Education, mobility and employers’ monopsony power: a search-theoretic analysis*

Bruno DECREUSE† Pierre GRANIER‡
GREQAM and Université de la Méditerranée

Abstract: We analyse the efficiency of schooling choices in a wage-posting search equilibrium model with on-the-job search. The workers have multidimensional skills and the search market is segmented by technology. Education determines the scope – or adaptability – of individual skills. Individuals get schooling to leave unemployment more quickly and to climb the wage ladder rapidly through job-to-job mobility – that is, to speed up job-shopping. Education reduces firms’ monopsony power in the wage determination by improving workers’ mobility. As a result, the wage distribution shifts rightward with aggregate schooling. However, the ratio of vacant jobs to job-seekers also falls in each sector. Either one or the other externality may dominate, implying, respectively, under- or over-education. A combination of minimum wage and schooling fee can decentralize the efficient allocation.

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

J.E.L. classification: J38, J64

---

*This paper has benefited from comments of seminar participants at EUREQua, GREQAM, the AFSE workshop of May 2002 in Lyon, the PAI-UAP workshop of June 2002 in Louvain-la-Neuve, and the EEA congress of August 2002 in Venice. We especially wish to thank, without implication, Pierre Cahuc, Olivier Charlot, Fabien Postel-Vinay, and Henri Sneessens.

†GREQAM, 2 rue de la charité, 13236 Marseille cedex 02, France. E-mail: decreuse@ehess.univ-mrs.fr
‡GREQAM, 2 rue de la charité, 13236 Marseille cedex 02, France. E-mail: granier@ehess.univ-mrs.fr
1 Introduction

There has been a strong revival in recent years of the notion that employers have some monopsony power in wage setting (see Bhaskar, Manning, and To, 2002, for a survey). Employers’ market power has two main causes: mobility costs and search frictions. On the one hand, workers have heterogenous preferences over the different jobs. This originates direct as well as opportunity costs to mobility, which are likely to reduce the labor supply elasticity. On the other hand, the information on available jobs is imperfect. This creates frictions in the allocation of workers to jobs, slowing down the mobility process. A major source of search frictions lies in the heterogeneity of the technologies used by firms, and in the corresponding skill requirements. Hence, the time needed to find a job suited to one’s characteristics is likely to increase with the diversity of skill requirements; so does employers’ market power.

One may however object that workers are seemingly passive in this approach. In this respect, our point is that individuals get an education before they enter the labour market, and education is likely to increases one’s technological mobility by improving the scope of a worker’s skills. Improved mobility means that higher educated workers find more easily a job while trying to contact alternative employers. As a result, firms facing a more educated workforce experience stronger limits to the use of their market power. This paper is devoted to a formalization of this argument.

A large body of empirical evidence suggests that schooling affects the mobility process on the labour market. On the one hand, unemployment rates decrease with education attainment in almost all OECD countries (see Drèze and Sneessens, 1997, and Layard and Nickell, 1999). Moreover, estimates from duration models generally report that workers with higher levels of formal education move into employment faster than workers with less education (see Devine and Kiefer, 1991, for a survey). On the other hand, the positive relationship between education and wages is indeed one of the most certain in Economics. Besides, a large part of wage growth over one’s career is due to occupational mobility. Empirical studies indicate that for young workers, job switching is a major component of life-cycle wage growth. Topel and Ward (1992) estimate this accounts for one third of the wage growth that occurs during the first decade of work experience. Taken together, these two findings suggest higher-educated workers benefit from better wages because they move faster on the wage ladder. The empirical results obtained by Galor and Sicherman (1990) and Sicherman (1990) confirm this view. They show that given an origin occupation, schooling increases the likelihood of occupational upgrading, either within or between firms.

In this paper, we develop the idea according to which the workers climb in the wage
ladder through job-shopping, while education improves workers’ mobility by making the workforce more adaptable to different technologies. Education, therefore, gives the workers some control over the speed of the job-shopping process. As a result, firms’ monopsony power decreases with education at the aggregate level. We focus on a set of three questions: Firstly, how do job creation and wage distribution respond to changes in the schooling level of the workforce? Secondly, what are the determinants of education? Thirdly, is the private return to schooling equal to, bigger or smaller than the social return?

The main assumption is that education provides adaptability since it improves the scope of an individual’s skills. Jobs differ in some technological dimension and in their corresponding skill requirement. Higher educated workers are more adaptable to various occupations requiring different knowledge and can apply on a wider range of jobs. We devote the full body of section 2 to an exposition of the arguments in favour of such a relationship between education and adaptability.

Charlot, Decreuse and Granier (2003) – hereafter CDG – embed this notion of education-induced adaptability in a multi-sector matching model with rent-sharing. They show that the incentives to schooling increase with the incidence and duration of unemployment spells. Moreover, some of the skills acquired through education are used to raise workers’ outside opportunities during the wage bargain. This originates over-education. However, this model abstracts from on-the-job search, and there is no job-to-job mobility. In addition, wages result from a situation of bilateral monopoly.

In the paper considered here, we build on Burdett and Mortensen (1998) – BM –, extended by Mortensen (2000). As in BM, workers search on the job, and firms set wages to attract applicants and reduce job turnover. As a consequence, an endogenous wage distribution emerges reflecting heterogeneity in workers’ status. We incorporate educational choices aimed at raising one’s mobility. We use this structure to answer the set of questions marked above.

1. **Unemployment and the wage dispersion provide incentives to schooling.** Education improves workers’ mobility and yields two types of returns. Status-mobility returns reflect the effect of education on one’s chance of leaving unemployment. Wage-mobility returns result from the effect of education on one’s chance of climbing the wage ladder. Importantly, both the cause and the main consequence of employers’ monopsony power provide incentives to schooling. The cause is unemployment, which raises the status-mobility returns to schooling. The individual schooling length is thus increasing in the two main determinants of unemployment: its incidence and duration. The consequence is wage dispersion, which increases the wage-mobility returns to schooling. Indeed, learning to climb the wage ladder is not very exciting when the ladder has only a few rungs.

2. **Education reduces unemployment and firms’ monopsony power, and lowers the
(sector-specific) tightness. More mobile workers benefit from a higher rate of contact while searching. They are less likely to be found unemployed or on the lower rungs of the wage ladder; they are also more likely to quit at given wage. Both effects act so as to reduce employers’ monopsony power. Firms are compelled to set higher wages: to attract more applicants, and to reduce job turnover. The wage offer distribution moves rightwards as a result. This is not so because education-induced human capital boosts the labour demand: education actually discourages firms’ entry on each sector. As expected profits are lower, the number of vacancies per applicant – the so-called (sector-specific) market tightness – decreases in each sector, even if higher educated workers benefit from better job opportunities.

3. Either under- or over-education may take place. We assume a social planner aiming at maximising the stationary output (net of search costs) chooses the schooling duration conditional on the decentralized labour market outcome – that is, conditional on the level of turnover the original Mortensen model predicts. Since education originates two conflicting externalities – it improves the wage offer distribution, but reduces the (sector-specific) market tightness – the social return to schooling may be either lower or larger than the private return to schooling. The positive externality is dominated when the minimum wage is sufficiently large and/or the elasticity of the matching technology with respect to vacancies is sufficiently low. Hence, the market incentives to schooling are too large. The inefficiencies due to schooling add to the inefficiency induced by poaching as in the original Mortensen model. We show that a combination of minimum wage and schooling fee can decentralize the efficient allocation.

Our main assumptions and results find some echoes in the literature.

The idea that education improves the scope of a worker’s skills has been first considered by Nelson and Phelps (1966). It has been recently developed by Aghion, Howitt and Violante (2002), and Kumar and Krueger (2002) in growth models. The other idea according to which occupational mobility drives wage mobility and can be related to schooling has already been suggested by Galor and Sicherman (1990). Our paper connects these two ideas and embed them in a search-theoretic framework.

Our paper is also related to the literature exploring the nexus between education and labour market performance within the matching theory. Several models in this field stress two main predictions. Firstly, unemployment implies that human capital remains idle over long/frequent spells. Therefore, it is detrimental to schooling investments (see e.g.

\[\text{We do so because the original Burdett-Mortensen model predicts excess turnover. The efficient level of turnover is zero in this model, because working for a poaching firm does not generate more wealth. However, the allocation without on-the-job search is a nonsense in the decentralized economy: the wage distribution collapses to the minimum wage.}\]
Laing, Palivos and Wang, 1995, Burdett and Smith, 2002). Secondly, 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 where a simultaneous rise in the duration of studies and in unemployment has been observed during the 70's the 80's. In Moen, the workers are ranked in a job queue. Investing in education improves one’s chance of being employed, but at the expense of the others. Over-education may result\(^2\). In our paper, investing in education improves employment perspectives, not because the higher educated benefit from a better ranking in the job queue, but because they benefit from a larger contact rate.

The outline of the paper is as follows. Section 2 argues in favour of the adaptability view of education. Section 3 describes the structure of our model. Section 4 examines the effects of schooling on unemployment, labour demand and wage distribution. Section 5 takes into account the endogenous determination of education. In section 6, we compare private and social returns to schooling and propose a policy to restore efficiency. Section 7 concludes.

2 The adaptability view of education

In this paper, education enhances one’s ability to benefit from job-shopping: the higher educated workers have access to more jobs because they are able to perform on more technologies. We refer to this effect as the adaptability effect of education. This section is devoted to this approach.

A preliminary clarification. By stating the higher educated are able to perform on more technologies, we do not necessarily mean individuals directly learn different technologies at school. The underlying idea is that general education enhances one’s capacity of learning. This capacity offers the possibility to perform many different types of tasks, allowing one to switch from one activity to another over the working life. Think for instance that students acquire a number of basic abilities in various fields during their schooling, the number of which increases with schooling duration. Jobs in the economy differ in their embodied technology and, therefore, in their skill requirement. A technology is defined as a particular combination of abilities and a worker can apply for a given job only if she

---

\(^2\) Dual matching models provide other examples of over-education. In Saint-Paul (1996), there may be some excess supply of skills as a rise in the number of educated alters firms’ incentives to allocate their vacancies between sectors, and causes a rise in unemployment for both educated and uneducated workers. In Charlot and Decreuse (2003), self-selection in education leads to composition effects which can be detrimental to job creation in each market segment.
embodies the required combination of abilities. As a higher educated embodies a larger number of abilities, she can combine them in a larger number of ways, and, therefore, can operate on a larger number of technologies.

What is adaptability? This view of education refers to the more general notion of adaptability, which complements the human capital theory initiated by Becker (1964). Actually, many economists agree with the idea that education must be viewed as an investment but not as a separate factor of production. They also embrace the principle that education equips a man to perform certain jobs. Since at least Nelson and Phelps (1966), it has been recognized that the direct contribution of education to physical production does not account for the total contribution to revenue and this is probably so because education improves workers’ adaptability. Nelson and Phelps indeed point out two essential dimensions of education: the ability to adapt to technological change and the ability to innovate. This view has received some support at the micro level, either from studies of US industries (Bartel and Lichtenberg, 1987), or plants (Dunne and Schmitz, 1995, Doms, Dunne and Troske, 1997).

Nelson and Phelps’ notion of adaptability has recently been pursued by Krueger and Kumar (2002) and Aghion, Howitt and Violante (2002) in growth models. The former assume that general (by opposition to vocational) education reduces the probability the workers suffer a loss of task-specific productivity following the introduction of a new technology. The latter assume new technologies are sector-specific and workers must be adaptable to implement any such new technology. The longer the schooling period when young, the more likely it is that the worker can use the new technology when old.

In our paper, educated workers are more adaptable not because they can perform better on new jobs but because they can form matches with more firms. A close assumption is used by LLoyd-Ellis (1999) in a growth model with endogenous technological change. In his model, minimum skill levels are required to implement new but equiproductive technologies and workers differ in the range of technologies they can implement. However, the skill acquisition process is exogenous. Our paper departs from these previous studies by introducing labour market imperfections, and by focusing on efficiency aspects instead of positive aspects.

Education and occupational upgrading. The more specific idea according to which education improves one’s ability to climb the wage ladder because of enhanced occupational mobility has already been suggested by Galor and Sicherman (1990). In their paper,

---

3See for instance the influential paper of Benhabib and Spiegel (1994). Considering a sample of 78 countries they conclude that “the role of human capital is indeed one of facilitating adoption of technology from abroad and creation of appropriate domestic technologies rather than entering on its own as a factor of production”.

---

6
occupational mobility is due to career mobility defined as mobility across a number of occupations within or across firms. They argue that a large part of the returns to schooling consists in a higher probability to reach high wage occupations. In this paper, we associate occupational mobility to a stochastic process of job-shopping. Empirical studies by Topel and Ward (1992) or Keith and McWilliams (1999) are in accordance. They find strong evidence in favor of job-shopping as a key determinant of occupational mobility. It suggests that the wage-mobility returns to schooling can be interpreted as returns to search capital: education provides wage-mobility returns because it raises search efficiency. The notion of adaptability provides a rationale for this view.

To simplify, expanding the type of jobs to which a given worker can apply is the only reason for a positive expected wage return to schooling in this model. Of course this reason is not the only explanation; rather, it should be seen as complement and not as substitute to the more traditional view based on the human capital theory. There is a large body of evidence that more experienced workers and/or persons with longer job tenures suffer persistent wage losses after displacement (Jacobson, Lalonde and Sullivan, 1993, Topel, 1991). These findings are commonly interpreted as an empirical support to (specific) human capital theory (Topel, 1991). However, these findings are also compatible with a simple on-the-job search framework (see e.g. Burdett, 1978, Manning, 2000).

3 The model

The demographics of the model has two main characteristics. First, $\delta$ 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 $\delta$ 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 pure rate of time preference $\rho$ tends to 0. Individuals are risk-neutral.

Production sector. There is a single final good which can be produced by means of a continuum of equiproducive technologies. The output flow produced by each technology is $y$ and the set of technologies is normalized to $[0,1]$. Advertising a vacancy implies posting a wage along with a given technology. Firms choose the wage level, and draw a technology from the uniform law on $[0,1]$. This process is irreversible, so that neither the wage nor the technology can be modified once a job offer is posted. We assume that ads for jobs convey sufficient information about the type of skills needed to perform the job.

---

4The reason why such technologies coexist is beyond the scope of this paper. One may think of intermediary goods, which are complementary in the final good production. Equivalently, there may be a continuum of final goods that are complementary in preferences.
Hence, the search market is segmented by technology and only those having acquired the technological knowledge to operate a given technology can apply for such jobs. Crucially, the number of technologies which can be operated 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 $H(T_i) \subset [0, 1]$ submarkets for which she has the required skills, but not on the remaining $1 - H(T_i)$ submarkets.

The technology $H$ is strictly increasing 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 to schooling. The concavity of function $H$ implies that the rate of return is decreasing in the schooling duration.

**Matching sector.** All the agents who have completed their schooling participate to the search process with equal intensity. On each (sub-)market $j$, the flow of matches $M_j$ between job-seekers and vacancies is given by the matching technology:

$$M_j = M(x_j, v_j)$$

where $x_j$ and $v_j$ are (resp.) the number of job-seekers and vacancies on market $j$. The function $M$ features 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: $M(x_1, 0) = M(0, x_2) = 0$ for all $x_1, x_2 \geq 0$, and the Inada conditions. We shall denote by $\alpha \equiv \alpha(\theta) \equiv \theta m'(\theta)/m(\theta)$ the elasticity of the matching technology with respect to the number of vacancies.

Matches are equiprobably distributed among the job-seekers, as well as among the vacancies. Denoting by $\theta_j \equiv v_j/x_j$ the market-specific tightness, the flow probability for a worker to locate a vacancy on market $j$ is $m(\theta_j) \equiv M(1, \theta_j)$. Symmetrically, the flow probability for a firm to meet an applicant is $\eta_j = m(\theta_j)/\theta_j$. The individual job-finding rate is then

$$\mu_i = \int_{j \in H(T_i)} m(\theta_j) d\theta$$

The law of large numbers implies $v_j = v$ and $x_j = \overline{H} N$, where $N$ is the workforce and $\overline{H} \equiv \mathbb{E}[H(T_i)]$ is the number of technologies which can be implemented on average by a “typical” worker. Consequently, market-specific tightness is $\theta_j = \theta$, contact rates are $\eta_j = \eta(\theta)$ and

$$\mu_i = H(T_i) m(\theta)$$
It is increasing in individual schooling duration. Mobility returns to schooling will therefore consist in a higher contact rate.

**Firms’ and workers’ gains.** Let $S(T_i)$ and $W(T_i, w)$ 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 minimum wage. The values $S(T_i)$ and $W(T_i, w)$ satisfy the two arbitrage equations:\(^5\)

\[
rS(T_i) = \mu_i \int_{w_0}^{w_{\text{max}}} [W(T_i, w) - S(T_i)] dF(w) \tag{4}
\]
\[
(r + q) W(T_i, w) = w + \mu_i \int_{w}^{w_{\text{max}}} [W(T_i, \tilde{w}) - W(T_i, w)] dF(\tilde{w}) + qS(T_i) \tag{5}
\]

where $F(w)$ is the wage offer distribution, and $w_{\text{max}}$ is the (endogenous) upper-bound of the wage distribution. We adopt the following minimum wage policy: $w_0 \equiv \beta y$, $\beta \in [0, 1]$. The parameter $\beta$ will be referred to as the minimum wage factor. We shall also note $\kappa \equiv 1 - \beta \in (0, 1]$. With (flow) probability $\mu_i$, 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 (flow) probability $q$ the match is destroyed and the worker steps into unemployment where she 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:

\[
pV = \max_{w \geq w_0} (-\gamma + \eta [\mathbb{E} T_i p(w \mid T_i) (J(T_i, w) - V)]) \tag{6}
\]
\[
(r + q) J(T_i, w) = y - w - \mu_i [1 - F(w)] [J(T_i, w) - V] + (q + \delta) V \tag{7}
\]

Vacant job slots induce a flow cost of posting a vacancy $\gamma > 0$. With (flow) probability $\eta$, 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 status 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 $p(w \mid T_i) = u(T_i) + (1 - u(T_i)) G(w \mid T_i)$, where $G(\cdot \mid T_i)$ is the distribution of wages among the employees of schooling duration $T_i$. Since the education attainment of the incoming worker is $a \text{ priori}$ unknown to the firm, the value of a filled job has to be taken with an expectation operator in equation (6). Filled jobs bring a revenue to the

---

\(^5\)Equation (4) 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.
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.

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

$$\mathbb{E}_{T_i} p(w \mid T_i) J(T_i, w) = \gamma/\eta(\theta)$$

(8)

From equations (4) and (7), we get

$$S(T_i) = \frac{\mu_i}{r + \mu_i} \int_{w_0}^{w_{\text{max}}} W(T_i, w) dF(w)$$

(9)

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

(10)

The value of a filled job decreases with the schooling duration of the worker, for she is more likely to quit to get a better wage.

4 Schooling reduces firms’ monopsony power

The goal of this section is to capture the effects of schooling on the supply side. Education essentially reduces the magnitude of the matching frictions faced by the workers. It lowers unemployment, and firms’ monopsony power. As a result, firms set higher wages, but (sector-specific) market tightness is reduced.

In the remainder of this section, $T$ denotes the schooling duration common to all the workers.

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 unemployment are:

$$\frac{dN}{dt} = \delta e^{-\delta T} - \delta N$$

(11)

$$\frac{dU}{dt} = \delta e^{-\delta T} + q (N - U) - (H(T) m(\theta) + \delta) U$$

(12)

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 the unemployment rate $u$ is given by the following ‘Beveridge curve’:

$$u = \frac{\delta + q}{\delta + q + H(T) m(\theta)}$$

(13)

Employment is residually defined by $L \equiv N - U$.

Wage distributions. The wage distributions $F$ and $G$ do 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

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

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 is \( \mu w e^{-\delta T} F(w) \): the job-finding rate times unemployment times the proportion of wage offers below \( w \). Equating these two flows and making use of the Beveridge curve \((13)\) yields

\[
G(w) = \frac{uF(w)}{u + (1-u)[1-F(w)]} 
\]

The probability that a given worker will accept the job offer is thus

\[
p(w) = \frac{uF(w)}{u + (1-u)[1-F(w)]} \]

The \emph{wage offer distribution} \( F(w) \) is then computed from the following free entry condition:

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

Using \((10)\) and solving for \( F(w) \), we get

\[
1 - F(w) = \frac{-2q+\gamma}{2Hm} \left( \frac{2q+\gamma+\delta}{2Hm} \right)^{1/2} 
\]

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

\[
w_{\max} = y - (r + q) \frac{\gamma}{\eta} 
\]

by virtue of \( F(w_{\max}) = 1 \). Finally, tightness can be derived from equation \((16)\) evaluated in \( w = w_0 \):

\[
\frac{\gamma}{\eta(\theta)} = \frac{\delta + q}{\delta + q + Hm(\theta)} \frac{y - w_0}{r + q + Hm(\theta)} 
\]

\emph{Equilibrium.} In the remainder, we only focus on the case where the rate of time preference \( \rho \) tends to 0. The supply side can be summarized by the following equations:

\[
G(w) = \frac{uF(w)}{u + (1-u)[1-F(w)]} 
\]

\[
F(w) = \frac{1 - \frac{y-w}{\kappa y}}{1 - u} 
\]

\[
w_{\max} = y \left( 1 - \kappa u^2 \right) 
\]

\[
\frac{\gamma}{\eta(\theta)} = \kappa u^2 \frac{y}{q + \delta} 
\]

where \( u \) is given by \((13)\). The three first equations characterize the wage distribution; the fourth determines (sector-specific) tightness. The equilibrium with exogenous schooling
has a recursive structure. Tightness and unemployment can be solved jointly from (13) and (23). The offer distribution, and then the wage distribution follow.

From the Beveridge curve, the unemployment rate can be solved as a decreasing function $u$ of tightness and schooling duration; alternatively, tightness can be solved as a decreasing function $\theta$ of unemployment rate and schooling duration. Formally,

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

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

Either one or the other of these functions can be used in equation (23) to characterize the equilibrium tightness, or unemployment rate.\(^6\)

**Proposition 1: Education, Wages and Employment**

Equation (13), (21), (22), (23) defines a unique positive vector $(\bar{w}_{\text{max}}, \bar{\theta}, \bar{u})$, and two functions $\bar{F} : [\beta y, w_{\text{max}}] \rightarrow [0, 1]$ and $\bar{G} : [\beta y, w_{\text{max}}] \rightarrow [0, 1]$.

Let $\bar{\theta} \equiv \bar{\theta}(T)$, $\bar{u} \equiv \bar{u}(T)$, $\bar{F}(w) \equiv \bar{F}(w, T)$, $\bar{G}(w) \equiv \bar{G}(w, T)$. For all $T > 0$,

(i) $\bar{w} (T) < 0$

(ii) $\bar{w}'_{\text{max}} (T) > 0$, $\bar{F}_T (w, T) < 0$ and $\bar{G}_T (w, T) < 0$

(iii) $\bar{\theta} (T) < 0$

By making the workers more adaptable, education accelerates the job-shopping process. The appropriate measure of the speed of the process is the rate of contact $H(T) m(\bar{\theta}(T))$, which increases with $T$. This has three major implications, summarized by points (i) to (iii) in the proposition.

(i) The unemployment rate is strictly decreasing in the education level of the workforce, revealing the higher job-finding rate of the workers.

(ii) The fact that workers are more mobile weakens firms’ monopsony power. Firms are compelled to offer higher wages, in order to increase the probability to meet a worker willing to accept the offer, and to reduce the probability the worker quits. This means that a highly educated workforce benefits from better wage offers and $\bar{F}_T (w, T) < 0$. Consequently, the wage distribution $\bar{G}$ shifts rightwards, reflecting both the shift of the wage offer distribution, and increased upward mobility in the distribution. Figure 1 depicts two wage offer distributions, corresponding to two schooling durations. The wage

---

\(^6\)Each solving method has its own merit. Solving in tightness allows to shed light on the fact that equilibrium tightness is decreasing in education, an externality called to an important role in section 6. Solving in unemployment rate immediately shows that equilibrium unemployment rate is decreasing in schooling duration; this will guarantee uniqueness in section 5.
offer distribution associated with the shorter schooling duration first-order stochastically dominates the other.

(iii) Increased adaptability reduces firms’ profits: firms are obliged to pay higher wages, and job quits are more frequent. This discourages firms’ entry and the ratio of vacancies to job-seekers falls in each sector. Another way of stating this point is the following: the elasticity $\varepsilon_{Hm}$ of the contact rate $Hm$ wrt $T$ is lower than the elasticity $\varepsilon_H$ of the learning technology $H$ wrt $T$. Formally,

$$\varepsilon_{Hm} = \frac{T d(Hm) / dT}{Hm} = T \lambda(T) + \alpha(\bar{\theta})\frac{T \bar{\theta}'(T)}{\bar{\theta}} < T \lambda(T) \equiv \varepsilon_H$$  \hspace{1cm} (26)

5 Accounting for the schooling incentives

In this section, we consider the schooling choice. Firstly, we focus on the incentives to schooling at given tightness and wage offer distribution. The respective roles of unemployment and wage dispersion are then emphasized. Secondly, we analyse the equilibrium with endogenous schooling.

The value of search. Each newborn individual sets individually her schooling duration to maximize her intertemporal discounted utility. The wage offer distribution $F$ is common to all workers – irrespective of their schooling duration. By contrast, the worker can alter her wage distribution $G_i(w) \equiv G(w \mid T_i)$ by changing the rate at which she receives job offers, either when she is employed or unemployed. Following the previous section, $G_i$ is defined by

$$G_i(w) = \frac{u_iF(w)}{u_i + (1 - u_i)[1 - F(w)]}$$ \hspace{1cm} (27)

where

$$u_i = \frac{\delta + q}{\delta + q + H(T_i)m(\theta)}$$ \hspace{1cm} (28)

is the rate of unemployment among individuals whose schooling duration is $T_i$. The wage distribution moves rightwards with $T_i$. The average wage of individuals whose education is $T_i$ is

$$\bar{w}_i \equiv \int_{w_0}^{w_{\text{max}}} w dG_i(w) = w_0 + \int_{w_0}^{w_{\text{max}}} [1 - G_i(w)] dw$$ \hspace{1cm} (29)

From stochastic dominance, it is strictly increasing in $T_i$. When the rate of time preference tends to 0, the value of search is worth

$$S(T_i) = (1 - u_i) \frac{\bar{w}_i}{\delta}$$ \hspace{1cm} (30)
The optimal schooling duration. Optimal schooling results from

$$\hat{T}_i \in \arg \max_{T_i \geq 0} e^{-\delta T_i} S(T_i)$$

(31)

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

$$S'(T_i) = - \frac{\partial u_i \bar{w}_i}{\partial T_i \delta} + \frac{1 - u_i \partial \bar{w}_i}{\delta \partial T_i}$$

(32)

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 (or, equivalently, the wage distribution the worker is facing).

**Proposition 2** A CHARACTERIZATION OF THE OPTIMAL SCHOOLING DURATION

The optimal schooling duration solves

$$w_0 (\lambda_i u_i - \delta) + \int_{w_0}^{w_{\text{max}}} [1 - G_i(w)] \left[ \lambda_i \frac{G_i(w)}{F(w)} - \delta \right] dw = 0$$

(33)

It has the following properties:

(i) $\delta u_i < \lambda_i u_i < \delta$

(ii) $\lambda_i$ tends to $\delta$ as $m(\theta)$ tends to 0, or $q$ tends to infinity; $\lambda_i$ tends to infinity as $m(\theta)$ tends to infinity

(iii) $\lambda_i u_i = \delta$ when the wage distribution is degenerate, i.e. when $w_0 = w_{\text{max}}$

The three properties we have highlighted concern the magnitude of education investment, the relationship of the schooling duration with the different components of unemployment (unemployment incidence and duration), and the impact of the wage dispersion. Property (i) shows that despite the apparent complexity of equation (33), the schooling duration remains bounded.

Property (ii) puts the emphasis on the effects of incidence and duration of unemployment on the incentives to schooling. It is difficult to exhibit analytically a systematic link\(^7\); however limit properties offer some intuitions. As education is aimed at improving workers’ mobility, it become useless in the perfectly competitive environment where the

\(^7\)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.
contact rate between workers and job offers tends to infinity (i.e. when tightness becomes very high). Conversely, the return to schooling is maximal when the matching process is so inefficient that it is virtually impossible to get a job\(^8\) (i.e. when tightness is very low), or when the job loss rate is arbitrarily large\(^9\).

Property (iii) shows that the optimal duration of schooling is minimal when the wage distribution collapses to \(w_0 = w_{\text{max}}\). The reason is that there is no wage-mobility return to schooling in this case. This suggests that in most plausible cases, the individual schooling duration should be increasing in the dispersion of the wage offer distribution.

**Equilibrium.** In equilibrium, the wage offer distribution is characterized by (21) and (22). Replacing in (33), and using (13), we get (after some algebra):

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

This offers a first relationship between education and unemployment rate. The free entry condition (23) and the Beveridge curve (13) together offer a second relationship:

\[
\gamma/\eta(\theta(u, T)) = \kappa u^2 y/(q + \delta)
\]

Equilibrium is defined by the intersection of these two curves: the optimal schooling equation (OS), and the free entry equation (FE).

**Proposition 3** There exists a unique equilibrium \((u^*, T^*)\).

(FE) defines the unemployment rate as a strictly decreasing function of the schooling duration. This negative relationship between unemployment and schooling reveals the fact that schooling improves the job-finding rate. Note, however, that it does not mean education raises the labour demand in each sector – these points are discussed in the previous section. (OS) defines the schooling duration as a strictly increasing function of the unemployment rate. According to proposition 2, education tends to increase with the components of unemployment. Yet the result cannot be established for any wage distribution. However, it is true in the case of the (labour market) equilibrium wage distribution. Therefore, education is strictly increasing in the unemployment rate. As (FE) and (OS) have opposite slopes, the equilibrium is unique.

\(^8\)Note that the maximisation problem is not continuous in \(m(\theta) = 0\), which implies \(S(T) = 0\) for all \(T\). Nevertheless, from the foc, \(\lambda(T)\) becomes arbitrarily close to \(\tau\) as \(m(\theta)\) tends to 0.

\(^9\)CDG obtain similar results in their model with ex-post rent-sharing. As far as one is willing to accept the adaptability view of education, it seems that unemployment tends to favourize education investments instead of discouraging them. This, of course, strongly departs from the basic search model with human capital.
Comparative statics. Two parameters are of great economic importance concerning job mobility: the job destruction rate – which determines how frequently status-mobility is needed – and the minimum wage (MW) – which determines the wage dispersion, and, therefore, the size of wage-mobility returns to schooling. The effects of these parameters are analyzed in proposition 4 and depicted by figure 2.

**Proposition 4** A characterization of equilibrium

There exist \((\beta_0, \beta_1), 0 < \beta_0 < \beta_1 < 1\), 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) \(du^*/d\beta > 0\) for all \(\beta \in [0, 1]\) 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 shifts upward. 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 \(\delta\).

[Figure 2]

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 shifts 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 raising status-mobility returns. 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\(^{10}\); this is depicted by figure 3.

[Figure 3]

Empirically, proposition 4 implies that two economies characterized by different degrees of wage rigidity (different \(\beta\)) feature different unemployment rates, but may well exhibit the same aggregate educational level. In the low MW country, employment is

---

\(^{10}\) 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.
higher and the wage distribution is 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\textsuperscript{11}.

We now turn to efficiency considerations.

6 Private and social mobility returns to schooling

This section is devoted to a comparison between social and private returns to schooling. Education originates two conflicting externalities (in terms of welfare): it compels firms to offer better wages, and it reduces the market tightness in each sector. Over- and under-education are, therefore, possible outputs. However, over-education is a more likely outcome than under-education. An adequate combination of minimum wage and schooling fee can decentralize the efficient allocation.

The second-best allocation. We consider the problem faced by a social planner who intends to set the schooling duration, taking into account firms’ willingness to enter the search market. Hence, we only focus on schooling-specific distortions, abstracting from other inefficiencies due to the inability of the wage-posting scheme to internalize search externalities. The objective of the social planner is standard, namely maximising the discounted value of the path of aggregate consumption (see Hosios, 1990, and Pissarides, 2000). When the rate of time preference $\rho$ tends to 0, maximising stationary consumption is equivalent to maximize the discounted value of future expected gains of a newborn individual. Therefore,

$$ T^p \in \arg \max_{T} e^{-\delta T} \left( 1 - u \right) \frac{\bar{w}}{\delta} $$

(34)

where $\bar{w} = (1 - \kappa u) y$. The maximisation is subject to $u \equiv \bar{u}(T) = u(\bar{\theta}(T), T)$, and $\theta \equiv \bar{\theta}(T)$, where $\bar{\theta}$ and $\bar{u}$ are defined in proposition 1, and $u$ is defined in equation (24). Like individuals, the planner considers status- and wage-mobility returns to schooling: education lowers the unemployment rate, and raises the speed of the job-shopping process. However, unlike individuals, the planner also takes into account induced changes in the stationary distribution of wage offers, as well as in the tightness common to all sectors.

The first-order condition to the maximisation problem is necessary; it yields:

$$ \delta \frac{1 - \kappa u^p}{u^p \lambda^p} = (1 - \alpha^p) \frac{1 - \kappa + 2\kappa (1 - u^p)}{1 - \alpha^p + 2\alpha^p (1 - u^p)} $$

(35)

\textsuperscript{11}Several papers have already claimed that the minimum wage may provide incentives to schooling (see e.g. Cahuc and Michel, 1996, or Agell and Lommerud, 1997) by raising employers’ reservation productivity. The mechanism is of course different in our paper.
This equation must be compared to its decentralized analogue:

$$\delta \frac{1 - \kappa u^*}{u^{*\lambda}} = 1 + \kappa + 2\kappa \frac{u^*}{1 - u^*} \ln u^*$$  \hspace{1cm} (36)

**Proposition 5** SOCIAL VERSUS PRIVATE RETURN TO SCHOOLING

(i) Both over- and under-education are possible outcomes

(ii) There is over-education if $\kappa \leq \alpha^p / (1 - \alpha^p)$

with $\alpha^p \equiv \alpha \left(\bar{\theta} (T^p)\right)$.

We make two points. Firstly, at a theoretical level, the private return to schooling may either be lower or larger than the social return. Secondly, over-education is a more likely outcome than under-education.

(i) By raising the speed of the job-shopping process, education reduces firms’ monopsony power. This implies two externalities. A positive externality according to which the longer the schooling duration, the better the wage offers. A negative externality according to which the longer the schooling duration, the lower the (sector-specific) market tightness. Since those externalities have opposite signs, it is quite intuitive that either one of them may dominate.

(ii) Over-education seems more plausible. Looking at condition (ii), there is over-education when the MW factor $\beta$ is higher than $(1 - 2\alpha^p) / (1 - \alpha^p)$. This is necessarily so whenever $\alpha^p \geq 1/2$, a rather consensual value. The reason why there is a condition over the minimum wage is obvious: the positive externality transits through higher wages. If the MW is “large” – positive is enough in most cases – the latter externality has a very small magnitude.

Over-education does not mean that the economic situation of the population would improve if individuals were shortening their schooling duration. Actually, unemployment would rise, and wages would decrease on average. But the decline would be modest, as the fall in job-finding rate due to decreased job mobility would be partially compensated by a higher tightness in each sector. As a result, individuals would spend much less time in the schooling system, while they would face a very small deterioration of their job prospects.

In CDG, the private incentives to schooling associated to the adaptability component of education are also inadequate. In their model, there is bilateral monopoly over wage setting. One of the workers’ incentives to schooling is to raise their outside options at the bargaining stage. Hence, schooling affects one’s wage, but does not alter the others’. The only externality is negative: tightness decreases in each sector, reflecting the fall in firms’ profitability. In our paper, firms have monopsony power in wage setting. Workers acquire skills to climb up faster in the wage distribution, but they do not realize that doing so, the distribution of wage offers faced by the others is altered. This originates a
positive externality that goes against the negative externality due to lower tightness in each sector.

Should the excess job turnover externality be considered seriously? There are two limits to the model: firms are precluded from making wage counter-offers, and jobs are not vertically differentiated. Indeed, excess job turnover is made possible by the fact firms do not react to wage offers received by their employees. If they were able to do so, they would raise the wage so as to meet the other’s proposal, up to job’s profitability. Such Bertrand competition between firms would of course reduce job turnover. However, a novel externality would result: workers would get skills through schooling to benefit from a higher chance to be contacted by another firm, and therefore receive a pay rise, though not quitting their job. Concerning the issue of vertical differentiation, it is clear that, in our model – as well as in the original BM model –, job-to-job transitions are essentially wasteful, because workers never improve the quality/productivity of their match by switching from job to job.

The “first-best” allocation. In this sub-section, we enlarge the planner’s set of controls. She now chooses both the schooling duration and the flow number of vacancies. As job-to-job transitions are socially wasteful since they raise congestion between job-seekers, the planner should normally forbid on-the-job search. But the resulting allocation would have nothing to do with the decentralized allocation, in which on-the-job search sustains the existence of a non-degenerate wage distribution. Therefore, we proceed as Mortensen (2000): the social planner takes as given that all workers search. The maximisation problem is

\[(T^*, \theta^*) \in \arg\max_{T, \theta} \left\{ e^{-\delta T} (1 - u) y - \gamma v \right\} \]  
}s.t. \[ u = u(\theta, T) \] and \[ \theta = v / (H(T) e^{-\delta T}) \].

The desired rate of unemployment \(u^*\) and schooling duration \(T^*\) solve the first-order conditions:

\[ \gamma \eta(\theta(u, T)) = \alpha u^2 y / (q + \delta) \]  
(38)

\[ (1 - \alpha) u \lambda(T) = \delta [1 - \alpha u] \]  
(39)

Equations (38) and (39) must be compared to equations (FE) and (OS), i.e. the market outcome.

It comes immediately that, in addition to the externalities induced by education, Mortensen’s (2000) result still holds. At given schooling duration, 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 a single firm cannot take into account. Hence, we can identify two distinct sources of inefficiency: the

\[ \text{See Postel-Vinay and Robin (2002) for the complete characterization of such a mechanism.} \]
one highlighted by Mortensen, according to which firms tend to create too many vacancies, and the novel externality, according to which workers are too mobile.

**Decentralizing first best.** There are two externalities, therefore two policy tools are required: the MW to affect the labour market, and a schooling fee to alter education behaviours. Let $\Sigma = \sigma y$ 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. Individual schooling results from:

$$\max_T \left( e^{-\delta T} S(T) - \int_0^T e^{-\delta z} \sigma y dz \right)$$

(40)

**Proposition 6 First best in the decentralized economy**

\[ \alpha^e = \alpha (\theta^e) \text{. The decentralized outcome is efficient provided that } \kappa = \hat{\kappa} \equiv \alpha^e \text{ and } \sigma = \hat{\sigma} \equiv 2 \frac{e^{\alpha^e}}{1 - \alpha^e} (1 - \alpha^e u^*) (1 - u^*) \left[ 1 + \frac{e^{\alpha^e}}{1 - \alpha^e} \ln u^* \right]. \]

The externality identified by Mortensen can be corrected by means of the MW, which lowers firms’ profits. The optimal MW factor is then $\beta = 1 - \alpha^e$. Taxing education is then necessary to reach the efficient allocation. It is a direct implication of condition (ii) of proposition 4: $\hat{\kappa} = \alpha^e$ implies $\hat{\kappa} \leq \alpha^e / (1 - \alpha^e)$.

7 Concluding comments

Employers derive some monopsony power in wage setting from workers’ imperfect ability to contact and work for alternative employers. In turn, such imperfect mobility jointly results from technological impossibilities and search frictions. In this paper, we suggest that education can reduce employers’ market power by offering mobility skills to the workers. Below, we highlight the market incentives to schooling may either be too high or too low.

These arguments take place within a wage-posting search equilibrium framework à la Burdett-Mortensen (1998) and Mortensen (2000). Workers have multidimensional skills and the search market is segmented by technology. Education is a time-consuming activity and determines the scope – or adaptability – of individual skills. Individuals get schooling to leave unemployment more quickly and to climb the wage ladder more rapidly through job-to-job mobility – that is, to speed up job-shopping. Because education lowers the effects of search frictions on workers’ mobility, it reduces employers’ monopsony power. Hence, the wage distribution shifts rightward when aggregate schooling increases, a positive externality. However, the ratio of vacant jobs to job-seekers also falls in each sector, a negative externality. A combination of minimum wage and schooling fee can decentralize the efficient allocation.
APPENDIX

A Proofs

The following lemma reveals useful throughout the different proofs.

**Lemma** Consider the functions $u$ and $\theta$ given by (24) and (25). Then,

(i) Properties of the function $\theta$

\[
\frac{\partial \theta}{\partial u} = -\frac{\theta}{\alpha u (1-u)}, \quad \frac{\partial \theta}{\partial T} = -\frac{\theta}{\alpha \lambda}, \quad \text{and} \quad \frac{\partial \theta}{\partial q} = \frac{\theta}{\alpha q + \delta}
\]

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

(ii) Properties of the function $u$

\[
\frac{\partial u}{\partial \theta} = -\frac{\alpha}{\theta} u (1-u), \quad \frac{\partial u}{\partial T} = -\lambda u (1-u), \quad \text{and} \quad \frac{\partial u}{\partial q} = \frac{u (1-u)}{q + \delta}
\]

Moreover, $\lim_{\theta \to 0} u = 1$, $\lim_{\theta \to \infty} u = 0$, $\lim_{T \to 0} u = 1$, $\lim_{T \to \infty} u = (q + \delta)/(q + \delta + m(\theta))$, and $\lim_{q \to \infty} u = 1$.

The above derivatives follow by direct calculation, and use of the fact that $\alpha \equiv \alpha(\theta) \equiv \theta m'(\theta)/m(\theta)$. The limits are then induced by the properties of the matching technology.

**Proposition 1** Consider the function $\phi_1$ such that

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

To prove existence and uniqueness of $\overline{u}, \overline{\theta}, \overline{F}, \overline{G}$, it is necessary and sufficient to show there exists a unique $\overline{u} \equiv \overline{u}(T) \in (0, 1)$ such that $\phi_1(\overline{u}, T) = 0$. Making use of the lemma, we get

\[
\lim_{u \to 0} \phi_1(u, T) = \infty \quad \text{and} \quad \lim_{u \to 1} \phi_1(u, T) = -\kappa \frac{y}{q + \delta}
\]

while

\[
\frac{\partial \phi_1(u, T)}{\partial u} = \frac{\gamma 1 - \alpha \frac{\partial \theta(u, T)}{\partial u}}{\eta \theta} - 2\kappa u \frac{y}{q + \delta} < 0
\]

The result follows. Then, $\overline{\theta}(T) = \theta(\overline{u}(T), T)$, while $w_{\max}, \overline{F}$ and $\overline{G}$ are given, resp, by (22), (21), and (20), with $u = \overline{u}$. 

21
To establish claims (i) to (iii), we make use of the implicit function theorem, according to which

$$\frac{\partial \phi_1 (\pi, T)}{\partial T} = \gamma \frac{1 - \alpha}{\theta} \frac{\partial \theta (u, T)}{\partial T} < 0$$

which has the sign of the numerator. As

$$\frac{\partial \phi_1 (u, T)}{\partial T} = \frac{1 - \alpha}{\eta} \frac{\partial \theta (u, T)}{\partial T} = \frac{1}{\eta} \frac{\partial \theta (u, T)}{\partial T}$$

we have \( \pi' (T) < 0 \). Then,

$$\bar{\pi}' (T) = \frac{\partial \theta (\pi, T)}{\partial u} \pi' (T) + \frac{\partial \theta (\pi, T)}{\partial T}$$

The two terms have opposite signs; a complete computation is required to conclude. But,

$$\pi' (T) = - \frac{(1 - \alpha) u (1 - u) \lambda}{1 - \alpha + 2 \alpha (1 - u)}$$

Hence,

$$\bar{\pi}' (T) = - \frac{2 (1 - u)}{1 - \alpha + 2 \alpha (1 - u)} \theta \lambda < 0$$

Finally,

$$\bar{\pi}'_{\max} (T) = -2 u \kappa y \pi' (T) > 0$$

$$\bar{F}_T (w, T) = F (w) \frac{\pi' (T)}{1 - u} < 0$$

$$\bar{G}_T (w, T) = \frac{u \bar{F}_T (w, T)}{(1 - u + u F (w))^2} + \frac{F (w) (1 - F (w)) \pi' (T)}{(1 - u + u F (w))^2} < 0$$

**proposition 2** 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. It writes down

$$\delta S (T_i) = S' (T_i)$$

As written in the text,

$$S' (T_i) = S (T_i) \left\{ - \frac{\partial u_i / \partial T_i}{1 - u_i} + \frac{\partial \bar{w}_i / \partial T_i}{\bar{w}_i} \right\}$$

But

$$\frac{\partial \bar{w}_i}{\partial T_i} = - \int_{u_0}^{v_{\max}} \frac{\partial G_i (w) / \partial T_i}{\partial T_i} dw$$

and

$$\frac{\partial G_i (w)}{\partial T_i} = G_i (1 - G_i) \frac{\partial u_i / \partial T_i}{u_i}$$
Since
\[ \frac{\partial u_i}{\partial T_i} = \frac{\partial u}{\partial T_i} = -\lambda_i u_i (1 - u_i) \] (58)
we finally get
\[ S_i(T_i) = \lambda_i S(T_i) \left\{ u_i + (1 - u_i) \int_{u_0}^{w_{\text{max}}} G_i(1 - G_i) dw \right\} \] (59)

Replacing (59) in (54) and manipulating the resulting equation yields (33).

(i) From (59),
\[ \lambda_i u_i S(T_i) < S_i(T_i) < \lambda_i S(T_i) \] (60)
These inequalities taken together with the first-order condition (54) yield the claim.

(ii) From the lemma, \( u_i \equiv u(\theta, T_i; q) \) is such that \( \lim_{\theta \to 0} u_i = \lim_{q \to \infty} u_i = 1 \) for all \( T_i \geq 0 \).
As a result, the wage distribution \( G_i \) given by (27) collapses to,
\[ \lim_{\theta \to 0} G_i(w) = \lim_{q \to \infty} G_i(w) = F(w) \] (61)
As \( \theta \) tends to 0, or \( q \) tends to infinity, the optimal schooling duration given by (33) therefore tends to the solution of \( \lambda(T_i) = \delta \).

(iii) When the wage distribution is degenerate, there is no longer a wage-mobility return to schooling and the first-order condition (54) collapses to
\[ \delta = -\frac{\partial u_i / \partial T_i}{1 - u_i} = \lambda_i u_i \] (62)
This closes the proof.

**Proposition 3** Consider the function \( \phi_2 \) such that
\[ \phi_2(u, T) \equiv [1 + \kappa + 2\kappa \ln u] u \lambda - \delta [1 - \kappa u] \] (63)
An equilibrium is a pair \((u^*, T^*)\) which solves \( \phi_1(u^*, T^*) = \phi_2(u^*, T^*) = 0 \).

Step 1. Let \( u_0 \) be such that \( \phi_1(u_0, \infty) = 0 \); let also \( T_1(u) \) be the (unique) solution of \( \phi_1(u, T_1(u)) = 0 \). Then \( T_1'(u) < 0 \) and
\[ \lim_{u \to u_0} T_1(u) = +\infty, \lim_{u \to 1} T_1(u) = 0 \] (64)
We know from (47) that \( \phi_1 \) is strictly decreasing in \( T \). In addition, \( \phi_1(u, 0) = \infty \) and \( \phi_1(u, \infty) < 0 \). Thus, 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 establishes the claim.
Step 2. Let $T_2(u)$ be the solution of $\phi_2(u, T_2(u)) = 0$. Then, $T_2'(u) > 0$ and

$$\lim_{u \to 0} T_2(u) = 0, \quad \lim_{u \to 1} T_2(u) = \lambda^{-1}(\delta)$$  \hspace{1cm} (65)

Taking the derivative of $\phi_2$ with respect to $T$, we get

$$[1 + \kappa + 2\kappa \ln u] u \lambda'(T) < 0$$  \hspace{1cm} (66)

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 3. Conclusion. The fact there exists a unique equilibrium follows directly from step 1 and step 2. Moreover, $T_2'(u^*) < 0 < T_2'(u^*)$. \hfill $\blacksquare$

**proposition 4** From the Lemma, there exists a unique equilibrium to which the implicit function theorem applies. Let $J$ denote the Jacobian matrix of function $\Phi \equiv (\phi_1, \phi_2)$ evaluated in equilibrium.

$$J = \begin{bmatrix}
\frac{\partial \phi_1}{\partial u} & \frac{\partial \phi_1}{\partial T} \\
\frac{\partial \phi_2}{\partial u} & \frac{\partial \phi_2}{\partial T}
\end{bmatrix}$$ \hspace{1cm} (67)

where partial derivatives are computed by means of the Lemma:

$$\frac{\partial \phi_1}{\partial u} \equiv -(1 + \alpha - 2\alpha u) \frac{\kappa}{\alpha} \frac{u}{1 - u} q + \delta y$$  \hspace{1cm} (68)

$$\frac{\partial \phi_1}{\partial T} \equiv -\frac{1 - \alpha}{\alpha} \kappa u^2 \frac{y}{q + \delta} \lambda$$  \hspace{1cm} (69)

$$\frac{\partial \phi_2}{\partial u} \equiv \frac{\delta}{u} + 2\kappa u \lambda \left[ \frac{1}{1 - u} \ln u + \frac{u}{(1 - u)^2} \ln u + \frac{1}{1 - u} \right]$$  \hspace{1cm} (70)

$$\frac{\partial \phi_2}{\partial T} \equiv \delta [1 - \kappa u] \frac{\lambda'}{\lambda}$$  \hspace{1cm} (71)

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

$$\begin{bmatrix}
\frac{du^*}{dx} \\
\frac{dT^*}{dx}
\end{bmatrix} = -J^{-1} \begin{bmatrix}
\frac{\partial \phi_1}{\partial x} \\
\frac{\partial \phi_2}{\partial x}
\end{bmatrix}$$ \hspace{1cm} (72)

where

$$\frac{\partial \phi_1}{\partial \kappa} \equiv -u^2 \frac{y}{q + \delta}$$ \hspace{1cm} (73)

$$\frac{\partial \phi_1}{\partial q} \equiv -\frac{\kappa}{\alpha} u^2 \frac{y}{(q + \delta)^2}$$ \hspace{1cm} (74)

$$\frac{\partial \phi_2}{\partial \kappa} \equiv \left[ 1 + 2 - \frac{u}{1 - u} \ln u \right] u \lambda + \delta u$$ \hspace{1cm} (75)

$$\frac{\partial \phi_2}{\partial q} \equiv 0$$ \hspace{1cm} (76)
and
\[
J^{-1} = \frac{1}{\det J} \begin{bmatrix}
\partial \phi_2 / \partial T & -\partial \phi_1 / \partial T \\
-\partial \phi_2 / \partial u & \partial \phi_1 / \partial u
\end{bmatrix}
\]
(77)
with \(\det J > 0\) since \(T_1^* (u^*) < 0 < T_2^* (u^*)\).

(i) The facts that \(du^* / dq > 0\), and \(dT^* / dq > 0\) come immediately. Moreover, \(u\) tends to 1 as \(q\) tends to infinity for all \(T \geq 0\) (i.e. even if \(T\) is actually \(+\infty\)). From (63), \(T^*\) then tends to \(\lambda^{-1}(\delta)\).

(ii) The fact that \(du^* / d\kappa < 0\) also comes immediately. Moreover
\[
\text{sign} \left\{ \frac{dT^*}{d\kappa} \right\} = \text{sign} \left\{ \frac{\partial \phi_2}{\partial u} \frac{\partial \phi_1}{\partial \kappa} - \frac{\partial \phi_1}{\partial u} \frac{\partial \phi_2}{\partial \kappa} \right\}
\]
(78)
Using equations (68) to (71), equations (73) and (75), multiplying the relevant expression by \(\frac{1-u}{u} \frac{q+\delta}{y}\) and grouping terms yields that \(dT^* / d\kappa\) has the sign of
\[
\Delta (\kappa) \equiv \kappa - \alpha + (\kappa + \alpha - 2\alpha \kappa) u + 2\alpha \kappa (\kappa - 1) u^2 + \frac{2\kappa u}{1-u} \ln u \{1 - \alpha - \alpha u + \alpha \kappa u\}
\]
(79)
Claim (ii) follows from the fact that \(\beta = 1 - \kappa\) and \(\Delta (\kappa) = -\alpha O (1 - u) < 0\) for \(\kappa\) 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 \(\beta\) tends to 1 implies that \(u\) tends to 0 for all \(T > 0\). \(\blacksquare\)

**proposition 5** Recall that \(u^p = \overline{u} (T^p)\) and \(u^* = \overline{u} (T^*)\). As the function \(\overline{u}\) is strictly decreasing in \(T\), the proof of proposition 5 simply requires to compare the rhs of equations (35) and (36) at given unemployment rate \(u\), having in mind there is over-(resp., under-)education iff the rhs of (35) is strictly lower (larger) than the rhs of (36). To this aim, consider the functions \(\psi_1\) and \(\psi_2\) such that
\[
\psi_1 (\alpha, \kappa, u) = \frac{(1-\alpha) [1 - \kappa + 2\kappa (1-u)]}{1-\alpha + 2\alpha (1-u)}
\]
(80)
\[
\psi_2 (\kappa, u) = 1 + \kappa + 2\kappa \frac{u}{1-u} \ln u
\]
(81)
The rhs of (35) is worth \(\psi_1 (\alpha^p, \kappa, u)\), while the rhs of (36) is simply \(\psi_2 (\kappa, u)\).

To establish claim (i), note first that the function \(f\) defined by
\[
f (u) = \frac{u}{1-u} \ln u
\]
(82)
It is strictly decreasing in \(u\), with \(f (0) = 0\), \(f (1) = -1\), and \(f (u) < -u\). Therefore,
\[
1 - \kappa < \psi_2 (\alpha, \kappa, u) < 1 - \kappa + 2\kappa (1-u)
\]
(83)
Consider \( \alpha^p \) as a parameter – it is so when the matching technology is Cobb-Douglas. The function \( \psi_1 \) is continuous and strictly decreasing in \( \alpha \), with, for all \( \kappa > 0 \),

\[
\psi_1 (0, \kappa, u) = 1 - \kappa + 2\kappa (1 - u) > \psi_2 (\kappa, u) \tag{84}
\]
\[
\psi_1 (1, \kappa, u) = 0 < \psi_2 (\kappa, u) \tag{85}
\]

It follows from (84) and (85) and the continuity of the function \( \psi_1 \) that \( T^s > T^* \) when \( \alpha^p \) is “small”, while \( T^s < T^* \) when \( \alpha^p \) is “large”. Hence, both over- and under-education are possible.

To prove (ii), note that \( \psi_1 \) can be written as follows

\[
\psi_1 (\alpha, \kappa, u) = 1 - \kappa + 2(1 - u) \frac{(1 - \alpha) \kappa - \alpha}{1 - \alpha + 2\alpha (1 - u)} \tag{86}
\]

Therefore, \( \psi_1 (\alpha, \kappa, u) < \psi_2 (\kappa, u) \) if \( \kappa \leq \alpha/1 - \alpha \). The result follows.

**proposition 6** Optimal schooling solves \( S' (T_i) - \sigma y = \delta S (T_i) \). Using (30), it comes:

\[
w_0 \{ \lambda_i u_i - \delta \} - \frac{\delta \sigma}{1 - u_i} + \int_{w_0}^{w_{\text{max}}} [1 - G_i (w)] \left[ \lambda_i \frac{G_i (w)}{F (w)} - \delta \right] dw = 0 \tag{87}
\]

In symmetric equilibrium, \( T_i = T \) for all individuals and the wage offer distribution is given by (17). We get (after some tedious algebra):

\[
\left[ 1 + \kappa + 2\kappa \frac{u}{1 - u} \ln u \right] u \lambda = \delta \left[ 1 - \kappa u + \frac{\sigma}{1 - u} \right] \tag{88}
\]

The decentralized equilibrium is then composed of equations (FE) and (88), while the efficient allocation is defined by equations (38) and (39). Both systems coincide whenever \( \kappa = \alpha^s \) and

\[
\left[ 1 + \alpha^s + 2\alpha^s \frac{u^s}{1 - u^s} \ln u^s \right] u^s \lambda^s - \delta \frac{\sigma}{1 - u^s} = (1 - \alpha^s) u^s \lambda^s \tag{89}
\]

By mean of equation (88), we see that (89) is equivalent to \( \sigma = \hat{\sigma} \).
References


28


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 (87), which we reproduce now for convenience:

\[
\begin{align*}
 w_0 \left\{ \lambda_i u_i - \delta \right\} + \frac{\delta}{q + \delta} \\
 + \int_{w_0}^{w_{\text{max}}} \frac{1 - F(w)}{u_i + (1 - u_i) (1 - F(w))} \left\{ \lambda_i u_i + (1 - u_i) (1 - F(w)) \right\} dw = 0
\end{align*}
\]

(90)

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

\[
1 - F(w) = \frac{-u + \left( \frac{y-w}{y-w_0} \right)^{1/2}}{1-u}
\]

(91)

We now compute the term $A_3$.

\[
A_3 = \frac{1}{1-u} \int_{w_0}^{w_{\text{max}}} \left[ 1 - u \left( \frac{y-w}{y-w_0} \right)^{-1/2} \right] \left[ \lambda u \left( \frac{y-w}{y-w_0} \right)^{-1/2} - \delta \right] dw
\]

\[
= -\frac{\lambda u^2}{1-u} \int_{w_0}^{w_{\text{max}}} \left( \frac{y-w}{y-w_0} \right)^{-1} dw + \frac{u(\lambda+\delta)}{1-u} \int_{w_0}^{w_{\text{max}}} \left( \frac{y-w}{y-w_0} \right)^{-1/2} dw
\]

(92)

But,

\[
\int_{w_0}^{w_{\text{max}}} (y-w)^{-1/2} dw = -2 (y-w_{\text{max}})^{1/2} + 2 (y-w_0)^{1/2}
\]

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

(93) (94)

Then,

\[
A_3 = -\frac{\lambda u^2}{1-u} (y-w_0) \ln \frac{y-w_0}{y-w_{\text{max}}} + \frac{2u}{1-u} (\lambda + \delta) (y-w_0)
\]

\[
-\frac{2u}{1-u} (\lambda + \delta) (y-w_0)^{1/2} (y-w_{\text{max}})^{1/2}
\]

(95)

\[
-\frac{\delta}{1-u} (y-w_{\text{max}})^{1/2} - \frac{\delta}{1-u} (y-w_0)
\]

The MW is worth $w_0 = (1-\kappa) y$. From (22), $y-w_{\text{max}} = \kappa u^2 y$. Using these two facts, we get

\[
A_3 = \frac{2\lambda u^2}{1-u} \kappa y \ln u + 2u (\lambda + \delta) \kappa y - \delta (1 + u) \kappa y
\]

(96)

Therefore,

\[
\frac{1-u}{y} A_3 = \frac{2\lambda u^2}{1-u} \kappa \ln u + 2u (1 - u) (\lambda + \delta) \kappa - \delta (1 - u^2) \kappa
\]

(97)

The (OS) curve results from:

\[
\frac{1-u}{y} (A_1 + A_2 + A_3) = 0
\]

(98)
Equations (90), (97) and (98) imply that
\[
(1 - u)(1 - \kappa) \lambda u - (1 - \kappa) (1 - \kappa) \delta - \delta \sigma + 2\kappa \lambda u^2 \ln u + 2\kappa u (1 - u) (\lambda + \delta) - \kappa \delta (1 - u^2) = 0 \tag{99}
\]

Grouping terms, we finally get the (OS) curve:
\[
\left[1 + \kappa + 2\kappa \frac{u}{1 - u} \ln u\right] u\lambda = \delta \left[1 - \kappa u - \frac{\sigma}{1 - u}\right] \tag{100}
\]

Equation (OS) is obviously a special case of (100) with \(\sigma = 0\).
Fig. 1: Education and the wage offer distribution

Fig. 2: Comparative statics
Fig. 3: Education and the minimum wage