i.e. even if T is actually $+\infty$). From (40), T^* then tends to $\lambda^{-1}(\delta)$. This shows point (i) of the proposition.

The fact that $du^*/d\kappa<0$ also comes immediately. Moreover

$$sign\left\{\frac{dT^*}{d\kappa}\right\}=sign\left\{\frac{\partial\phi_2}{\partial u}\frac{\partial\phi_1}{\partial\kappa}-\frac{\partial\phi_1}{\partial u}\frac{\partial\phi_2}{\partial\kappa}\right\} \quad (54)$$

Using equations (44) to (47), equations (49) and (51), multiplying the relevant expression by $\alpha\frac{1-u}{u}\frac{q+\delta}{y}$ and grouping terms yields that $dT^*/d\kappa$ has the sign of

$$\Delta(\kappa)\equiv\kappa-\alpha+(\kappa+\alpha-2\alpha\kappa)u+2\alpha\kappa(\kappa-1)u^2+2\kappa\frac{u}{1-u}\ln u\{1-\alpha-\alpha u+\alpha\kappa u\} \quad (55)$$

Point (ii) of the proposition follows from the fact that $\beta=1-\kappa$ and $\Delta(\kappa)=-\alpha.O(1-u)<0$ for κ sufficiently small, while $\Delta(1)=(1-\alpha)\left[1+u+2\frac{u}{1-u}\ln u\right]>0$ for all $u\in[0,1)$ ($=0$ for $u=1$). The limit properties are obvious, since β tends to 1 implies that u tends to 0 whatever $T>0$. ■

proposition 3 Optimal schooling solves $S'(T_i)-\sigma y=rS(T_i)$. Using (12), it comes:

$$\begin{aligned}
&\tfrac{w_0\mu_i}{(q+\delta+\mu_i)^2}\left\{\lambda_i-\delta-\tfrac{\delta}{q+\delta}\mu_i\right\}-\tfrac{\delta}{q+\delta}\sigma+\\
&\int_{w_0}^{w_{\max}}\tfrac{(1-F(w))\mu_i}{[q+\delta+(1-F(w))\mu_i]^2}\left\{\lambda_i-\delta-\tfrac{\delta}{q+\delta}(1-F(w))\mu_i\right\}dw=0
\end{aligned} \quad (56)$$

In symmetric equilibrium, $T_i = T$ for all individuals and the wage offer distribution is given by (21). Using the fact that the unemployment rate is worth $u = (q+\delta)/(q+\delta+Hm)$, we get (after some tedious algebra):

$$\left[1+\kappa+2\kappa\frac{u}{1-u}\ln u\right]u\lambda = \delta\left[1-\kappa u+\frac{\sigma}{1-u}\right] \tag{57}$$

The decentralized equilibrium is then composed of equations (25) and (57), while the social optimum is still defined by equations (26) and (27). Both systems coincide whenever $\beta = 1-\alpha^p$ and

$$\left[1+1-\alpha+2(1-\alpha)\frac{u^p}{1-u^p}\ln u^p\right]u^p\lambda^p - \delta\frac{\sigma}{1-u^p} = (1-\alpha^p)\,u^p\lambda^p \tag{58}$$

By mean of equation (57), we see that (58) is equivalent to $\sigma = \widehat{\sigma}$.

side calculations *(Not to be published).* This note is devoted to the derivation of the (OS) curve. We consider the most general case with a schooling fee $\Sigma \equiv \sigma y$. Individual schooling solves equation (56), which we reproduce now for convenience:

$$\begin{aligned}&\underbrace{\frac{w_0\mu_i}{(q+\delta+\mu_i)^2}\left\{\lambda_i-\delta-\frac{\delta}{q+\delta}\mu_i\right\}}_{A_1}+\underbrace{-\frac{\delta}{q+\delta}\sigma}_{A_2}\\&+\underbrace{\int_{w_0}^{w_{\max}}\frac{(1-F(w))\,\mu_i}{[q+\delta+(1-F(w))\,\mu_i]^2}\left\{\lambda_i-\delta-\frac{\delta}{q+\delta}(1-F(w))\,\mu_i\right\}dw}_{A_3} = 0\end{aligned} \tag{59}$$

In symmetric equilibrium, $T_i = T$ and the wage offer distribution is the following:

$$1-F(w) = \frac{-(q+\delta)+(q+\delta+Hm)\left(\frac{y-w}{y-w_0}\right)^{1/2}}{Hm} \tag{60}$$

We now compute the term A_3.

$$\begin{aligned}A_3 &= \int_{w_0}^{w_{\max}}\frac{-(q+\delta)+(q+\delta+Hm)\left(\frac{y-w}{y-w_0}\right)^2}{(q+\delta+Hm)^2\left(\frac{y-w}{y-w_0}\right)}\left\{\lambda-\frac{\delta}{q+\delta}(q+\delta+Hm)\left(\frac{y-w}{y-w_0}\right)^{1/2}\right\}\\&= -\frac{q+\delta}{(q+\delta+Hm)^2}\lambda\int_{w_0}^{w_{\max}}\left(\frac{y-w}{y-w_0}\right)^{-1}dw-\frac{\delta}{q+\delta}(w_{\max}-w_0)\\&+\frac{\delta}{q+\delta+Hm}\int_{w_0}^{w_{\max}}\left(\frac{y-w}{y-w_0}\right)^{-1/2}dw+\frac{\lambda}{q+\delta+Hm}\int_{w_0}^{w_{\max}}\left(\frac{y-w}{y-w_0}\right)^{-1/2}dw\end{aligned} \tag{61}$$

But,

$$\int_{w_0}^{w_{\max}}(y-w)^{-1/2}dw = -2(y-w_{\max})^{1/2}+2(y-w_0)^{1/2} \tag{62}$$

$$\int_{w_0}^{w_{\max}}(y-w)^{-1}dw = \ln\left[\frac{y-w_0}{y-w_{\max}}\right] \tag{63}$$

Then,

$$
\begin{aligned}
A_3 &= -\tfrac{q+\delta}{(q+\delta+Hm)^2}\lambda\left(y-w_0\right)\ln\tfrac{y-w_0}{y-w_{\max}} + 2\tfrac{\delta}{q+\delta+Hm}\left(y-w_0\right) \\
&-2\tfrac{\delta}{q+\delta+Hm}\left(y-w_0\right)^{1/2}\left(y-w_{\max}\right)^{1/2} + 2\tfrac{\lambda}{q+\delta+Hm}\left(y-w_0\right) \\
&-2\tfrac{\lambda}{q+\delta+Hm}\left(y-w_0\right)^{1/2}\left(y-w_{\max}\right)^{1/2} + \tfrac{\delta}{q+\delta}\left(y-w_{\max}\right) - \tfrac{\delta}{q+\delta}\left(y-w_0\right)
\end{aligned} \tag{64}
$$

Multiplying A_3 by $\frac{q+\delta+Hm}{y-w_0}$ and grouping term yields:

$$
\tfrac{q+\delta+Hm}{y-w_0}A_3 = -2\left[\left(\tfrac{y-w_{\max}}{y-w_0}\right)^{1/2}-1\right](\lambda+\delta) + \tfrac{q+\delta}{q+\delta+Hm}\lambda\ln\tfrac{y-w_{\max}}{y-w_0} + \delta\tfrac{q+\delta+Hm}{q+\delta}\left[\tfrac{y-w_{\max}}{y-w_0}-1\right] \tag{65}
$$

From the free entry equation (23) and the definition of the upper-bound of the wage distribution, we get

$$
\left(\frac{y-w_{\max}}{y-w_0}\right)^{1/2} = \frac{q+\delta}{q+\delta+Hm} \tag{66}
$$

Equation (65) together with (66) yields:

$$
\tfrac{q+\delta+Hm}{y-w_0}A_3 = 2\lambda\tfrac{Hm}{q+\delta+Hm} + 2\tfrac{q+\delta}{q+\delta+Hm}\lambda\ln\tfrac{q+\delta}{q+\delta+Hm} - \tfrac{\delta}{q+\delta}\tfrac{H^2m^2}{q+\delta+Hm} \tag{67}
$$

The (OS) curve results from:

$$
\frac{q+\delta+Hm}{y-w_0}\left(A_1+A_2+A_3\right)=0 \tag{68}
$$

Equations (59), (67) and (68) imply that

$$
\begin{aligned}
&\tfrac{w_0}{y-w_0}\tfrac{Hm}{q+\delta+Hm}\left\{\lambda-\delta-\tfrac{\delta}{q+\delta}Hm\right\} - \tfrac{\delta}{q+\delta}\sigma y\tfrac{q+\delta+Hm}{y-w_0} \\
&+2\lambda\tfrac{Hm}{q+\delta+Hm} + 2\left(q+\delta\right)\tfrac{\lambda}{q+\delta+Hm}\ln\tfrac{q+\delta}{q+\delta+Hm} - \tfrac{\delta}{q+\delta}\tfrac{H^2m^2}{q+\delta+Hm} \quad = 0
\end{aligned} \tag{69}
$$

Noting that $u=\frac{q+\delta}{q+\delta+Hm}$, dividing each term by $(1-u)$, and using the minimum wage policy $w_0=\beta y$, we get the (OS) curve:

$$
\left[1+\kappa+2\kappa\frac{u}{1-u}\ln u\right]u\lambda = \delta\left[1-\kappa u-\frac{\sigma}{1-u}\right] \tag{70}
$$

Equation (25) is obviously a special case of (70) with $\sigma=0$.

References

[1] Acemoglu, D.T., 1996. A microfoundation for social increasing returns to human capital accumulation. Quarterly Journal of Economics 61, 779-804

[2] Becker, G.S., 1964. Human capital: a theoretical and empirical analysis with special reference to education. Columbia University Press

[3] Burdett, K., Mortensen, D.T., 1998. Wage differentials, employer size, and unemployment. International Economic Review 39, 257-273

[4] Card, D., 1999. The causal effect of education on earnings. Handbook of Labor Economics 3, 1801-1864

[5] Devine, T.J., Kiefer, N.M., 1991. Empirical labor Economics: the search approach. New York, Oxford University Press

[6] Drèze, J H., Sneessens, H., 1997. Technological development, competition from low-wage economies and low-skilled unemployment. In D.J. Snower and G. de la Dehesa, eds., Unemployment policy: Government options for the labour market, Cambridge; New York and Melbourne: Cambridge University Press

[7] Galor, O., Sicherman, N., 1990. A theory of career mobility. Journal of Political Economy 98, 169-192

[8] Hosios, A., 1990. On the efficiency of matching and related models of search and unemployment. Review of Economic Studies 57, 279-298

[9] Laing, D., Palivos, T., Wang, P., 1995. Learning, matching and growth. Review of Economic Studies 62, 115-129

[10] Layard, R., Nickell, S., 1999. Labor market institutions and economic performance. Handbook of Labor Economics 3C, 3029-3084

[11] Manning, A., 2000. Movin'on up: interpreting the earnings-experience profile. Bulletin of Economic Research 52, 261-295

[12] Moen, E.R., 1999. Education, ranking and competition for jobs. Journal of Labor Economics 17, 694-723

[13] Mortensen, D.T., 2000. Equilibrium unemployment with wage-posting: Burdett-Mortensen meet Pissarides. In H. Bunzel, B.J. Christiansen, P. Jensen, N.M. Kiefer and D.T. Mortensen, eds., Panel data and structural labor market models, Amsterdam: Elsevier

[14] Pissarides, C., 1990. Equilibrium unemployment theory. Basil Blackwell

[15] Saint-Paul, G., 1996. Dual labor market: a macroeconomic perspective. MIT Press

[16] Sicherman, N., 1990. Education and occupational mobility. Economics of Education Review 9, 163-179

[17] Thurow, L.C., 1975. Generating Inequality. Basic Books, New York

[18] Topel, R., Ward, M., 1992. Job mobility and the career of young men. Quarterly Journal of Economics 107, 441-479

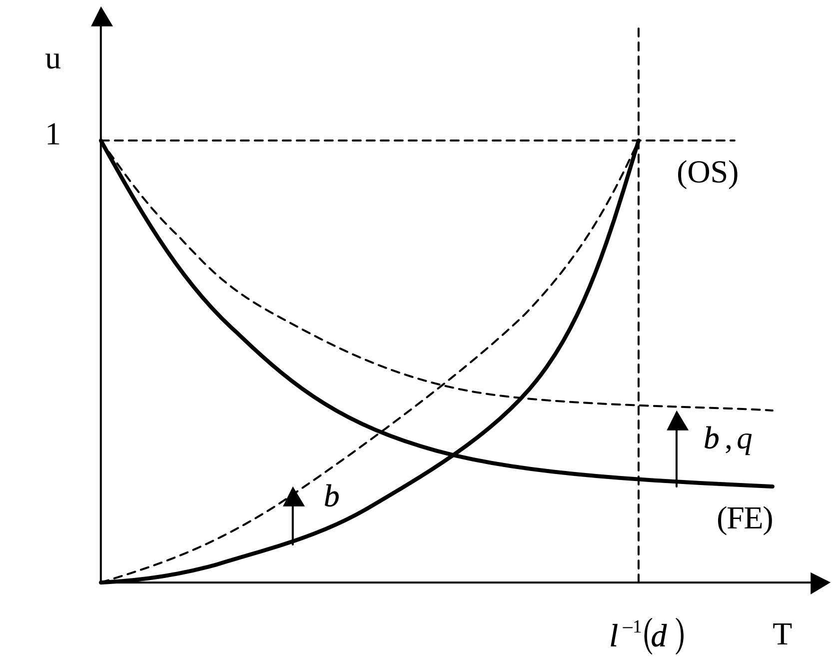

Fig.1 : Comparative statics

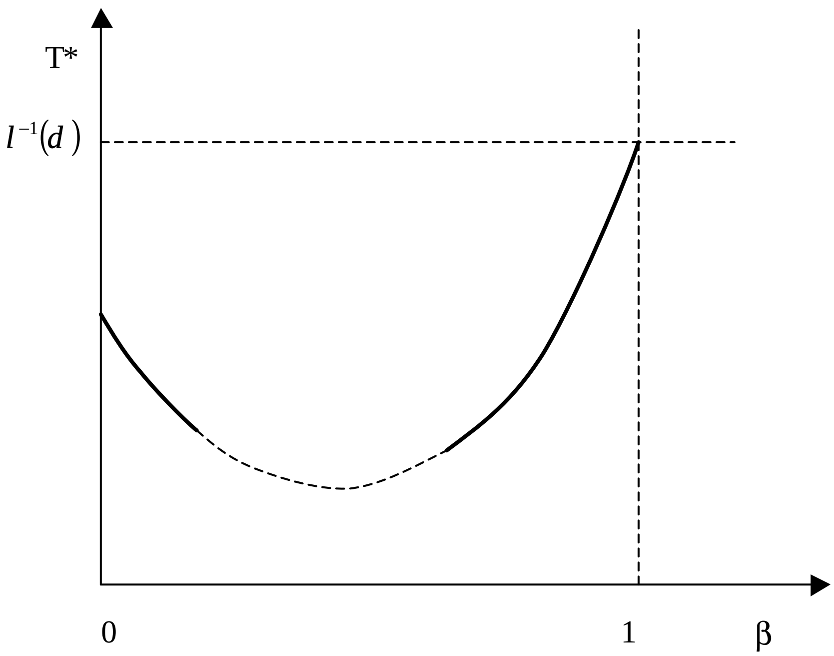

Fig.2 : Education and the minimum wage