DISCUSSION PAPER SERIES

IZA DP No. 8341

Formal Education Versus Learning-by-Doing

Frédéric Gavrel
Isabelle Lebon
Thérèse Rebière

July 2014

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Formal Education Versus Learning-by-Doing

Frédéric Gavrel
CREM-CNRS, University of Caen and TEPP-CNRS

Isabelle Lebon
CREM-CNRS, University of Caen and TEPP-CNRS

Thérèse Rebière
LIRSA-CNAM and IZA

Discussion Paper No. 8341
July 2014

IZA

P.O. Box 7240
53072 Bonn
Germany

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

IZA Discussion Paper No. 8341
July 2014

ABSTRACT

Formal Education Versus Learning-by-Doing*

The efficiency of educational choices is studied in a search-matching model where individuals face a tradeoff: acquiring formal education or learning while on the job. When their education effort is successful, newcomers directly obtain a high-skill job; otherwise, they begin with a low-skill job, learn-by-doing and then search while on-the-job for a high-skill job. Low-skill firms suffer from hold-up behavior by high-skill firms. The low-skill sector is insufficiently attractive and individuals devote too much effort to formal education. A self-financing tax and subsidy policy restores market efficiency.

JEL Classification: H21, I20, J21, J64, J68

Keywords: formal education, learning-by-doing, market efficiency, on-the-job search, search unemployment

Corresponding author:

Thérèse Rebière
LIRSA-CNAM
Department of Economics
40 rue des Jeûneurs
Case 1D2P30
75002 Paris
France
E-mail: therese.rebiere@cnam.fr

* We would like to thank Björn Brügemann, Gary Fields, and Bernd Fitzenberger for their helpful comments on earlier versions of that work. The usual caveat applies.

1 Introduction

Formal education is not the only way to acquire skills that give workers the opportunity of gaining a good job. Learning-by-doing in a low-skill job and then searching (while on-the-job) for a high-skill job is another way of reaching the same goal. Do workers choose the right amount of formal education when faced with this trade-off? If not, what kind of public policy should be implemented?

Although human capital is generally measured by the amount of formal education, many skills are best learned on-the-job through participating in the production process. Consequently, training here also determines workers' productivity. In the absence of learning-by-doing, workers would always hold a job equivalent to their education level, without any prospect of improvement.

Using US data, Andersson *et al.* (2005) find that 15% to 20% of workers with a high school diploma or less had escaped low-wage employment after nine years. This result seems to indicate that there is a stepping-stone effect toward better paid jobs (Connolly and Gotschalk 2001) which is more likely to occur when workers voluntarily change jobs (Sicherman and Galor 1990, Gotschalk 2001, Holzer 2004), and even more likely when employment is gained at a higher-wage firm (Andersson *et al.* 2003). Such a spring-board effect is demonstrated for Germany by Knabe and Plum (2013) who state that the rate of transition to a high-paid job conditional on first accepting a low-paid job is particularly significant for low-skilled workers. Likewise, empirical evidence for France shows that, since the 1980s, upward professional mobility has improved, especially for low-skilled workers: 29% of blue collar workers in 1998 had experienced upward mobility between 1998 and 2003, against 19% between 1980 and 1985. We can think of a secretary becoming an executive secretary, or an unskilled worker in a routine occupation becoming skilled in a lower technical occupation in sectors such as manufacturing, industrial crafts, construction, and warehousing and transport; and then transitioning from lower to intermediate technical occupations, such as technician to foreman or supervisor (Monso 2006). Most of those occupations could also be entered by obtaining a diploma in the relevant field of competence. This evidence is consistent with our framework, in which educated workers gain a well-paid job directly, whereas workers with a lower level of education have to train themselves on-the-job before gaining a better-paid job.

The fact remains that, during the past few decades, more and more individuals have chosen to reinforce their effort in formal education (see for instance Machin (1996), Acemoglu (2002), Mincer (1994, 2003), and Moscarini and Vella (2008)). Did these private educational choices lead to an efficient outcome? The purpose of this paper is to shed some light on this issue. We argue that individuals tend to put too much emphasis on formal education, compared with training in the workplace. The reason for this does not involve educational decisions themselves. This distortion originates in the fact that firms with high-skill jobs underestimate the social cost of filling their vacancies with workers previously employed in low-skill jobs in which they have practised learning-by-doing. Firms create too many high-skill jobs. In response to this hold-up behavior, job creation is suboptimal in the low-skill sub-market. As a result, high-skill jobs are too appealing, and individuals make too great an effort to acquire formal education. This creates a problem requiring government involvement.

To assess the consistency of our argument we use a two-sector search-matching model in which workers have a finite life expectancy (Moen and Rosén 2004, Gavrel *et al.* 2010). In contrast to these previous papers, our model assumes that workers can become skilled via formal education. Before entering the labor market, new workers decide on the amount of effort to devote to formal education. If they succeed in acquiring the required skills, they directly join the pool of applicants for good jobs. If they fail, they have to search for a low-skill job, and then begin to learn while on-the-job. When the learning-by-doing process comes to its end, workers are endowed with the same skills as (formally) educated workers (following Arrow (1962)). They then can join the pool of applicants for good jobs.

First, we describe a (decentralized) stationary equilibrium of the labor market and its efficiency conditions. Assuming that firms internalize the well-known congestion effect (Hosios 1990, Pissarides 2000), high-skill job creation appears to be too high; whereas low-skill job creation, as well as individuals' educational choices, are constrained efficient. In other words, they are optimal for the equilibrium value of the tightness of the high-skill sub-market. This means that inefficiency derives entirely from an excessive creation of high-skill vacancies. Next, we compare the decentralized equilibrium with a social optimum. The results validate the consistency of our argument: low skill jobs are too few in number, and individuals put too much emphasis upon formal education.

Second, we show that a Tax and Subsidy Policy (TSP) can decentralize the social optimum. Taxes must be levied on (filled) good jobs. They ensure that perceived hiring costs coincide with social costs. However, these taxes distort low-skill job creation as well as educational choices. In order to restore market efficiency, these taxes must be dedicated to the funding of two kinds of compensatory transfer. One is allocated to firms in the low-skill sub-market when they lose workers leaving them for a better job. The other is a reward that workers receive if their formal education is successfully completed. Rewarding graduates is necessary since the taxes which have to be levied on high-skill jobs excessively reduce the surplus for a match with such jobs, hence the returns to formal education for workers.

Economists have been interested in the efficiency of human capital investment for a long time. A controversial issue, going back to Pigou (1912), concerns governmental involvement that seeks to enhance skills. Since firms would not have an interest in investing in workers' skills because of the risk that their experienced workers might quit for external opportunities, government subsidies appear to be a necessary measure for the improvement of training and of schooling. By contrast, Becker (1964) pointed out that the solution for human capital inefficiency may be improved loan markets rather than government regulation and training subsidies. A competitive labor market implies that workers are the only ones who have an incentive to invest in their general training, bearing the cost themselves, either directly or by taking a wage cut. Hence, the appropriate amount of investment for an efficient market would be undertaken unless workers are credit-constrained.

More recently, labor theory has re-examined the issue of educational choices in the presence of market imperfections. Our paper is a contribution to this literature. Acemoglu and Pischke (1999) argue that search frictions allow us to account for employer-provided on-the-job training, since firms are able to recover such investment in human capital. Moen (1999) studies the efficiency of educational choices. Firms rank their applicants and hire the best, while workers use formal education to compete for jobs. Here, the education effort can be too intensive. The same result is presented by Charlot and Decreuse (2007), in which workers self-select their educational choices (see also Charlot and Decreuse (2005)). Workers of low ability place too much emphasis on the value of a higher formal education for gaining a job, even though education can be costly. Such inefficient behavior leads the authors to suggest that educational subsidies be prohibited.

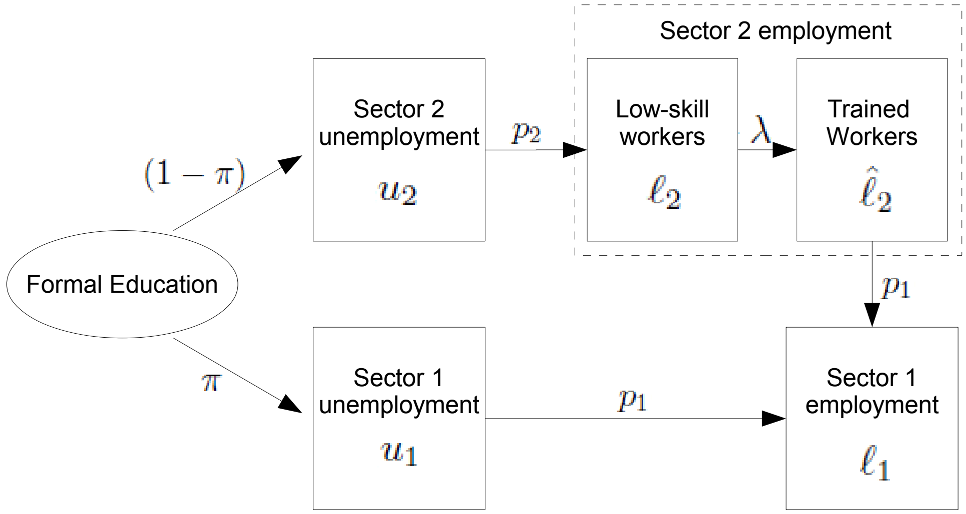
The paper is organized as follows: Section 2 outlines the analytical framework. We define a labor market decentralized (stationary) equilibrium in section 3. Section 4 studies market efficiency and states two main results: a decentralized equilibrium is constrained efficient in terms of low-skill job creation and educational choices but inefficient in terms of high-skill job creation; the *laissez-faire* situation is inefficient. In section 5, we outline a self-financed fiscal policy which rewards educational success and leads to a social optimum. In section 6, we calibrate the model with US empirical evidence. Finally, section 7 contains some concluding comments.

2 Analytical framework

The economy consists of two types of agents: workers and firms. Firms are infinitely-lived whereas workers have a finite life expectancy of $1/m$. Time is continuous and parameter m measures the workers' labor market exit rate. Each worker who leaves the market is replaced with a newcomer. The measure of the total labor force is constant and normalized to one. All agents are risk-neutral and discount future payoffs at rate r ($r \geq 0$).

The labor market is segmented into two interacting sub-markets (sectors arranged into a hierarchy). Sector 2 offers low-skill jobs, while sector 1 offers high-skill jobs. Workers decide on their effort in formal education e when entering the economy. If their effort is successful (which occurs with the probability π), workers will enter the pool of applicants for high-skill jobs (high-skill unemployment); whereas workers with unsuccessful effort will enter the pool of applicants for low-skill jobs (see figure 1). The probability π is assumed to be an increasing and concave function $\pi(e)$ of the education effort e ($\pi'(.) > 0$, $\pi''(.) < 0$). Workers with low-skill jobs will therefore have to engage in a learning-by-doing process in order to become skilled enough to be employable in a high-skill firm. The expected duration of this process is denoted by $(1/\lambda)$. Workers thus acquire the required skills at a Poisson rate λ. When the learning period comes to an end, workers engage in an on-the-job search process, hoping to get a high-skill job. The incentive to look for a high-skill job is the wage differential between sectors.

Figure 1: Workers' flows



When entering the labor market firms choose the sub-market i ($i = 1, 2$) in which they will operate. They then create a single job in their chosen sub-market. Frictions exist that prevent the instantaneous matching of jobs with workers. Firms thus have to pay a cost, c, in order to keep their vacancy open. When matched with a worker, jobs yield output y_1 in sector 1, $\hat{y}_2$ in sector 2 when workers are trained, and y_2 when workers are novice (with $y_1 > \hat{y}_2 > y_2$). Wages are negotiated. Workers have a bargaining power of β and firms have a bargaining power of $(1-\beta)$. Sector 1 offers the wage w_1; whereas sector 2 offers the wage w_2 when workers are untrained, and the wage $\hat{w}_2$ when workers have learned by doing.

Job creation results from the usual assumption of free entry in both sectors. Market frictions in sector-i are summarized in a constant-returns matching function that defines the arrival rate of workers to job vacancies $q_i(\theta_i)$ with $q_i'(\theta_i) < 0$. The arrival rate of job offers to searching workers is $p_i = \theta_i q_i$ with $p_i'(\theta_i) > 0$ where θ_i is the sub-market tightness.

2.1 High-skill jobs

2.1.1 Asset values

In sub-market 1, the lifetime utility of an employed worker, called W_1, satisfies:

$$(r+m)(W_1 - U_1) = w_1 - (r+m)U_1 \tag{1}$$

where U_1 is the lifetime utility of a high-skilled worker when unemployed. We have:

$$(r+m)U_1 = d + p_1(W_1 - U_1) \tag{2}$$

with d being the value of leisure.
Regarding sector-1 firms, the value of a filled job, called J_1, verifies:

$$(r+m)(J_1 - V_1) = y_1 - w_1 - (r+m)V_1 \tag{3}$$

where V_1 is the asset value of a sector-1 firm whose job is vacant. We have:

$$rV_1 = -c + q_1(J_1 - V_1) \tag{4}$$

2.1.2 Private surplus and market tightness

Firms are unlikely to commit to wages. Following Shimer (2006), we assume that when an on-the-job seeker meets a firm, she must reject her current job and bargain with the other employer with no possibility of reverting to her old job. Therefore, the worker's threat point is always unemployment rather than the value of her previous job. This assumption follows Mortensen (2003, p.99) who pointed out that making counteroffers is not the norm in many labor markets. The outside option of a worker in sector 1 cannot be the lifetime utility of an employed worker in sector 2 who chooses to keep her low-skill job, but unemployment in sector 1[1]. When a worker and a firm meet and agree to form a match, the private surplus $S_1 = [W_1 - U_1] + [J_1 - V_1]$ of this match is shared between the worker and the firm according to their bargaining power. From equations (1) and (3), we deduce that the (private) surplus of a match in sub-market 1, satisfies:

$$(r+m)S_1 = y_1 - (r+m)(U_1 + V_1)$$

As the wage w_1 stems from static Nash bargaining, we have:

$$\beta S_1 = [W_1 - U_1] \tag{5}$$

As already mentioned, in both sub-markets job creation results from the assumption of free-entry ($V_1 = 0$). We thus have:

$$(r + m + \beta p_1(\theta_1))S_1 = y_1 - d \tag{6}$$

Consequently, by using (4), the market tightness θ_1 is determined by the following equilibrium equation:

$$-c + q_1(\theta_1)(1-\beta)S_1 = 0 \tag{7}$$

This equilibrium equation is equivalent to the reduced form of the basic matching model (Pissarides (2000)). Hence, an increase in parameters c, β, d, r and m lowers the market tightness θ_1, whereas an increase in the output y_1 stimulates job creation in this sub-market.

[1] Other outside options are nevertheless studied in section 4.4.2

2.2 Low-skill jobs

2.2.1 Asset values

When the training period comes to an end, the output of a worker in a low-skill job rises from y_2 to $\hat{y}_2$ and the worker begins to search (while on the job) for a high-skill vacancy. Her outside opportunities are defined by the lifetime utility of an unemployed worker in sub-market 1 (utility U_1). As Nash bargaining is static, the wage jumps from w_2 to $\hat{w}_2$. It means that we first need to define the asset values associated with a match between a low-skill job and a trained worker (hereafter referred to as an on-the-job seeker). So let $\hat{W}_2$ be the lifetime utility of such a worker. Using (5), one can show that this asset value satisfies:

$$(r + m + p_1)(\hat{W}_2 - U_1) = \hat{w}_2 + p_1 \beta S_1 - (r + m)U_1 \tag{8}$$

Regarding sector 2 firms, the value of a low-skill job when matched with an on-the-job seeker, called $\hat{J}_2$, verifies:

$$(r + m + p_1)(\hat{J}_2 - V_2) = \hat{y}_2 - \hat{w}_2 - rV_2 \tag{9}$$

where V_2 is the value of a sector 2 vacancy.

From equations (8) and (9), we derive the (private) surplus of a match of a sector 2 firm with an on-the-job seeker, $\hat{S}_2$. Knowing that $\hat{S}_2 = [\hat{W}_2 - U_1] + [\hat{J}_2 - V_2]$, the private surplus $\hat{S}_2$ satisfies:

$$(r + m + p_1)\hat{S}_2 = \hat{y}_2 + p_1 \beta S_1 - (r + m)U_1 - rV_2 \tag{10}$$

Under the assumption of free-entry ($V_2 = 0$), the substitution of (2) into (10) yields:

$$(r + m + p_1(\theta_1))\hat{S}_2 = \hat{y}_2 - d \tag{11}$$

We can now define the asset values associated with a match between a sector 2 firm and a newcomer.

As Nash bargaining implies that:

$$\hat{W}_2 - U_1 = \beta \hat{S}_2,$$

we obtain the result that the lifetime utility of an unskilled worker when holding a sector 2 job, W_2, satisfies:

$$(r + m + \lambda)(W_2 - U_2) = w_2 + \lambda \beta \hat{S}_2 + \lambda U_1 - (r + m + \lambda)U_2 \tag{12}$$

where U_2 is the value of unemployment in this sub-market. We have:

$$(r + m)U_2 = d + p_2(W_2 - U_2) \tag{13}$$

On the firms' side, the value of a job when held by a newcomer verifies:

$$rJ_2 = y_2 - w_2 - m(J_2 - V_2) + \lambda(\hat{J}_2 - J_2) \tag{14}$$

Under the assumptions of free-entry ($V_2 = 0$) and Nash bargaining, the latter equation can be rewritten as:

$$(r + m + \lambda)J_2 = y_2 - w_2 + \lambda(1 - \beta)\hat{S}_2 \tag{15}$$

2.2.2 Private surplus and market tightness

The private surplus of an untrained worker matched with a sector 2 firm is such that $S_2 = [W_2 - U_2] + [J_2 - V_2]$. From equations (12) and (15), we deduce S_2 as a function of $\hat{S}_2$:

$$(r + m + \lambda)S_2 = y_2 + \lambda\hat{S}_2 + \lambda U_1 - (r + m + \lambda)U_2 \tag{16}$$

Finally, by using (2) and (13), one can see that equation (16) can be rewritten as follows:

$$\frac{(r + m + \lambda)(r + m + \beta p_2(\theta_2))}{r + m} S_2 = y_2 + \lambda\hat{S}_2 - d + \frac{\lambda}{r + m}\beta p_1(\theta_1)S_1 \tag{17}$$

According to equation (17), the tightness of sub-market 2 is a function of the tightness of sub-market 1 via the term $\beta p_1 S_1$. Equilibrium in sector 2 thus depends on the equilibrium in sector 1. This results from the fact that workers' asset values in sector 2 depend on workers' asset values in sector 1. This one-way interdependence will play a crucial role in the efficiency study.

As a result, the assumption of free-entry determines the market tightness θ_2 by the following equilibrium equation:

$$-c + q_2(1 - \beta)S_2 = 0 \tag{18}$$

where the cost of keeping a vacancy open, c, is assumed to be the same in both sub-markets.

2.3 Educational choices

When entering the economy, a new worker decides on the amount of effort to devote to formal education. Her effort, denoted by e, determines the probability π of becoming a high-skilled worker. If she succeeds, she enters the pool of applicants for high-skill jobs; if she fails, she must search for a low-skill job and must begin a learning-by-doing process after finding one. Remember that the probability π is an increasing and concave function $\pi(e)$ of effort e such that $\pi'(.) > 0$, $\pi"(.) < 0$.

Education effort is then obtained by maximizing the following objective:

$$ED \equiv -e + \pi(e)U_1 + (1 - \pi(e))U_2 \tag{19}$$

We obtain the following first order condition:

$$\pi'(e)(U_1 - U_2) - 1 = 0 \tag{20}$$

For obvious reasons, the effort e increases with the difference $(U_1 - U_2)$. From the concavity of function $\pi(.)$ we deduce that the second order condition is satisfied.
Using equations (2) and (13), we can rewrite the optimality condition as follows:

$$\pi'(e)\beta(p_1 S_1 - p_2 S_2) - (r + m) = 0 \tag{21}$$

Education effort is an increasing function of the private surplus S_1, whereas it is a decreasing function of the private surplus S_2. In other words, workers would have an incentive to increase (reduce) their education effort if the gain from holding a high-skill (low-skill) job increases.

3 Equilibrium

3.1 Definition

In sum, a labor market equilibrium can be defined as follows:

Definition 1. *A labor market equilibrium is a set of values* $(S_1^*, \theta_1^*, \hat{S_2}^*, S_2^*, \theta_2^*, e^*)$ *which jointly satisfy equations (6), (7), (11), (17), (18) and (21).*

From market tightness and probability π^*, the employment and unemployment levels in both sub-markets can be deduced by using the conditions for flow-equilibrium.

3.2 Employment and unemployment levels

In a steady state, employment and unemployment levels are deduced from the flow-equilibrium conditions.
In sub-market 1, high-skill unemployment u_1 and high-skill employment ℓ_1 are obtained from the following equations:

$$m\pi = (m + p_1)u_1 \tag{22}$$

$$m\ell_1 = p_1(u_1 + \hat{\ell}_2) \tag{23}$$

where $\hat{\ell}_2$ is the number of on-the-job seekers (*i.e.* the level of high-skill employment in sub-market 2).
In sub-market 2, low-skill unemployment u_2, low-skill employment ℓ_2 and high-skill employment $\hat{\ell}_2$ are derived from the following conditions:

$$m(1 - \pi) = (m + p_2)u_2 \tag{24}$$

$$m\ell_2 + \lambda\ell_2 = p_2 u_2 \tag{25}$$

$$(m + p_1)\hat{\ell}_2 = \lambda\ell_2 \tag{26}$$

With v_i denoting vacant jobs in the labor sub-market i, the sub-market tightness of sector 1 is given by $\theta_1 = v_1/(u_1 + \hat{\ell}_2)$ and the sub-market tightness of sector 2 is given by $\theta_2 = v_2/u_2$. From these flow-equilibrium conditions, we derive the impacts of variables θ_1, θ_2 and π on all employment and unemployment levels. Table 1 reports these partial derivatives. The variable η_i $(i = 1, 2)$ denotes the elasticity of rate q_i with respect to market tightness θ_i (in absolute value).

Table 1: Partial derivatives of employment and unemployment levels

	u_2	ℓ_2
θ_1	0	0
θ_2	$-\frac{m(1-\pi)q_2(1-\eta_2)}{(m+p_2)^2}$	$\frac{m^2(1-\pi)q_2(1-\eta_2)}{(m+\lambda)(m+p_2)^2}$
π	$-\frac{m}{m+p_2}$	$-\frac{mp_2}{(m+\lambda)(m+p_2)}$

	$\hat{\ell}_2$	u_1	ℓ_1
θ_1	$-\frac{\lambda \ell_2 q_1(1-\eta_1)}{(m+p_1)^2}$	$-\frac{m\pi q_1(1-\eta_1)}{(m+p_1)^2}$	$\frac{(u_1+\hat{\ell}_2)q_1(1-\eta_1)}{m+p_1}$
θ_2	$\frac{\lambda}{m+p_1}\frac{\partial \ell_2}{\partial \theta_2}$	0	$\frac{\lambda p_1}{m(m+p_1)}\frac{\partial \ell_2}{\partial \theta_2}$
π	$-\frac{\lambda m p_2}{(m+p_1)(m+\lambda)(m+p_2)}$	$\frac{m}{m+p_1}$	$\frac{p_1}{m+p_1} + \frac{p_1}{m}\frac{\partial \hat{\ell}_2}{\partial \pi}$

The tightness in sub-market 1 is independent of that in sub-market 2. However owing to the interactions between the two sub-markets, high-skill employment depends on the transition rates in sectors 1 and 2. Therefore high-skill employment depends on job creation in the low-skill sub-market.

4 Efficiency

We now study the welfare properties of a decentralized equilibrium (Definition 1). As in Gavrel *et al.* (2010), firms do not internalize the social cost of hiring a high-skilled worker coming from the low-skill sector. The creation of high-skill jobs thus appears to be too high. Due to this hold-up phenomenon, job creation is suboptimal in the low-skill sub-market. As a consequence, educational choices are inefficient; workers devote too much effort to formal education.

Similar to Hosios (1990) and Pissarides (2000), let us consider a social planner who is only subject to search frictions and can redistribute income at no cost. In this case, the efficiency criterion is the social surplus. For the sake of expositional simplicity, the interest

rate r is assumed to be equal to zero. This assumption allows us to compare steady states according to the social surplus per period.
Denoted by Σ, the social surplus per head and per period is given by:

$$\Sigma = \ell_1 y_1 + \ell_2 y_2 + \hat{\ell}_2 \hat{y}_2 + (u_1 + u_2)d - \theta_1(u_1 + \hat{\ell}_2)c - \theta_2 u_2 c - me \tag{27}$$

Notice that in (27) the last term, me, measures the cost of formal education, per period. In what follows, for methodological reasons, we will assume that the usual Hosios condition holds true in both sub-markets, that is:

$$\eta_1 = \eta_2 = \beta$$

The matching functions are therefore Cobb-Douglas.

4.1 High-skill job creation

Let us first study the efficiency of high-skill job creation. Using Table 1, one can show that the derivative of the surplus Σ with respect to θ_1 has the same sign as:

$$HS \equiv (1 - \eta_1)q_1 \left[y_1 - \left(\frac{\hat{\ell}_2}{u_1 + \hat{\ell}_2} \hat{y}_2 + \frac{u_1}{u_1 + \hat{\ell}_2} d \right) \right] - (m + \eta_1 p_1)c \tag{28}$$

Under the Hosios condition and with $r = 0$, the decentralized equilibrium in sub-market 1 can be rewritten (see equations (6) and (7)):

$$(1 - \eta_1)q_1(y_1 - d) - (m + \eta_1 p_1)c = 0$$

As $\hat{y}_2 > d$, we have:

$$\frac{\hat{\ell}_2}{u_1 + \hat{\ell}_2} \hat{y}_2 + \frac{u_1}{u_1 + \hat{\ell}_2} d > d$$

This implies $HS < 0$ in a decentralized equilibrium. Sector 1 firms thus create too many vacancies. This results means that $\theta_1^S < \theta_1^*$, where θ_1^S denotes the social optimum value of the tightness of sub-market 1.

When the relevant outside option for sector 1 workers (coming from sector 2) is the asset value of high-skill unemployment, the wage bargaining of on-the-job seekers with sector 1 firms yields an inefficiency that is not solved by the Hosios rule.

The intuition behind this inefficiency result is that, with static Nash bargaining, sector 1 firms underestimate the (social) opportunity cost of a match realized with a worker coming from sector 2. This cost is given by the output $\hat{y}_2$ which is higher than the value of leisure, d. As a consequence, high-skill job creation is the more inefficient, the larger the share of on-the-job seekers in the pool of applicants for high-skill jobs. In other words, sector 2 firms suffer from hold-up behavior from sector 1 firms.
This distortion clearly depends on the manner in which one defines the outside option for skilled workers. We come back to this issue in section 4.4 where we study the consequences of alternative threat points.

4.2 Low-skill job creation

One can show that the derivative of the social surplus Σ with respect to the market tightness θ_2 has the same sign as (see A):

$$\begin{aligned} LS \equiv \quad & (1-\eta_2)q_2 \frac{\lambda}{m(m+p_1)}[p_1(y_1-d) - m\theta_1 c] \\ & +(1-\eta_2)q_2\left[y_2 - d + \frac{\lambda(\hat{y}_2 - d)}{m+p_1}\right] - \frac{(m+\lambda)(m+\eta_2 p_2)}{m}c \end{aligned} \tag{29}$$

We shall state that LS is equal to zero in a decentralized equilibrium (Definition 1).

For $r = 0$, combining equilibrium equations (11) (17) and (18) gives:

$$(1-\beta)q_2\frac{\lambda}{m}\beta p_1 S_1 + (1-\beta)q_2\left[y_2 - d + \frac{\lambda(\hat{y}_2-d)}{m+p_1}\right] - \frac{(m+\lambda)(m+\beta p_2)}{m}c = 0 \tag{30}$$

Let us consider the first term of the previous equation:

$$(1-\beta)q_2\frac{\lambda}{m}\beta p_1 S_1$$

From the equilibrium equation (7), we deduce:

$$\beta p_1 S_1 = p_1 S_1 - \theta_1 c$$

With Nash bargaining, part of the surplus goes to the firms. However, under the assumption of free entry, profits are dedicated to the funding of vacancy costs.
On the other hand, from the definition of private surplus S_1, it follows that:

$$\beta p_1 S_1 = y_1 - d - m S_1$$

We then obtain:

$$p_1 S_1 - \theta_1 c = y_1 - d - m S_1 \Leftrightarrow S_1 = \frac{y_1 - d + \theta_1 c}{m + p_1}$$

Consequently, the first term of the equilibrium equation for tightness θ_2 can be rewritten as:

$$(1-\beta)q_2\frac{\lambda}{m}\beta p_1 S_1 = (1-\beta)q_2\frac{\lambda}{m(m+p_1)}[p_1(y_1-d) - m\theta_1 c]$$

Under the Hosios condition, this proves that the derivative of the social surplus with respect to the sub-market tightness θ_2 is zero in a decentralized equilibrium (hence for an inefficient decentralized equilibrium value of θ_1).

The decentralized equilibrium value of the tightness of sub-market 2, θ_2^*, does not coincide with its social optimum value, θ_2^S. The value of θ_2^* indeed depends on the tightness of sub-market 1, θ_1^*. We have shown in the previous section that θ_1^* takes a value which is higher than its optimal value θ_1^S. Thus θ_2^* differs from θ_2^S even if $LS = 0$ in a decentralized equilibrium. We compare those two values in section 4.4. $LS = 0$ can however be read as a "constrained" efficient result of sub-market 2 in the sense of Moen and Rosén (2004): job creation is socially optimal in sub-market 2 if the social planner chooses the same tightness as the one determined in sub-market 1 in a decentralized equilibrium, θ_1^*.

At first glance, this (constrained) efficiency result might look surprising as, via on-the-job search, high-skill employment depends positively on low-skill job creation (see Table 1). In fact, an increase in high-skill employment raises the output of sector 1 and it also generates an increase in vacancy costs. What we have shown is that sector 2 firms correctly evaluate and internalize these consequences through the term $\beta p_1 S_1$. It is worth noting that this result is driven by the assumption of free entry which ensures that sector 2 firms correctly internalize the vacancy costs generated by the transition of their trained workers toward high-skill employment.

4.3 Educational choices

One can check that for a nil interest rate, the derivative of the social surplus with respect to the education effort has the same sign as:

$$\begin{aligned} E \equiv \quad & \left[1 - \frac{\lambda p_2}{(m+\lambda)(m+p_2)}\right]\left[\frac{p_1(y_1 - d)}{m+p_1} - \frac{m\theta_1 c}{m+p_1}\right] \\ & - \frac{m}{m+\lambda}\left[\frac{p_2}{m+p_2}(y_2 - d + \lambda\hat{S}_2) - \frac{m+\lambda}{m+p_2}\theta_2 c\right] - \frac{m}{\pi'(e)} \end{aligned} \tag{31}$$

Here also we shall state that E is equal to zero in a decentralized equilibrium. In other words, the education effort appears to be constrained efficient in the same sense as θ_2^* is. In order to verify the previous statement, let us first consider the quantity $\beta p_1 S_1$ (see B for detailed calculus). From the definition of the private surplus S_1 (see equation (6)) and from the equilibrium equation (7), we obtain (for $r = 0$):

$$\frac{p_1}{m+p_1}(y_1 - d) - \frac{m}{m+p_1}\theta_1 c = \beta p_1 S_1 \tag{32}$$

Substitution of equation (32) into equation (31) then yields:

$$E = \beta p_1 S_1 - \frac{m}{m+\lambda}\left[\frac{p_2}{m+p_2}\left(y_2 - d + \lambda\hat{S}_2 + \frac{\lambda\beta p_1 S_1}{m}\right) - \frac{m+\lambda}{m+p_2}\theta_2 c\right] - \frac{m}{\pi'(e)} \tag{33}$$

Let us now consider the quantity $\beta p_2 S_2$ (see B for detailed calculus). From the definition of the private surplus S_2 (see equation (17)) and by using equation (18), we have (for $r = 0$):

$$\frac{p_2}{m+p_2}\left[y_2 - d + \lambda\hat{S}_2 + \frac{\lambda\beta p_1 S_1}{m}\right] - \frac{m+\lambda}{m+p_2}\theta_2 c = \frac{m+\lambda}{m}\beta p_2 S_2 \tag{34}$$

Substitution of (34) into (33) finally yields:

$$E = \beta p_1 S_1 - \beta p_2 S_2 - \frac{m}{\pi'(e)}$$

This shows that in a decentralized equilibrium, the derivative of the social surplus with respect to the education effort is nil (see equilibrium equation (21)). We can write $e^* = e^C$.[2] For the same reason given in the previous section, e^* does not coincide with the social optimum e^S. It is worth noting that the constrained efficiency of educational choices also results from the condition of free entry. Because profits go to the funding of job creation, this condition makes individuals correctly internalize the social return to education.

4.4 Social optimum and decentralized equilibrium: comparison and discussion

We now examine how the decentralized equilibrium is located relative to the social optimum. Next, we provide a discussion of our welfare results by considering alternative the threat points of trained workers in wage bargaining with high-skill firms.

4.4.1 Market efficiency

A social optimum can be defined as follows:

Definition 2. *A social optimum is a set of values* $(\theta_1^S, \theta_2^S, e^S)$ *which jointly satisfy* $HS = LS = E = 0$.

The following proposition summarizes our (constrained) efficiency results:

Proposition 1. *A decentralized labor market equilibrium is constrained efficient in terms of low-skill job creation* ($\theta_2^* = \theta_2^C$) *and education effort* ($e^* = e^C$)*, but inefficient in terms of high-skill job creation* ($\theta_1^* > \theta_1^S$)*. The decentralized equilibrium does not coincide with a social optimum.*

It is worth noting that the efficiency of the tightness of sub-market 2 only holds for the (decentralized) equilibrium value of the market tightness θ_1^*. One can verify that in the absence of on-the-job search, a decentralized equilibrium would coincide with a social optimum (under the Hosios condition).

The constrained efficiency results enable us to explain why the decentralized equilibrium (Definition 1) is not a social optimum (Definition 2). Let us underline how the decentralized equilibrium is located relative to the social optimum. We already know that job creation is beyond its optimal level in sub-market 1 ($\theta_1^S < \theta_1^*$): market tightness θ_1^* is too high. What can be said about low-skill job creation and the education effort of newcomers? Under the Hosios' condition ($\eta_1 = \eta_2 = \beta$), we state the following proposition:

[2]Notice that the Hosios condition was not used in stating this point. Furthermore this result remains true whatever the workers' bargaining strength β is.

Proposition 2. *Relative to a social optimum, low-skill job creation* (θ_2^*) *is too low in a decentralized equilibrium* ($\theta_2^S > \theta_2^*$)*, and individuals' education effort* (e^*) *is too high* ($e^S < e^*$)*.*

Proof: See C.

As already shown, low-skill job creation is not optimal as θ_1^* is too high compared to its social optimum value θ_1^S. Since a significant level of high-skill job creation facilitates transitions of trained workers toward high-skill jobs, sector 2 firms create less jobs than required at the social optimum: $\theta_2^*(=\theta_2^C) < \theta_S^S$. Similarly, a significant number of sector 1 vacancies encourages newcomers to display a significant education effort (which is higher than the social optimal level) so that they apply to high-skill jobs directly. We thus have $e^*(=e^C) > e^S$.

4.4.2 Discussion

In order to reach a more precise understanding of our welfare analysis, let us consider a (very) hypothetical market structure in which the decentralized equilibrium would be efficient. First, similar to Albrecht *et al.* (2006) as well as Gautier and Wolthoff (2009), suppose that, in the bargaining with a high-skill firm, the threat point of trained workers, $\hat{U}_1$, is directly deduced from their productivity on a low-skill job.

$$r\hat{U}_1 = \hat{y}_2 - m\hat{U}_1$$

Expressed in words, when a trained worker initially matched with a low-skill firm encounters a high-skill firm, the high-skill firm (Bertrand) competes for the same worker with the low-skill firm. Consequently, the high-skill firm will win and pay a wage beyond the productivity level of the worker at the low-skill firm. This clearly requires that sector 1 firms can commit to wages.

In this case, the private surplus of a match with a trained worker in sector 1 will obviously reflect its social costs and benefits. Job creation will be efficient in sector 1. Is this sufficient to ensure the efficiency of a decentralized equilibrium? Consideration of the efficiency of job creation in sub-market 2 and of educational choices shows that this is not the case. The reason for this is that two different levels of wages would prevail in sector 1: high wages for trained on-the-job seekers having a threat point equal to $\hat{U}_1$, and lower wages for high-skill unemployed having a threat point based on U_1. The profitability of high-skill firms, and high-skill job creation, would therefore depend on the probability of recruiting one type of worker or the other. Since low-skill firms do not integrate the existence of lower-wage sector 1 workers, they underestimate the profitability of high-skill firms. This distortion implies that the gain a trained worker expects from a match with a high-skill firm no longer reflects the vacancy costs that her transition to sector 1 generates. As a result, low-skill job creation is too high compared with its social optimal value[3].

[3]See additional information which can be found in an unpublished Appendix available from the authors.

As shown in D, with this threat point, market efficiency also requires the segmentation of the high-skill sub-market according to the applicants' type (educated or trained). In the high-skill graduate sector, firms only hire workers who have successfully completed their formal education; whereas in the trainees' high-skill sector, firms only recruit trained workers coming from sub-market 2. Similar to Sattinger (2006), segmenting sub-market 1 eliminates this distortion, since each segment of the sub-market is composed of workers with identical wages[4]. As a result low-skill firms correctly evaluate the social costs and benefits of the departure of their trained workers. Segmenting the high-skill sector restores efficiency, but this market structure is not rational. As Sattinger would put it, the high-skill sub-market is "irreducible".

On the other hand, in conformity with a scenario sometimes used in the literature (Cahuc *et al.* (2006)), the workers' threat could be deduced from the lifetime utility of a low-skilled worker ($\hat{W}_2$). Under this assumption, high-skill job creation would no longer be efficient since the wage of trained sector 2 workers is lower than their productivity. High-skill firms would not internalize the entire loss of low-skill firms whose employees quit. They would therefore create too many vacancies compared to what is socially efficient. This scenario is presented in the unpublished Appendix.

Our discussion led us to consider different threat points of trained workers hired in sector 1. Assuming that a trained worker could bargain over her wage according to the low-skill firm productivity ($\hat{y}_2$) or even according to her lifetime utility while employed in a low-skill job ($\hat{W}_2$), she would obtain a higher wage than that of a worker who successfully completed her formal education. Beyond the argument presented by Shimer (2006) and Mortensen (2003), this result allows us to justify our choice in terms of a unique threat point equal to U_1 for all workers in the high-skill sector. It is indeed not reasonable to suppose that workers who failed in their formal education would obtain higher wages than educated workers.

5 Optimal public policy

The *laissez-faire* situation is not an optimum. What then should a government do? We now present a self-financed Taxes and Subsidies Policy (TSP) leading to a social optimum. The same assumptions as above have been adopted. The interest rate is equal to zero and the Hosios condition holds on both sub-markets.

Which public policy shall be implemented? From the previous analysis of inefficiency sources, we could think that the decentralized equilibrium would be efficient if firms in sub-market 1 could offer a direct transfer to firms in sub-market 2 in order to compensate for the loss high-skill firms impose on low-skill firms when their trained employees leave for sub-market 1. This relationship between firms, which would have to be imposed by the social planner, does not however restore market efficiency (the proof is available in

[4]Note that this distortion does not exist in our model as all high-skilled workers have the same threat point.

the unpublished Appendix available from the authors). Such a transfer policy raises a problem similar to the one we identified when the trained workers' threat point is deduced from their productivity on a low-skill job. Job creation becomes excessive in sector 2.

As a direct transfer between firms cannot restore labor market efficiency, the next section focuses on a more complex public policy.

5.1 Taxing sector 1

As previously highlighted, job creation is too high in sub-market 1. The government can decentralize the social optimum by implementing an appropriate fiscal policy. We now prove that in order to restore efficiency, a tax τ could be levied in sub-market 1. Thus, the value of a filled job J_1 (see equation (3)) now depends on τ:

$$(r+m)(J_1 - V_1) = y_1 - \tau - w_1 - (r+m)V_1 \tag{35}$$

By comparison between (7) and the optimal condition (28), we obtain the result that the tax would restore sub-market efficiency if it is equal to:

$$\tau = \frac{\hat{\ell}_2}{u_1 + \hat{\ell}_2}\hat{y}_2 + \frac{u_1}{u_1 + \hat{\ell}_2}d - d = \frac{\hat{\ell}_2}{\hat{\ell}_2 + u_1}(\hat{y}_2 - d) > 0 \tag{36}$$

Let α be the share, in sub-market 1 employment, of workers coming from sub-market 2:

$$\alpha = \frac{\hat{\ell}_2}{\hat{\ell}_2 + u_1}$$

The tax can therefore be written as:

$$\tau = \alpha(\hat{y}_2 - d)$$

and S_1 is now given by:

$$S_1 = \frac{y_1 - (\alpha\hat{y}_2 + (1-\alpha)d)}{m + \beta p_1}$$

Therefore, sector-1 equilibrium (7) becomes:

$$0 = -c + q_1(1-\beta)\frac{y_1 - (\alpha\hat{y}_2 + (1-\alpha)d)}{m + \beta p_1} \tag{37}$$

Equation (37) coincides with the optimality condition in sector 1 (28). With a tax τ, high-skill job creation becomes efficient. In short the Pigovian tax τ makes sector 1 firms internalize the real cost of hiring a worker coming from sector 2. However, implementing this tax does not only restore efficiency in sector 1, it also modifies efficiency results for sector 2 and for educational choices: job creation in sector 2 is no longer efficient and the same holds for the education effort.

These distortions lead to the tax τ being dedicated to the funding of two compensatory transfers. The first, denoted by σ_q is allocated to sector 2 firms when a worker quits her low-skill job. The transfer σ_q is given by:

$$\sigma_q = \frac{\tau}{m} \tag{38}$$

The second transfer, denoted by σ_e, is allocated to (entrant) workers whose education effort e is successful. The transfer σ_e is given by:

$$\sigma_e = \frac{p_1}{m + p_1}\sigma_q \tag{39}$$

Before showing that these transfers offset the distortions created by the tax τ, we need to verify that the policy is self-financing. As there are $m\pi$ workers whose effort e is successful and $p_1\hat{\ell}_2$ quit, the government's expenditures are equal to:

$$m\pi\sigma_e + p_1\hat{\ell}_2\sigma_q = \left(\frac{m\pi}{m + p_1} + \hat{\ell}_2\right) p_1\sigma_q$$

From equations (22), (23), and (38), we deduce that:

$$\left(\frac{m\pi}{m + p_1} + \hat{\ell}_2\right) p_1\sigma_q = (u_1 + \hat{\ell}_2)p_1\sigma_q = m\ell_1\sigma_q = \ell_1\tau$$

As the government's receipts are given by $(\ell_1\tau)$ per period, this shows that the government's balanced budget constraint is satisfied for this self-financed TSP.

5.2 Subsidizing Sector 2

By restoring efficiency in sector 1 one has reduced job creation in sector 2 in excess of the efficiency level. The reduction of the value of a sector 1 job does reduce the future opportunities of sector 2 employees. It thus decreases the value associated with their low-skill job. As a consequence, the private surplus associated with a filled sector 2 job is reduced. In order to restore efficiency in the overall labor market, we propose to subsidize sector 2 firms whose workers leave for sector 1. With the compensatory transfer σ_q (see equation (38)), the private surplus S_2 now satisfies (for $r = 0$):

$$\frac{(m + \lambda)(m + \beta p_2)}{m} S_2 = y_2 - d + \lambda\frac{\hat{y}_2 - d}{m + p_1} + \lambda\frac{p_1}{m + p_1}\sigma_q + \frac{\lambda}{m}\beta p_1 S_1$$

On the other hand, using equation (37) one can see that, with the tax, the quantity $(\beta p_1 S_1)$ is now given by:

$$\beta p_1 S_1 = \frac{p_1(y_1 - d) - m\theta_1 c}{m + p_1} - \frac{p_1}{m + p_1} m\sigma_q \tag{40}$$

Combining the two previous equations yields:

$$\frac{(m + \lambda)(m + \beta p_2)}{m} S_2 = y_2 - d + \lambda\frac{\hat{y}_2 - d}{m + p_1} + \frac{\lambda}{m}\frac{p_1(y_1 - d) - m\theta_1 c}{m + p_1}$$

Substituting S_2 into equation (18) shows that transfer σ_q enables the restoration of the efficiency of low-skill vacancy creation (see equation (29) for $\beta = \eta_2$).

5.3 Rewarding educational success

With the reward σ_e defined by equation (39), the private optimality condition (20) has to be rewritten as follows (for $r = 0$):

$$\pi'(e)\left(U_1 + \frac{p_1}{m+p_1}\sigma_q - U_2\right) - 1$$

or

$$\pi'(e)\left(\beta p_1 S_1 + \frac{mp_1}{m+p_1}\sigma_q - \beta p_2 S_2\right) = m$$

Using equation (40), the previous equation can be rewritten as:

$$\left(\frac{p_1}{m+p_1}(y_1 - d) - \frac{m}{m+p_1}\theta_1 c - \frac{mp_1}{m+p_1}\sigma_q + \frac{mp_1}{m+p_1}\sigma_q - \beta p_2 S_2\right) - \frac{m}{\pi'(e)} = 0$$

or

$$\left(\frac{p_1}{m+p_1}(y_1 - d) - \frac{m}{m+p_1}\theta_1 c - \beta p_2 S_2\right) - \frac{m}{\pi'(e)} = 0$$

As the efficiency of θ_2 is restored (despite the tax), $\beta p_2 S_2$ remains equal to:

$$\frac{m}{m+\lambda}\left[\frac{p_2}{m+p_2}\left(y_2 - d + \lambda\hat{S}_2 + \frac{\lambda\beta p_1 S_1}{m}\right) - \frac{m+\lambda}{m+p_2}\theta_2 c\right]$$

Therefore the social optimality condition (33) holds true.

At first glance, the idea of rewarding educational success might look counterintuitive, since one could point out that, according to Proposition 2, education effort is lower in a social optimum than in a decentralized equilibrium. The reason for this is that without subsidies, private educational choices are no longer efficient for the optimum value of submarket tightness θ_1^S (computed with the tax τ). In the absence of a reward, the return to education (the opportunity to get a "good" job) would be too weak, thus leading to a reduction in formal education below its optimal level. The reward compensates for this effect.

The following proposition summarizes the above results:

Proposition 3. *With the TSP ($\tau; \sigma_e; \sigma_q$) the decentralized equilibrium is a social optimum ($\theta^*_{1(TSP)} = \theta_1^S, \theta^*_{2(TSP)} = \theta_2^S, e^*_{(TSP)} = e^S$).*

6 Calibration

In this section we verify that the previous TSP restores market efficiency through a calibration of the model using U.S. labor market evidence. We therefore assume that the Hosios condition holds and that the discount rate is nil ($r = 0$). We consider that the low-skill group is constituted of individuals whose educational attainment is a high-school diploma with no college attendance, or less. The high-skill group is therefore composed of individuals with some college or associate degree, plus all college graduates.

6.1 Baseline scenario

The unemployment rates presented in the theoretical model have to be understood as the unemployment of young workers only, since the model does not compute any job separation rate. For this reason we match the unemployment rates of the calibration with that of educated and non-educated young workers aged 16 to 24 years old for 2006, given by the U.S. Bureau of Labor Statistics (BLS). The rates are computed as the youth unemployed population divided by the total labor force[5]. We obtain from the same source a share of educated workers in the total labor force π of about 42%. The probability of leaving permanently the labor market (dying) is fixed at $m = 0.0018$ which represents about 42.7 years in the labor force (Ljungqvist and Sargent 1998). Equations (22) and (24) can then be used to retrieve the average monthly job finding rates p_1 and p_2. Table 2 gives the fixed values of variables and parameters and table 3 gives the inferred values.

Table 2: Parameters and variables whose value is fixed

Name	Description	Value	Source
β	worker's bargaining power	0.5	Pissarides (2000)
η_1	elasticity sub-market 1	0.5	Pissarides (2000)
η_2	elasticity sub-market 2	0.5	Pissarides (2000), Hosios (1990)
m	permanent exist rate	0.0018	Ljungqvist and Sargent (1998)
M_1	matching parameter 1	0.1020	Hagedorn *et al.* (2010)
M_2	matching parameter 2	0.1640	Hagedorn *et al.* (2010)
y_1	high-skill productivity	100	*normalized*
y_2	low-skill prod. (untrained)	50.39	Hagedorn *et al.* (2010)
d	non-market output	35.77	Hall and Milgrom (2008), *see text*
π	share of educated in LF	0.4200	BLS, year 2006
u_1	high-skill youth unemployment	0.0116623	BLS, year 2006
u_2	low-skill youth unemployment	0.0162453	BLS, year 2006
AHL	low/high-skill transition share	0.20	Andersson *et al.* (2005)

[5]That is to say - educated (for u_1) or non-educated (for u_2) - youth unemployment rates times the share of the youth labor force in overall labor force.

As a standard feature, bargaining power β and the elasticities η_1 and η_2 are assumed to be equal to 0.5 (the Hosios condition holds). This choice of calibration will allow us to observe the impact on the labor market structure of a TSP as proposed in the theoretical part of this work. High-skill productivity y_1 is normalized to 100.

We assume the matching functions to be Cobb-Douglas with matching parameters M_1 and M_2, so that $h_1 = M_1 v_1^{\eta_1}(u_1 + \hat{\ell}_2)^{1-\eta_1}$ and $h_2 = M_2 v_2^{\eta_2} u_2^{1-\eta_2}$ are the matching function of respectively the high-skill sub-market and the low-skill sub-market. The matching parameters values are set according to Hagedorn *et al.* (2010) who present the calibration of a search-matching model with the same two subgroups (high-skill and low-skill individuals) using data from the CPS between 1979 and 2006. We can then deduce the tightness of each sub-market θ_1 and θ_2 as well and the probability that a firm fills a job q_1 and q_2.

Hagedorn *et al.* (2010) give a marginal product ratio of 1.9846 between skilled and unskilled workers. We use this ratio to retrieve the value for y_2. According to Hall and Milgrom (2008) non-market activities, which includes the value of leisure, represent a fraction 0.71 of the average productivity. In order to minimize the presence of potential unemployment benefits, which are not taken into account in our modeling, we chose to fix the value of d as a proportion of y_2.

Table 3: Parameters and variables whose value is inferred

Name	Description	Value	Source
$\hat{y}_2$	low-skill prod. (trained)	85	*see text*
p_1	high-skill job transition rate	0.06302422	equation (22)
p_2	low-skill job transition rate	0.06246457	equation (24)
θ_1	tightness sector 1	0.38178124	matching function definition
θ_2	tightness sector 2	0.14507072	matching function definition
$\hat{\ell}_1$	low-skilled leaving for a high-skill job	0.1160	*see text*
ℓ_2	low-skill employ. (untrained)	0.44444164	equations (23), (25), (26), (41)
$\hat{\ell}_2$	low-skill employ. (trained)	0.00331301	equations (23), (25), (26), (41)
ℓ_1	high-skill employment	0.52433769	equations (23), (25), (26), (41)
λ	training transition rate	0.00048322	equations (23), (25), (26), (41)
c_1	vacancy cost sub-market 1	1.591335268E+2	equation (7)
c_2	vacancy cost sub-market 2	1.607701929E+2	equation (18)
$\pi'(e)$	marginal success rate of education	0.000048088	equation (21)
e	education effort	1.206115E+4	*see text*

The probability λ of upgrading the skills of low-skilled workers is obtained knowing that the share of overall low-skilled workers in the labor force must be equal to the following equation:

$$(1 - \pi) = u_2 + \ell_2 + \hat{\ell}_2 + \hat{\ell}_1 \tag{41}$$

Where $\hat{\ell}_1$ is the share of low-skilled workers who succeed in obtaining a high-skill job. According to Andersson *et al.* (2005) between 15% and 20% of workers with a low level of education had escaped low wage employment after nine years. We use this proportion to estimate $\hat{\ell}_1 = AHL \times (1 - \pi)$ where $AHL = 0.2$ is the proportion of low-skilled workers escaping from low-skill jobs. We then deduce the training transition rate λ and the employment levels ℓ_2, $\hat{\ell}_2$, and ℓ_1 from equations (23), (25), (26), and (41).

Regarding the education technology, we retain the specification $\pi(e) = \frac{e-\phi}{e}$. Knowing the baseline value for π and $\pi'(e)$ enables us to retrieve the parameter ϕ and the value of education effort e.

The remaining parameters are the cost of a vacancy, and the trained workers' productivity $\hat{y}_2$. In order to ensure an equilibrium in each sub-market, we assume from now on the existence of two different vacancy costs depending on the sub-market, so that c_i is the cost of vacancy in sector i. Finally, we arbitrarily fix a value for productivity $\hat{y}_2$ of 0.85 times the high-skill output (sensitivity test could be performed on the basis of this parameter but such a test goes beyond the scope of this work). This choice ensures a progressiveness in the workers' productivity ($y_2 < \hat{y}_2 < y_1$). A trained worker from sub-market 2 will indeed search on-the-job for a high-skill job rather than choosing high-skill unemployment (this implies $\hat{y}_2 < d$), and low-skill wages will remain below high-skill wages even for trained workers (this implies $\hat{y}_2 < y_1$). Using the preceding values in equations (7) and (18) gives the values for c_i.

6.2 The Tax and Subsidy Policy (TSP)

The TSP requires the imposition of a tax τ on the output of high-skill firms, a subsidy σ_q to low-skill firms whose workers quit for a high-skill job, and a reward σ_e for educational success. Table 4 synthesizes the values of variables at the baseline and after the TSP has been implemented.

As expected, the decentralized equilibrium is not a social optimum since $HS < 0$. The tax imposed on high-skill firms that is required to restore a social optimum creates a distortion which is restored by the implementation of a subsidy to low-skill firms, whose workers quit for high-skill jobs, in addition with a reward for educational success.

The calibration of the policy that would restore the social optimum gives us the opportunity of studying the consequences of such a policy for different variables in the economy. We can appraise the importance of the tax and the subsidies required to restore the efficiency based on US empirical evidence. The tax that should be implemented represents about 10.9% of high-skill firms' output, and provokes a reduction in the surplus (S_1) of about 9.1%. The reason why such a tax is necessary to lower the over-creation of high-skill vacancies and to restore efficiency in sub-market 1 lies in the strong differential between d (35.77) and $\hat{y}_2$ (85) in our baseline scenario.

Since the product of the tax is split among a small number of agents, the subsidies distributed are important. This creates a strong increase in the surplus of low-skill firms whose employees are trained ($\hat{S}_2$). The increase of this surplus is almost compensated by the reduction of the high-skill surplus (small rise of S_2), hence job creation rises only in

Table 4: Impact of the TSP

Variable	Description	Baseline	TSP value
p_1	high-skill job transition rate	0.06302422	0.05727193
p_2	low-skill job transition rate	0.06246457	0.06249696
q_1	high-skill job filling rate	0.16507940	0.18165966
q_2	low-skill job filling rate	0.43058012	0.43035694
θ_1	tightness sub-market 1	0.38178124	0.31527051
θ_2	tightness sub-market 2	0.14507072	0.14522122
v_1	vacancy sub-market 1	0.00571730	0.00517869
v_2	vacancy sub-market 2	0.00235672	0.00235923
e	education effort	1.206115E+4	1.205476E+4
π	share of educated in LF	0.42000000	0.41969280
u_1	high-skill youth unemployment	0.01166231	0.01278860
u_2	low-skill youth unemployment	0.01624534	0.01624576
ℓ_2	low-skill employment (untrained)	0.44444164	0.44468365
$\hat{\ell}_2$	low-skill employment (trained)	0.00331301	0.00363760
$\hat{\ell}_1$	high-skill employment (trained)	0.11600000	0.11574020
ℓ_1	high-skill employment (educated)	0.52433769	0.52264440
S_2	match surplus sub-market 2 (untrained)	7.46760879E+2	7.471481339E+2
$\hat{S}_2$	match surplus sub-market 2 (trained)	7.59354001E+2	1.301303E+4
S_1	match surplus sub-market 1	1.927963E+3	1.751996E+3
Σ	social surplus	53.1095346	53.1170959
HS	efficiency high-skill	-0.8988586	0
LS	efficiency low-skill	0	0
E	efficiency education effort	0	0
τ	tax on sector 1 firms	-	10.9008278
σ_q	subsidy to firms whose workers quit	-	6.056015E+3
σ_e	reward to educational success	-	5.871481E+3

the low-skill sub-market. In the same way, the impact of the large size of the subsidy allocated to individuals as a reward for educational success is offset by the reduction of the attractiveness of the high-skill sub-market. The education effort (e), too high with regard to the social optimum in the baseline scenario, lowers slightly after the TSP is implemented. The share of high-skill newcomers (π) follows the same path.

Despite a reduction in the number of high-skilled workers, the decline of high-skill job creation leads to a reduction in high-skill tightness, as well as to a small increase in unemployment. All in all, the unemployment level of young workers increases slightly after the imposition of the TSP. We also observe a reduction of 0.32% in the most productive jobs in the economy, together with a limited rise in low-skill employment (+9.8% for trained workers and +0.05% for untrained workers). Those observations should count against the improvement of the efficiency of the labor market. However, the rise in the total surplus of the economy (social surplus) is made possible through the reduction of the social cost of high-skill vacancies ($\theta_1(u_1 + \hat{\ell}_2)c_1$ and $\theta_2 u_2 c_2$) together with a reduction in the social cost of education (me).

7 Conclusion

In many countries governments subsidize formal education and/or training through different channels. For example, the first Clinton administration made skill upgrading a major priority. Are these subsidies justified?

We set up a model in which workers face a tradeoff between acquiring formal education (thus having the opportunity to obtain a good job directly) and learning-by-doing in a low-skill job, then searching (while on-the-job) for a high-skill job.

We have stated that workers do not choose the appropriate amount of formal education when faced with this tradeoff. Even if the decentralized equilibrium is constrained efficient in terms of low-skill job creation and educational choices, it is inefficient in terms of high-skill job creation. Because high-skill job creation is too high and induces hold-up behavior which penalizes low-skill jobs, a tax must be levied on high-skill firms. Therefore, educational choices and low-skill job creation will no longer be constrained efficient. A self-financing Tax and Subsidy Policy restores market efficiency. The tax should finance two compensatory transfers: a subsidy to low-skill jobs (whose workers quit), and a reward aiming at encouraging education effort. In a decentralized equilibrium, workers tend to put too much stress on formal education. However, according to our results, subsidies to education make sense, even without credit constraint, since the tax levied on high-skill firms leads to an inefficient amount of effort devoted to formal education.

We assessed the impacts of such a theoretical scenario through a calibration of the model with US empirical evidence. From a practical viewpoint, our results have to be treated with caution, as one could object that this tax/transfer policy would be difficult to implement when the government does not have perfect knowledge of each firm's output.

It is worth noting that, empirically, the path along which low-skilled workers advance to high-skill jobs differs between sectors in the economy. Even if most low-skill transitions require a movement from one firm to another in sectors such as in manufacturing, industrial crafts, construction, and warehousing, internal promotions is more usual in some sectors or occupations. This is for instance the case for the French transport industry (Alonzo and Chardon 2006). Such a case cannot be taken into account in our framework since each firm owns only one job.

It could be objected that the nature of the job could shift from low-skill characteristics to high-skill ones when the worker becomes trained. This argument would however deny the existence of reorganization costs faced by the firm during this shift. If these costs are too high, such a job transformation becomes unthinkable, hence nullifying the objection. In addition, high-skill firms would not internalize the fact that opening additional vacancies in sector 1 forces low-skill firms to bear reorganization costs. As a consequence, job creation in the high-skill sub-market is likely to remain excessive.

One could also argue that in the presence of a reward allocated to graduates, the educational system could be prompted to award more diplomas in exchange for an increase in the level of tuition given. This phenomenon would affect the skill level of graduates.

To conclude, we would like to emphasize an unexpected result of our study. Economists usually believe that the share of educated workers would not be sufficiently significant when there are search-frictions. Search-frictions create rents, which implies that a part of the return to education goes to firms. Therefore the incentive for workers to invest in their formal education would be too low. We have shown that this widespread view may be incorrect. Without on-the-job search, educational choices are perfectly efficient, despite search-frictions. The reason for this is that firms have to pay for the cost of creating jobs that educated workers will hold. As already mentioned, in our setting the inefficiency of the education effort in the *laissez-faire* situation does not come from private educational choices, but from the presence of the hold-up phenomenon.

A Optimality condition for low-skill jobs

The derivative of the social surplus (equation (27)) with respect to the tightness of sub-market 2 can be written as:

$$\begin{aligned}
\frac{\partial \Sigma}{\partial \theta_2} &= q_2(1-\eta_2)\Bigg[\frac{\lambda p_1}{m(m+p_1)}\frac{m^2(1-\pi)}{(m+\lambda)(m+p_2)^2}y_1+\frac{m^2(1-\pi)}{(m+\lambda)(m+p_2)^2}y_2 \\
&+\frac{\lambda}{(m+p_1)}\frac{m^2(1-\pi)}{(m+\lambda)(m+p_2)^2}\hat{y}_2-\frac{m(1-\pi)}{(m+p_2)^2}d-\frac{\lambda}{(m+p_1)}\frac{m^2(1-\pi)}{(m+\lambda)(m+p_2)^2}\theta_1 c \\
&+\frac{m(1-\pi)}{(m+p_2)^2}\theta_2 c\Bigg]-\frac{m(1-\pi)}{(m+p_2)}c \\
&= \frac{m^2(1-\pi)}{(m+\lambda)(m+p_2)^2}\Bigg[q_2(1-\eta_2)\Bigg(\frac{\lambda p_1}{m(m+p_1)}y_1+y_2+\frac{\lambda}{(m+p_1)}\hat{y}_2-\frac{(m+\lambda)}{m}d \\
&-\frac{\lambda}{m+p_1}\theta_1 c+\frac{m+\lambda}{m}\theta_2 c\Bigg)-c\frac{(m+p_2)(m+\lambda)}{m}\Bigg] \\
&= \frac{m^2(1-\pi)}{(m+\lambda)(m+p_2)^2}\Bigg[q_2(1-\eta_2)\Bigg(y_2-d+\frac{\lambda}{m+p_1}(\hat{y}_2-d)-\frac{\lambda p_1}{m(m+p_1)}d \\
&+\frac{\lambda[p_1 y_1-\theta_1 cm]}{m(m+p_1)}\Bigg)-c(m+\eta_2 p_2)\frac{m+\lambda}{m}\Bigg]
\end{aligned}$$

As the first term is always positive, the sign of the optimality condition $\frac{\partial \Sigma}{\partial \theta_2}$ is given by:

$$LS \equiv q_2(1-\eta_2)\left(y_2-d+\frac{\lambda(\hat{y}_2-d)}{m+p_1}+\frac{\lambda}{m(m+p_1)}[p_1(y_1-d)-\theta_1 cm]\right)-c(m+\eta_2 p_2)\frac{m+\lambda}{m}$$

B "Constrained" efficiency of educational choices

We show that the optimality condition for educational choices (31) is equal to zero in a decentralized equilibrium.

Let us consider the quantity $\beta p_1 S_1$. From equation (6), we deduce (for $r=0$):

$$\beta p_1 S_1 = \frac{p_1}{m+p_1}(y_1-d)+\frac{m}{m+p_1}(y_1-d)-mS_1$$

By using equilibrium equation (7), we obtain:

$$\frac{m}{m+p_1}(y_1-d)-mS_1 = -\frac{m}{m+p_1}p_1(1-\beta)S_1 = -\frac{m}{m+p_1}\theta_1 c$$

The result is:

$$\frac{p_1}{m+p_1}(y_1-d)-\frac{m}{m+p_1}\theta_1 c = \beta p_1 S_1$$

Substitution of the previous equation into equation (31) then yields:

$$E = \beta p_1 S_1 - \frac{m}{m+\lambda}\left[\frac{p_2}{m+p_2}\left(y_2-d+\lambda\hat{S}_2+\frac{\lambda\beta p_1 S_1}{m}\right)-\frac{m+\lambda}{m+p_2}\theta_2 c\right]-\frac{m}{\pi'(e)}$$

Let us consider the quantity $\beta p_2 S_2$. From equation (17), we deduce (for $r = 0$):

$$\frac{m+\lambda}{m}\beta p_2 S_2 = y_2 - d + \lambda \hat{S}_2 + \frac{\lambda \beta p_1 S_1}{m} - (m+\lambda) S_2$$

The latter equation can be rewritten as follows:

$$\begin{aligned} \frac{m+\lambda}{m}\beta p_2 S_2 = \quad & \frac{p_2}{m+p_2}\left[y_2 - d + \lambda \hat{S}_2 + \frac{\lambda \beta p_1 S_1}{m}\right] \\ & + \frac{m}{m+p_2}\left[y_2 - d + \lambda \hat{S}_2 + \frac{\lambda \beta p_1 S_1}{m}\right] - (m+\lambda) S_2 \end{aligned}$$

By using equilibrium equation (18), we obtain:

$$\begin{aligned} \frac{m}{m+p_2}\left[y_2 - d + \lambda \hat{S}_2 + \frac{\lambda \beta p_1 S_1}{m}\right] - (m+\lambda) S_2 \quad &= \quad \frac{m+\lambda}{m+p_2}(m+\beta p_2) S_2 - (m+\lambda) S_2 \\ &= \quad -\frac{m+\lambda}{m+p_2} p_2 (1-\beta) S_2 = -\frac{m+\lambda}{m+p_2}\theta_2 c \end{aligned}$$

We thus have:

$$\frac{p_2}{m+p_2}\left[y_2 - d + \lambda \hat{S}_2 + \frac{\lambda \beta p_1 S_1}{m}\right] - \frac{m+\lambda}{m+p_2}\theta_2 c = \frac{m+\lambda}{m}\beta p_2 S_2$$

Substitution of the previous equation into (33) finally yields:

$$E = \beta p_1 S_1 - \beta p_2 S_2 - \frac{m}{\pi'(e)}$$

The optimality condition (31) coincides with the decentralized equilibrium (21).

C Proof of proposition 2

First part of proposition 2

Proof. We first prove that tightness θ_2^* is too low relative to the social optimum. Let us consider the system composed of the following two equations in (θ_1, θ_2):

$$(1-\eta) q_1 [y_1 - d - a(\hat{y}_2 - d)] - (m + \eta p_1) c = 0 \tag{42}$$

$$\begin{aligned} & q_2 (1-\eta)(y_2 - d) - \frac{m+\lambda}{m} c \eta p_2 - (m+\lambda) c = 0 \\ & + q_2 (1-\eta) \frac{\lambda}{m}\left[\frac{p_1}{m+p_1}(y_1 - d) + \frac{m}{m+p_1}(\hat{y}_2 - d) - \frac{m}{m+p_1}\theta_1 c\right] \end{aligned} \tag{43}$$

This system is parameterized with the scalar a which takes its values in the interval $[0, a^\sigma]$. The limit a^σ is the value of ratio $\frac{\hat{\ell}_2}{u_1 + \hat{\ell}_2}$. Equation (43) is obtained by equalizing LS (defined by (29)) to zero, which is the (necessary) efficiency condition for tightness θ_2.

Let us consider (42). This equation (implicitly) determines tightness θ_1 as a function in parameter a, denoted by $\theta_1(a)$. For $a = 0$, this equation describes the decentralized equilibrium. The social optimum is obtained for $a = a^\sigma$.

As $q_1'(\theta_1) < 0$ and $p_1'(\theta_1) > 0$, parameter a has a negative impact on tightness θ_1 ($\theta_1'(a) < 0$). In accordance with proposition 1, we find that the decentralized equilibrium value of tightness θ_1^* is greater than its optimal value: $\theta_1(a^\sigma)(= \theta_1^S) < \theta_1(0)(= \theta_1^*)$.

The system composed of equations (42) and (43) also determines tightness θ_2 as an implicit function in a, $\theta_2(a)$. Let us consider equation (43). Its left hand side, LS, can be written as:

$$LS(.) = LS(\theta_1(a), \theta_2)$$

In order to deduce the impact of parameter a on tightness θ_2, we must sign the partial derivative of $LS(.)$ with respect to θ_1 and θ_2. Let us note $H(\theta_1)$ the term of equation (43) between brackets. The derivative of $H(\theta_1)$ with respect to θ_1 has the same sign as:

$$-q_1(1-\eta)(1-a)(\hat{y}_2 - d)$$

As the function $LS(.)$ is bounded by the social optimum ratio $\frac{\hat{\ell}_2}{u_1+\hat{\ell}_2}$, we know that $a < 1$. We therefore have $H'(\theta_1) < 0$. The partial derivative of $LS(.)$ with respect to θ_1 is thus negative.

Let us consider the impact of θ_2 on $LS(.)$. One must first determine the sign of $H(\theta_1)$, where $H(\theta_1)$ can be rewritten as:

$$\begin{aligned} H &= \frac{1}{m+p_1}\left[p_1(y_1 - d) + m(\hat{y}_2 - d) - m\theta_1 c\right] \\ &= \frac{1}{m+p_1}\left[\theta_1\left((m+\eta p_1)q_1\frac{y_1 - d - a(\hat{y}_2 - d)}{m+\eta p_1} - mc\right) + (m + ap_1)(\hat{y}_2 - d)\right] \end{aligned}$$

Knowing that (42) is equivalent to:

$$mc = m(1-\eta)q_1\frac{y_1 - d - a(\hat{y}_2 - d)}{m+\eta p_1}$$

We obtain by substitution:

$$H = \frac{1}{m+p_1}\left[\eta(m+p_1)p_1\frac{y_1 - d - a(\hat{y}_2 - d)}{m+\eta p_1} + (m+ap_1)(\hat{y}_2 - d)\right] > 0$$

$H(\theta_1) > 0$ and $q_2'(\theta_2) < 0$, therefore the partial derivative of $LS(.)$ with respect to θ_2 is (strictly) negative (see equation (43)).

The impact of parameter a on θ_2 is obtained by equalizing to zero the differential of $LS(.)$:

$$dLS(.) = \frac{\partial LS}{\partial \theta_1}\theta_1'(a)da + \frac{\partial LS}{\partial \theta_2}d\theta_2 = 0$$

where

$$\frac{d\theta_2}{da} = -\frac{\frac{\partial LS}{\partial \theta_1}\theta_1'(a)}{\frac{\partial LS}{\partial \theta_2}} > 0$$

The derivative of the implicit function $\theta_2(a)$ is thus (strictly) positive. As a consequence the optimal value of θ_2^S is higher than its decentralized equilibrium value $(\theta_2(a^\sigma)(=\theta_2^S) > \theta_2(0)(=\theta_2^*))$. This proves the first part of proposition 2. □

Second part of proposition 2

Proof. We now show that the education effort e^* is too high relative to the social optimum e^S. The necessary optimality condition for effort e (equation (31)) can be written as:

$$E = \frac{p_1(y_1 - d) - m\theta_1 c}{m + p_1} - \frac{m}{m+\lambda}\left[\frac{p_2}{m+p_2}\left(y_2 - d + \frac{\lambda}{m}H\right) - \frac{m+\lambda}{m+p_2}\theta_2 c\right] - \frac{m}{\pi'(e)} = 0 \tag{44}$$

where p_i and θ_i (for $i = 1, 2$), as well as $H(\theta_1)$, are deduced from equations (42) and (43).

Taking (42) and (43) into account, equation (44) determines e as an implicit function $e(a)$ of parameter a. The decentralized equilibrium value of e is given by $e(0)$ and its optimal value by $e(a^\sigma)$. We must sign the derivative of $e(a)$. The left side term, E, of (44) can be written as:

$$E(.) = E(\theta_1(a), \theta_2(a), e)$$

For a given H, the derivative of $E(.)$ with respect to θ_1 has the same sign as:

$$(1-\eta)q_1(y_1 - d) - (m + \eta p_1)c$$

We deduce from equation (42) that the previous equation is (strictly) positive for $a > 0$. As the term H is a decreasing function of θ_1, we thus deduce that the derivative of $E(.)$ with respect to θ_1 is (strictly) positive.

The derivative of $E(.)$ with respect to θ_2 has the same sign as:

$$-q_2(1-\eta)\left(y_2 - d + \frac{\lambda}{m}H\right) + \frac{m+\lambda}{m}(m + \eta p_2)c$$

We deduce from equation (43) that the previous equation is nil. The derivative of $E(.)$ with respect to θ_2 is therefore nil.

The impact of parameter a on education effort $e(a)$ is obtained by equalizing the differential $E(.)$ to zero:

$$dE(.) = \frac{\partial E}{\partial \theta_1}\theta_1'(a)da + \frac{\partial E}{\partial \theta_2}\theta_2'(a)da + \frac{m}{\pi'(e)^2}\pi"(e)de = 0$$

where

$$\frac{de}{da} = -\frac{\frac{\partial E}{\partial \theta_1}\theta_1'(a)}{\frac{m}{\pi'(e)^2}\pi"(e)}$$

The function $\pi(.)$ being concave, the derivative of the implicit function $e(a)$ is thus (strictly) negative. As a consequence, the optimal value of education effort e^S is lower than its decentralized equilibrium value $(e(a^\sigma)(= e^S) < e(0)(= e^*))$. This proves the second part of proposition 2. □

D Efficient outcome: high-skill segmentation with low-skill productivity as a threat point for high-skilled trainees

In order to identify and compare the sources of inefficiency presented in our model, we present a situation which would be efficient, even if in many respects unrealistic.

A segmented high-skill sub-market

In a segmented high-skill sub-market, some firms open vacancies only to workers who have succeeded in their formal education process, while the other high-skill firms open vacancies only to workers who learned "by doing" in sub-market 2. For the sake of simplicity we will refer to the sectors respectively as the graduates' and the trainees' high-skill sectors, to distinguish between these two high-skill sub-markets. The productivity is assumed to be identical for all high-skill firms.

Graduates high-skill sector

In this sub-sub-market, the tightness is equal to $\theta_1 = \frac{v_1}{u_1}$. The arrival rate of job offers to unemployed graduates is $p_1 = p_1(\theta_1)$, and the arrival rate of workers to vacancies is $q_1 = q_1(\theta_1)$. The asset values of a firm whose job is vacant or filled, the lifetime utility of a high-skilled unemployed graduate, and that of a graduate employee are given by:

$$rV_1 = -c + q_1 J_1$$
$$(r+m)(J_1 - V_1) = y_1 - w_1 - rV_1$$
$$(r+m)U_1 = d + p_1(W_1 - U_1)$$
$$(r+m)(W_1 - U_1) = w_1 - (r+m)U_1$$

The static Nash bargaining $(W_1 - U_1) = \beta S_1$, and the assumption of free-entry $V_1 = 0$, gives the following surplus:

$$(r+m)S_1 = y_1 - (r+m)U_1 \quad \Leftrightarrow (r+m+\beta p_1)S_1 = y_1 - d$$

The tightness θ_1^* is thus determined by the following equilibrium equation:

$$c = q_1(1-\beta)S_1$$

Trainees high-skill sector

In this sub-sub-market, the tightness is equal to $\hat{\theta}_1 = \frac{\hat{v}_1}{\hat{\ell}_2}$. The arrival rate of sector 1 job offers to trainees is $\hat{p}_1 = \hat{p}_1(\hat{\theta}_1)$, and the arrival rate of workers to vacancies is $\hat{q}_1 = \hat{q}_1(\hat{\theta}_1)$. The asset values of a firm whose job is vacant or filled are the following:

$$r\hat{V}_1 = -c + \hat{q}_1 \hat{J}_1$$

$$(r+m)(\hat{J}_1 - \hat{V}_1) = y_1 - \hat{w}_1 - r\hat{V}_1$$

When a trained sector 2 employee meet a sector 1 firm, her current sector 2 employer tries to retain her, and so prevent a zero asset value, by offering her the highest possible wage. This proposed wage is thus equal to the employee's productivity $\hat{y}_2$. The threat point of the worker in Nash bargaining with the sector 1 firm will therefore be her sector 2 lifetime utility determined by such a wage:

$$r\tilde{W}_2 = \hat{y}_2 + p_1(\hat{W}_1 - \tilde{W}_2) - m\tilde{W}_2$$

With free-entry $\hat{V}_1 = 0$, the surplus is then given by:

$$(r + m + \beta\hat{p}_1)\hat{S}_1 = y_1 - \hat{y}_2$$

and the job creation equation which determines $\hat{\theta}_1^*$ is:

$$c = \hat{q}_1(1-\beta)\hat{S}_1$$

The low-skill sector

As usual, the lifetime utility of unemployed and employed untrained low-skilled workers can be written as:

$$(r+m)U_2 = d + p_2(W_2 - U_2)$$

$$(r+m)W_2 = w_2 + \lambda(\hat{W}_2 - W_2)$$

With Nash bargaining, we have:

$$(r+m)U_2 = d + p_2\beta S_2$$

At the end of the learning-by-doing process, the worker can apply to a trainee high-skill job. She therefore negotiates her wage with her current sector 2 employer on the basis of her productivity, $\hat{y}_2$, but taking into account her ability to apply for a high-skill job. Her threat point is therefore the lifetime utility of an unemployed worker in sector 1, denoted by $\breve{U}_1$.

$$(r+m)\breve{U}_1 = d + \beta\hat{p}_1\breve{S}_1$$

where $\breve{S}_1$ is the surplus obtained by the hiring of an unemployed in a high-skill job. We have:

$$(r + m + \beta\hat{p}_1)\breve{S}_1 = y_1 - d$$

This lifetime utility and this surplus are only potential since there is no sector 1 unemployment in the trainees' high-skill sector. We can nevertheless write:

$$(r+m+\lambda)W_2 = w_2 + \lambda(\hat{W}_2 - \breve{U}_1) + \lambda\breve{U}_1$$

where $(\hat{W}_2 - \breve{U}_1)$ is the surplus of a trained sector 2 worker. We have:

$$(\hat{W}_2 - \breve{U}_1) = \beta\hat{S}_2$$

The lifetime utility of a low-skilled worker holding a sector 2 job satisfies:

$$(r+m+\lambda)(W_2 - U_2) = w_2 + \lambda(\hat{W}_2 - \breve{U}_1) + \lambda\breve{U}_1 - (r+m+\lambda)U_2$$

And the asset value of a filled job is given by:

$$rJ_2 = y_2 - w_2 - m(J_2 - V_2) + \lambda(\hat{J}_2 - J_2)$$

$$\Leftrightarrow (r+m+\lambda)J_2 = y_2 - w_2 + \lambda(1-\beta)\hat{S}_2$$

The surplus of a sector 2 job filled with a newcomer can be written as:

$$(r+m+\lambda)S_2 = y_2 + \lambda\hat{S}_2 + \lambda\breve{U}_1 - (r+m+\lambda)U_2$$

$$\Leftrightarrow \frac{(r+m+\lambda)(r+m+\beta p_2(\theta_2))}{r+m}S_2 = y_2 + \lambda\hat{S}_2 - d + \frac{\lambda}{r+m}\beta\hat{p}_1\breve{S}_1$$

At the end of the learning-by-doing process, the value of a low-skill firm becomes:

$$r\hat{J}_2 = \hat{y}_2 - \hat{w}_2 - (m+\hat{p}_1)(\hat{J}_2 - V_2)$$

$$\Leftrightarrow (r+m+\hat{p}_1)(\hat{J}_2 - V_2) = \hat{y}_2 - \hat{w}_2 - rV_2$$

While for the worker who still has no contact with a high-skill firm, we have:

$$r\hat{W}_2 = \hat{w}_2 + \hat{p}_1(\hat{W}_1 - \hat{W}_2) - m\hat{W}_2$$

$$\Leftrightarrow (r+m+\hat{p}_1)(\hat{W}_2 - \breve{U}_1) = \hat{w}_2 + \hat{p}_1(\hat{W}_1 - \hat{W}_2) + \hat{p}_1(\hat{W}_2 - \breve{U}_1) - (r+m)\breve{U}_1$$

The surplus becomes:

$$\begin{aligned}
(r+m+\hat{p}_1)\hat{S}_2 &= \hat{y}_2 - d + \hat{p}_1(\hat{W}_1 - \hat{W}_2) + \hat{p}_1(\hat{W}_2 - \breve{U}_1) - \beta\hat{p}_1\breve{S}_1 \\
\Leftrightarrow (r+m+\hat{p}_1)\hat{S}_2 &= \hat{y}_2 - d + \hat{p}_1(\hat{W}_1 - \tilde{W}_2) + \hat{p}_1(\tilde{W}_2 - \breve{U}_1) - \beta\hat{p}_1\breve{S}_1 \\
\Leftrightarrow (r+m+\hat{p}_1)\hat{S}_2 &= \hat{y}_2 - d + \beta\hat{p}_1\hat{S}_1 + p_1(\tilde{W}_2 - \breve{U}_1) - \beta\hat{p}_1\breve{S}_1
\end{aligned}$$

As $(r+m)(\tilde{W}_2 - \breve{U}_1) = \hat{y}_2 - d + \beta\hat{p}_1(\hat{S}_1 - \breve{S}_1)$, we have:

$$(r+m+\hat{p}_1)(r+m)\hat{S}_2 = (\hat{y}_2 - d)(r+m) + (r+m)\beta\hat{p}_1(\hat{S}_1 - \breve{S}_1) + \hat{p}_1(\hat{y}_2 - d + \beta\hat{p}_1(\hat{S}_1 - \breve{S}_1))$$

$$\Leftrightarrow (r+m)\hat{S}_2 = (\hat{y}_2 - d) + \beta\hat{p}_1(\hat{S}_1 - \breve{S}_1)$$

Knowing that $(r+m+\beta\hat{p}_1)(\hat{S}_1 - \breve{S}_1) = -(\hat{y}_2 - d)$, we can write:

$$(r+m)(r+m+\beta\hat{p}_1)\hat{S}_2 = (r+m+\beta\hat{p}_1)(\hat{y}_2 - d) - \beta\hat{p}_1(\hat{y}_2 - d)$$

$$\Leftrightarrow (r+m+\beta\hat{p}_1)\hat{S}_2 = (\hat{y}_2 - d) \Leftrightarrow \hat{S}_2 = -(\hat{S}_1 - \breve{S}_1)$$

And finally:

$$\frac{(r+m+\lambda)(r+m+\beta p_2(\theta_2))}{r+m} S_2 = y_2 - d + \lambda \hat{S}_2 + \frac{\lambda}{r+m}\beta\hat{p}_1\breve{S}_1$$

$$\Leftrightarrow \frac{(r+m+\lambda)(r+m+\beta p_2(\theta_2))}{r+m} S_2 = y_2 - d + \lambda \frac{(\hat{y}_2 - d)}{(r+m+\beta\hat{p}_1)} + \frac{\lambda}{r+m}\beta\hat{p}_1 \frac{(y_1 - d)}{(r+m+\beta\hat{p}_1)}$$

The sector 2 job creation equation which determines θ_2^* is then given by:

$$V_2 = 0 \Leftrightarrow c = (1-\beta)q_2 S_2$$

Efficiency

Due to the high-skill sector segmentation, the social surplus per capita and per period has to be rewritten as:

$$\Sigma = \ell_1 y_1 + \hat{\ell}_1 y_1 + \ell_2 y_2 + \hat{\ell}_2 \hat{y}_2 + (u_1 + u_2)d - \theta_1 u_1 c - \hat{\theta}_1 \hat{\ell}_2 c - \theta_2 u_2 c - me$$

Replacing each labor market employment state by its equation gives:

$$\begin{aligned}\Sigma \; = \; & \frac{p_1\pi}{(m+p_1)} y_1 + \frac{\lambda p_2 \hat{p}_1 (1-\pi)}{(m+\lambda)(m+p_2)(m+\hat{p}_1)} y_1 + \frac{m p_2 (1-\pi)}{(m+\lambda)(m+p_2)} y_2 \\ & + \frac{m\lambda p_2 (1-\pi)}{(m+\lambda)(m+p_2)(m+\hat{p}_1)} \hat{y}_2 + \left(\frac{m\pi}{(m+p_1)} + \frac{m(1-\pi)}{m+p_2}\right) d \\ & - \theta_1 c \frac{m\pi}{(m+p_1)} - \hat{\theta}_1 c \frac{m\lambda p_2 (1-\pi)}{(m+\lambda)(m+p_2)(m+\hat{p}_1)} - \theta_2 c \frac{m(1-\pi)}{m+p_2} - me\end{aligned}$$

Graduates high-skill sector

One can show that the derivative of Σ with respect to θ_1 has the same sign as:

$$HS \equiv (1-\eta_1)p_1(y_1 - d) - \theta_1 c (m + \eta_1 p_1)$$

Knowing that the equilibrium equation for graduates' high-skill job creation is given by:

$$c = q_1(1-\beta)\frac{y_1 - d}{(r+m+\beta p_1)}$$

and under the Hosios condition ($\beta = \eta_1$), with r assumed to be equal to zero, we have:

$$HS = 0$$

Job creation is thus efficient in the graduates' high-skill sector ($\theta_1^* = \theta_1^S$).

Trainees high-skill sector

One can show that the derivative of Σ with respect to $\hat{\theta}_1$ has the same sign as:

$$\hat{HS} \equiv (1-\hat{\eta}_1)\hat{p}_1(y_1 - \hat{y}_2) - \hat{\theta}_1 c (m + \hat{\eta}_1 \hat{p}_1)$$

Knowing that the equilibrium equation for trainees' high-skill job creation is given by:

$$c = \hat{q}_1(1-\beta)\frac{y_1 - \hat{y}_2}{(r+m+\beta\hat{p}_1)}$$

And under the Hosios condition ($\beta = \hat{\eta}_1$), with r assumed to be equal to zero, we have:

$$\hat{HS} = 0$$

Job creation is also efficient in the trainees' high-skill sector ($\hat{\theta}_1^* = \hat{\theta}_1^S$).

<u>*Low-skill sector*</u>

One can show that the derivative of Σ with respect to θ_2 has the same sign as:

$$\begin{aligned} LS \;\equiv\; & (1-\eta_2)p_2\left[\frac{\lambda\hat{p}_1(y_1-\hat{y}_2)}{(m+\lambda)(m+\hat{p}_1)} + \frac{\lambda\hat{y}_2}{(m+\lambda)} + \frac{my_2}{(m+\lambda)} - d - \hat{\theta}_1 c\frac{m\lambda}{(m+\lambda)(m+\hat{p}_1)}\right] \\ & -c\theta_2(m+\eta_2 p_2) \end{aligned}$$

Knowing that sector 2 equilibrium equation is given by:

$$c = q_2(1-\beta)\frac{r+m}{(r+m+\lambda)(r+m+\beta p_2)}\left[y_2 - d + \lambda\frac{(\hat{y}_2-d)}{(r+m+\beta\hat{p}_1)} + \frac{\lambda}{r+m}\beta\hat{p}_1\frac{(y_1-d)}{(r+m+\beta\hat{p}_1)}\right]$$

And taking into account the equilibrium equation for trainees' high-skill job creation, LS satisfies:

$$\begin{aligned} LS \equiv (1-\eta_2)p_2 \quad & \left[\ \frac{\lambda\hat{p}_1(y_1-\hat{y}_2)}{(m+\lambda)(m+\hat{p}_1)} + \frac{\lambda\hat{y}_2}{(m+\lambda)} + \frac{my_2}{(m+\lambda)} - d \right. \\ & \left. -\hat{\theta}_1\hat{q}_1(1-\beta)\frac{y_1-\hat{y}_2}{(r+m+\beta\hat{p}_1)}\frac{m\lambda}{(m+\lambda)(m+\hat{p}_1)}\right] \\ -\frac{\theta_2 q_2(1-\beta)(m+\eta_2 p_2)(r+m)}{(r+m+\lambda)(r+m+\beta p_2)} \quad & \left[y_2 - d + \lambda\frac{(\hat{y}_2-d)}{(r+m+\beta\hat{p}_1)} + \frac{\lambda}{r+m}\beta\hat{p}_1\frac{(y_1-d)}{(r+m+\beta\hat{p}_1)}\right] \end{aligned}$$

We can simplify the terms in $(y_2 - d)$. After some manipulation, the expression can be rewritten as:

$$\begin{aligned} \Leftrightarrow\quad LS \equiv (1-\eta_2)p_2 & \left[\frac{\lambda\hat{p}_1(y_1-\hat{y}_2)}{(m+\lambda)(m+\hat{p}_1)} + \frac{\lambda(\hat{y}_2-d)}{(m+\lambda)} - \hat{\theta}_1\hat{q}_1(1-\beta)\frac{y_1-\hat{y}_2}{(r+m+\beta\hat{p}_1)}\frac{m\lambda}{(m+\lambda)(m+\hat{p}_1)}\right] \\ -\theta_2 q_2(1-\beta)\frac{r+m}{(r+m+\lambda)} & \left[\lambda\frac{(\hat{y}_2-d)}{(r+m+\beta\hat{p}_1)} + \frac{\lambda}{r+m}\beta\hat{p}_1\frac{(y_1-\hat{y}_2+\hat{y}_2-d)}{(r+m+\beta\hat{p}_1)}\right] \end{aligned}$$

One can show that the terms in $(y_1 - \hat{y}_2)$ and the terms in $(\hat{y}_2 - d)$ are both equal to zero.

Under the Hosios condition ($\beta = \eta_2$) and r assumed to be equal to zero, we therefore have:

$$LS = 0$$

Job creation is efficient in the low-skill sector ($\theta_2^* = \theta_2^S$).

We can also demonstrate, in the same way as in paragraph 4.3, that educational choices are efficient ($E = 0$) (for $r = 0$). The scenario described in this Appendix is thus such that the decentralized equilibrium is efficient (simultaneously $HS = 0$, $\hat{HS} = 0$, $LS = 0$, and $E = 0$ which means $\theta_1^* = \theta_1^S$, $\hat{\theta}_1^* = \hat{\theta}_1^S$, $e^* = e^S$, $\theta_2^* = \theta_2^S$) without involving any economic policy.

References

Acemoglu, D. Technical Change, Inequality, and The Labor Market. *Journal of Economic Literature* 2002; 40; 7-72.

Acemoglu, D. and J-S. Piscke. Beyond Becker: Training in Imperfect Labor Markets. *Economic Journal* 1999; 109; 112-142.

Albrecht, J.W., P.A. Gautier and S.B. Vroman. Equilibrium Directed Search with Multiple Applications. *Review of Economic Studies* 2006; 73(4); 869-891.

Alonzo, P. and O. Chardon, "Quelle carrière professionnelle pour les salariés non qualifiés?" *Données Sociales*, la société française, Insee Références (2006), 265-272.

Andersson, F., H.J. Holzer and J. Lane. Worker Advancement in the Low-Wage Labor Market: The Importance of Good Jobs. *Brookings Institution Policy Brief* 2003, Center on Urban and Metropolitan Affairs. Washington, D.C.: Brookings Institution, 2003.

Andersson, F., H.J. Holzer and J. Lane. *Moving Up or Moving On: Who Advances in the Low-Wage Labor Market?* New York: Russell Sage Foundation; 2005.

Arrow, K.J. The Economic Implications of Learning by Doing. *Review of Economic Studies* 1962; 29(3); 155-173.

Becker, G.S. *Human Capital*. Chicago: The University of Chicago Press; 1964.

Cahuc, P., F. Postel-Vinay and J-M. Robin. Wage Bargaining with On-the-job Search: Theory and Evidence. *Econometrica* 2006; 74(2); 323-364.

Connolly, H. and P. Gottschalk. Stepping stone jobs: Theory and evidence. Boston MA, Economics Department, Boston College, Working Paper 427; 2001.

Charlot, O. and B. Decreuse. Self-selection in education with matching frictions. *Labour Economics* 2005; 12; 251-267.

Charlot, O. and B. Decreuse. Over-education for the rich, under-education for the poor: a search-theoretic microfoundation. Paper no. 3624, MPRA; 2007.

Gautier, P.A. and R.P. Wolthoff. Simultaneous Search with Heterogeneous Firms and Ex Post Competition. *Labour Economics* 2009; 16(3); 311-319.

Gavrel, F., I. Lebon and T. Rebiere. Career paths, unemployment, and the efficiency of the labor market: Should youth employment be subsidized? *Journal of Public Economic Theory* 2010; 12; 533-560.

Gottschalk, P. Wage Mobility within and between Jobs. LoWER Working Papers wp1, AIAS, Amsterdam Institute for Advanced Labour Studies; 2001.

Hagedorn, M., Manovskii I. and S. Stetsenko. Taxation and Unemployment in Models with Heterogeneous Workers. Working paper; 2010.

Hall and Milgrom, 2008. Hall, R. and P. Milgrom. The Limited Influence of Unemployment on the Wage Bargain. *American Economic Review* 2008; 98(4); 1653-1674.

Holzer, H.J. Encouraging Job Advancement Among Low-Wage Workers: A New Approach. *Brookings Institution Policy Brief: Welfare Reform and Beyond* 30, Washington, DC: Brookings Institution; 2004.

Hosios, A. On the Efficiency of Matching and Related Models of Search and Unemployment. *Review of Economic Studies* 1990; 57(2); 279-298.

Knabe, A. and A. Plum. "Low-Wage Jobs - Springboard to High-Paid Ones? *LABOUR: Review of Labour Economics and Industrial Relations* 2013; 27(3); 310-330.

Ljungqvist, L. and T.J. Sargent. The European Unemployment Dilemma. *Journal of Political Economy* 1998; 106(3); 514-550.

Machin, S. Changes in the relative demand for skills in the UK labor market". In *Acquiring Skills: Market Failures, Their Symptoms and Policy Responses*, ed. Alison L. Booth and Dennis J. Snower. Cambridge: Cambridge University Press; 1996.

Mincer, J. Investment in U.S. Education and Training. Working Paper no.4844, NBER; 1994.

Mincer, J. Technology and the Labor Market. *Review of Economics of the Household* 2003; 1(4); 249-272.

Moen, E.R. Education, Ranking, and Competition for Jobs. *Journal of Labor Economics* 1999; 17(4); 694-723.

Moen, E.R. and A. Rosén. Does Poaching Distort Training? *Review of Economic Studies* 2004; 71(4); 1143-1162.

Monso, O. Changer de groupe social en cours de carrière: davantage de mobilité depuis les années quatre-vingt. *INSEE Premiere*, no. 1112, division Emploi, Insee; 2006.

Mortensen, D. *Wage Dispersion: Why are Similar Workers Paid Differently?* Cambridge, MA: MIT Press; 2003.

Moscarini, G. and F. Vella. Occupational Mobility and the Business Cycle. Working Paper no. 13819, NBER; 2008.

Pigou, A.C. *Wealth and welfare*. London: Macmillan; 1912.

Pissarides, C.A. *Equilibrium Unemployment Theory*. 2nd edition. Cambridge, MA: MIT Press; 2000.

Sattinger, M. Overlapping Labour Markets. *Labour Economics* 2006; 13(2); 237-257.

Shimer, R. On-the-job Search and Strategic Bargaining. *European Economic Review* 2006; 50; 811-830.

Sicherman, N. and O. Galor. A Theory of Career Mobility. *Journal of Political Economy* 1990; 98(1); 169-192.