Over-education for the rich, under-education for the poor: a search-theoretic microfoundation*

Olivier CHARLOT
LEAD, Université Antilles-Guyane and CIRPEE†

Bruno DECREUSE
GREQAM, IDEP, and Université de la Méditerranée‡

May 2010

Abstract: This paper studies the efficiency of educational choices in a two sector/two schooling level matching model of the labor market where a continuum of heterogeneous workers allocates itself between sectors depending on their decision to invest in education. Individuals differ in working ability and schooling cost, the search market is segmented by education, and there is free entry of new firms in each sector. Self-selection in education causes composition effects in the distribution of skills across sectors. This in turn modifies the intensity of job creation, implying the private and social returns to schooling always differ. Provided that ability and schooling cost are not too positively correlated, agents with large schooling costs – the ‘poor’ – underinvest in education, while there is overinvestment among the low schooling cost individuals – the ‘rich’. We also show that education should be more taxed than subsidized when the Hosios condition holds.

Keywords: Ability; Schooling cost; Heterogeneity; Matching frictions; Efficiency

J.E.L. classification: I20, J21, J60

---

*We thank Antoine d’Autume, Pierre Cahuc, Etienne Lehmann, Bertrand Wigniolle, participants in various conferences and seminars, and two referees and an editor of this review for their comments and suggestions. The usual disclaimer applies.
†LEAD, Université Antilles-Guyane, and CIRPEE. UFR SJE, Campus de Fouillole, BP 270, 97157 Pointe-à-Pitre Cedex, Guadeloupe. E-mail: ocharlot@univ-ag.fr
‡GREQAM – 2, rue de la charité, 13236 Marseille cedex 02, France. E-mail: decreuse@univmed.fr
1. Introduction

Self-selection in education gives rise to composition effects in the distribution of skills across education groups. The characteristics of the agents who educate generally differ from the characteristics of those who do not. In a competitive labor market, such composition effects shape within and between-group wage dispersion. However, they do not alter the efficiency of educational choices. This property does not hold in a frictional labor market, where composition effects may distort the incentives to schooling. In a model where agents only differ with respect to labor market ability, Charlot and Decreuse (2005) show that self-selection in education is inefficient, and too many workers are willing to acquire education. The purpose of the present paper is to reconsider their result in the realistic case where agents not only differ in labor market ability, but also in schooling cost. Our main conclusion is that provided ability and schooling cost are not too positively correlated, agents with large schooling costs – the ‘poor’ – underinvest in education, while there is overinvestment among the low schooling cost individuals – the ‘rich’. We also study the combination of labor market and education policies that decentralizes the efficient allocation. There, we suggest that education should be more taxed than subsidized.

Our model is based on three main features. First, there is worker-job heterogeneity. There are two schooling levels, educated/uneducated, and two production sectors, one for the high-skill (complex) jobs, and the other for the low-skill (simple) jobs. There are two sources of heterogeneity among workers: ability and schooling costs. This assumption is the major departure from our previous contribution (Charlot and Decreuse, 2005). We shall refer to agents endowed with low schooling costs as the ‘rich’, and to those with high schooling costs as the ‘poor’.\footnote{This is of course a short-cut as we consider a non-degenerate joint distribution of characteristics.} Second, there are matching frictions and the search market is segmented by education. The number of contacts between employers and job-seekers is driven by a constant returns to scale matching technology. The search market is fully segmented by education/technology, but not by ability: only the educated can do the high-skill jobs, and firms target their high-skill positions towards the educated. Workers target their applications towards the jobs they want to perform, but search is random within each sector. Wages are determined by Nash bargaining, and sectoral job creation is endogenized by means of a free-entry condition. Free entry implies sectoral job creation depends on the composition of education groups. Third, there is self-selection in education. Agents face a binary schooling choice, i.e. being educated or uneducated, and select themselves on the basis of their private costs and returns to schooling. Only those whose ability is sufficiently high and/or whose schooling cost is sufficiently low choose to invest in education. At the individual level, self-selection depends on market factors, such
as productivity differences as in Roy’s (1951) model of job assignment, and unemployment rate differences across sectors, which is a straightforward consequence of the incorporation of search frictions in a Roy-type economy.

In a frictionless (Walrasian) environment, a worker’s return to education depends on her own innate characteristics and on sectoral productivity differences, but it is not altered by the composition of education groups. Educational choices are therefore efficient. In a frictional economy, the composition of education groups alters wage and employment opportunities, and, in turn, self-selection is driven by sectoral wage and employment differences. As a result, educational choices are generally inefficient.

The reason for inefficiency is as follows. Wage and employment opportunities depend on the intensity of job creation. Rent-sharing implies that the profitability of a filled job rises with worker’s ability. Under free entry, job creation in each sector depends on the average labor market ability across the pool of job-seekers.\(^2\) Anyone whose ability is above (below) the sector-specific mean ability generates a positive (negative) sector-specific externality by boosting (reducing) job creation, and therefore driving wage and employment opportunities upwards (downwards). Hence, high-ability individuals underestimate the return to schooling, while the less able over-estimate it. Provided that schooling cost and ability are not too positively correlated, agents with large schooling costs – the ‘poor’ – underinvest in education, while there is overinvestment among the low schooling cost individuals – the ‘rich’. Hence, search frictions provide a rationale to the following claim: there are too many rich investing in education, which depreciates the return to schooling, and crowds out poorer and abler individuals from schooling.

However, due to congestion externalities, over-education may prevail at all schooling costs. Assume individuals only differ in their schooling costs. In this case, the private return to participation in a given sector is lower than the social return unless the so-called Hosios condition holds – workers’ bargaining power must be equal to the elasticity of the matching function with respect to the pool of unemployed. If the difference between private and social return to participation is much smaller in the high-skill sector than in the low-skill one, then the private return to schooling exceeds the social return.

Finally, we turn to education policy. We assume the planner only observes individual schooling cost. Efficiency requires that some of the high ability individuals with high schooling cost must be attracted to education, while some of the low ability individuals with low schooling costs must be deterred from education. This can be done with two instruments: a subsidy increasing in schooling costs, combined with a lump-sum tuition

\(^2\) The empirical evidence shows that wages are positively correlated with profitability. For instance, using a matched panel to control for worker and firm heterogeneity, Abowd, Kramarz and Margolis (1999) demonstrate that high-skill workers are paid more and that profitability is higher for firms with more skilled workers.
cost. The optimal policy is therefore redistributive, as the ‘poor’ are more subsidized than they are taxed while the ‘rich’ are more taxed than subsidized. We show that the total amount of taxes is larger than the total amount of subsidies. Education, therefore, should be more taxed than subsidized. Indeed, the social return to schooling is strictly increasing and strictly concave in ability. The negative externality imposed by poorly talented individuals tends to be bigger than the positive externality induced by high ability persons. This result completes our earlier contribution with a single schooling cost where we show that education should always be taxed.

The search literature already provides a number of reasons why there may be inefficient educational investment. Acemoglu (1996) shows that rent sharing translates into a hold-up problem for the workers. In Snower (1995) and Saint-Paul (1996), there may be some excess supply of skills as a rise in the number of educated alters firms’ incentives to allocate vacancies between sectors. In Moen (1999), investing in education improves one’s ranking in the job queue, but at the expense of the others. In the multi-dimensional skill model of Charlot, Decreuse and Granier (2005), workers invest in education to improve their chance of being employed, but also to raise their outside opportunities during the wage bargain. This latter return to education is not matched by a social gain and workers over-invest in education.

The externality we put forward in the current paper is based on two-sided heterogeneity. In models where there is a unique search market for all skill levels, Laing, Palivos and Wang (1995), and Burdett and Smith (2002) demonstrate that there are social increasing returns to schooling. Acemoglu (1999) and Albrecht and Vroman (2002) consider a model with two skill levels and two job types, where exogenous changes in the distribution of skills alter skilled/unskilled workers’ wage and employment prospects. This occurs because a change in the proportion of skilled workers modifies firms’ incentives to advertise high rather than low skill positions. This externality does not arise in our paper as the search place is segmented by education.

Our paper highlights a case of ‘market adverse selection’. The return to participation in each sector depends on the intensity of job creation, which, under free entry, depends on the mean ability of job seekers in that sector. As the return to participation in each sector depends on sector-specific mean ability, all the agents tend to be attracted by the sector where the mean ability is the highest and where wage and employment opportunities are the best. At the market level, the ablest workers create wage and employment externalities for their less talented co-workers. This situation occurs because search is random within

---

3In signalling models (see Arrow, 1973, and Spence, 1973), high ability workers also provide a wage externality to low ability workers, but at the firm level. In our paper, firms perfectly observe individual characteristics at the recruitment stage, and could reject an application to wait for a better match. In this perspective, the inefficiency of the decentralized allocation is due to employers’ failure at the sector
sectors: the market is only segmented by education levels and not by education and ability. To appreciate the restrictiveness of such assumptions, it is worth investigating two polar cases.

On the one hand, the search market may be totally unsegmented. An externality similar to that highlighted by Burdett and Smith (2002) would then be at work. Increasing the share of educated individuals would raise the probability of meeting such workers. High-skill vacancies would then become more profitable. This would boost job creation. Such an externality would lead to under-education. In our paper, the search market is segmented by job type. The above mentioned externality cannot arise, because each worker who gets education also gets the right to participate in the market for high-skill jobs.

On the other hand, search could be directed and the search market may be fully segmented, that is both by schooling level and by ability. The composition of education groups does not alter the schooling return in such a case, as individuals are homogenous within each cell of the search market. Of course, this is no longer the case in our contribution, because low ability workers benefit from the presence of abler individuals in their market segment. Low ability workers have incentives to participate in the market segments where wage and employment opportunities are expected to be better.\(^4\)

Finally, non-frictional sorting models of education (such as those surveyed in Fernandez, 2001) also provide a number of examples where the composition of education groups matters for efficiency. At the family level, there can be marital sorting influencing the way in which human capital and other abilities are intergenerationally transmitted. In the schooling system, there can be some peer effects involved. At the production level, there are some complementarities between workers of different skill levels which may not be internalized within firms. In the city, the composition of the neighborhood induces some important externalities through taxation and the provision of schooling. Closest to us, Fershtman, Murphy and Weiss (1996) build a model in which agents also differ in ability and schooling cost (“nonwage income” in their terminology). Beyond the access to high-skill/best-paid positions, education also provides the workers with social status. In their view, the demand for social status increases with the average ability of the high-skill level to coordinate on the rejection of low ability workers’ applications. Of course, this depends on the organization of the search market as we argue below.

\(^4\)This is also due to the way wages are set in our paper, i.e. to rent-sharing. In the two ability level matching model of Inderst (2005), firms set wages and make them contingent on workers’ ability (see also Shi, 2001). The search market partitions into two market segments. Our paper adopts the view that such contracts are not enforceable, either because ability is not observable (or at least not verifiable) by a court, or because workers have some market power. In the former case, Inderst shows that the search market is no longer segmented (see also the model of Lang and Manove, 2003, with continuous types).
workers. This entails a composition externality, so that the “wrong” individuals may be led to acquire schooling. In our model, the composition externality vanishes when market frictions disappear. Search frictions provide a different yet realistic channel through which composition effects can be held responsible for the crowding-out from schooling of the poor by the rich.

The paper is organized as follows: section 2 presents the basic setup. Section 3 studies the properties of the Walrasian environment, whereas section 4 considers a frictional labor market. Section 5 deals with the design of the efficient policy. Section 6 concludes.

2. The model

2.1. Environment

We focus on the steady state of a continuous time matching model with two sectors, heterogeneous workers and endogenous education choices. Time is continuous. At each instant, \( \delta > 0 \) agents are born. This ensures new cohorts enter the population and make their education decision at each instant. Agents are risk neutral, discount time at rate \( \rho \), and have a constant risk of dying (retirement) \( \delta > 0 \). Hence, the effective rate of discount is \( r \equiv \rho + \delta \), and the total population is normalized to 1.

**Heterogeneity and education.** There are two schooling levels: educated/uneducated. Agents are heterogeneous with respect to two innate characteristics: schooling cost \( c \) and labor market ability \( a \) (ability, for short). Those who pay the schooling cost get education; those who do not remain uneducated. Heterogenous schooling costs may reflect the existence of credit constraints, or different aptitudes to learn at school. Though we are aware these two interpretations may actually lead to drastically different policy recommendations, our results are derived under the assumption private and social schooling costs coincide. The externalities we point out are therefore independent of the standard market failure due to credit market imperfections.\(^5\) In the same way, heterogeneous abilities may be attributed to different innate productive capacities, as well as to different social backgrounds. Schooling cost and ability are distributed across individuals within cohorts according to the stationary joint distribution of cdf \( \Phi \) and pdf \( \phi \).

**Assumption 1** The function \( \phi : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+ \) is such that (i) \( \phi \) is positive and continuous, (ii) \( \mathbb{E}(a) = \int_c \int_a \phi(a, c) \, adac < +\infty \), (iii) \( \mathbb{E}(a \mid a \geq bc + d) > \mathbb{E}(a \mid a \geq d) \) for all \( b > 0, \, d \geq 0 \).

\(^5\)In section 5 where we deal with policy issues, it is explicitly assumed that there are no credit constraints.
Part (i) rules out the possibility of a mass point in the distribution of ability and schooling cost. Unlike our former model with a single schooling cost, there is a non-degenerate distribution of schooling cost at each ability level. Part (ii) states that the supply of productive skills is finite. Part (iii) requires that education and schooling cost are not too positively correlated. Consider individuals whose ability $a$ is larger than $d$. Among, those agents, select those whose ability is larger than (a proportion of) the schooling cost. Part (iii) requires that those agents have a higher mean ability than those whose ability is lower than the schooling cost. This assumption seems reasonable. However, it does not always hold. For instance, if agents with a high ability are also endowed with a very high schooling cost, selecting those who have a higher ability than schooling cost means selecting a group of low ability persons. A weak version of requirement (iii) is needed to prove the existence of a decentralized allocation with positive education. We also need it to characterize the centralized allocation.

*Production technologies.* There are two sectors producing a single final good: a high-skill, high-productivity sector where complex tasks are performed and a low-skill, low-productivity sector where simple tasks are performed. Sectors are indexed by $i = h, l$, where $h$ refers to the high-skill sector, and $l$ to the low-skill sector. Only the educated workers can do the high-skill jobs. In each sector, there is an endogenous mass of firms all endowed with a single job, which can be either vacant, or filled. As in the standard model (see Pissarides, 2000), holding a vacancy involves a flow cost $\gamma > 0$. The cost is the same in each sector. Our results would still hold provided that the search cost in the high-skill sector is not too high compared to the cost in the low-skill sector. The key property is that the sector-specific cost must not be proportional to the sector-specific mean ability.

Output depends on firms’ and workers’ characteristics. Denoting by $y_{ia}$ the output of a type $a$ worker in a sector $i$ job, we have

$$y_{ia} = A_ia$$

where $A_i$ is a sector-specific productivity parameter, $A_h > A_l > 0$. Complex jobs are more productive than simple jobs. This implies education has a productive role: it allows the workers to perform on more complex and more productive technologies. There are no peer effects at the production stage: the ability of other workers does not affect one’s productivity.

*Matching sector.* The search market is segmented by sector. Each individual is allowed

\(^6\)We follow the tradition initiated by early matching models where $\gamma$ is viewed as an advertising and recruitment cost. One may reasonably consider that such a cost is larger in the high-skill sector than in the low-skill one. Alternatively, firms may purchase a capital unit at the time of job creation and recoup it gradually over time as in Albrecht and Vroman (2002). The price of the capital unit should not differ from a sector to another.
to participate in either one of the two sectors, but not to both. In each sector, unemployed and vacancies are brought together by pair according to a constant returns to scale (CRS) matching technology. The flow of (sectoral) matches $M_i$ between $u_i$ unemployed and $v_i$ vacancies is given by:

$$M_i \equiv m_0 M(u_i, v_i)$$

with $m_0 > 0$ a scale parameter sizing the magnitude of search frictions. Matching is random in each sector: matches are equiprobably distributed between unemployed – irrespective of ability – as well as between vacancies. Due to CRS, the flow probability of a worker to match a vacancy $\mu_i$ and the flow probability of a vacancy to match a worker $\eta_i$ only depend on the ratio of vacancies to job-seekers, the market tightness $\theta_i \equiv v_i/u_i$. Hence, $\mu_i \equiv m_0 M_i/v_i = \mu(\theta_i)$, and $\eta_i \equiv m_0 M_i/u_i = \mu(\theta_i)/\theta_i$.

**Assumption 2** The function $\mu$ is strictly increasing, strictly concave, satisfies the boundary conditions $\mu(0) = 0$ and $\lim_{\theta \to \infty} \mu(\theta) = \infty$, and the Inada conditions $\lim_{\theta \to 0^+} \mu'(\theta) = \infty$, and $\lim_{\theta \to 0^-} \mu'(\theta) = 0$.

These standard properties imply the function $\eta$ is strictly decreasing from infinity to 0.

**Flows and stocks.** Let $\Omega_{it}$, $i = h, l, e$ denote the sub-set of $\mathbb{R}_+ \times \mathbb{R}_+$ that induces, respectively, the choice of schooling (and, therefore, of participating to the high-skill sector), participating to the low-skill sector, and non-participating at time $t$. Since schooling is costly, non-participating workers do not get education and the intersection of the various subsets taken by pairs is empty, that is $\Omega_{it} \cap \Omega_{jt} = \emptyset$ for $i \neq j$. In addition, $\cup_i \Omega_{it} = \Omega$.

The size of the labor force $n_i$ on sector $i = h, l$ evolves according to

$$dn_i/dt = \delta \int \int_{\Omega_{it}} \phi(a, c) dcd\alpha - \delta n_i$$

The only source of separation is the workers’ death (retirement). The rate of unemployment $u_i$ and the mass-numbers of employees $l_i$ in sector $i$ obey the following law of motion:

$$\frac{du_i}{dt} = \frac{\delta \int \Omega_{it} \phi(a, c) dcd\alpha}{n_i} - (\mu_i + \delta) u_i$$

$$l_i = n_i (1 - u_i)$$

Entrants make up the inflow into unemployment in each sector, while the outflow is equal to the sum of hires and the number of deaths. In steady-state, $\Omega_{it} = \Omega_i$ for all $t$ and
\[ \frac{dn_i}{dt} = du_i/dt = 0. \] It follows that

\[ n_i = \int \int \int_{\Omega_i} \phi(a, c) \, dcda \quad (2.6) \]
\[ u_i = \frac{\delta}{\mu_i + \delta} \quad (2.7) \]
\[ l_i = \frac{\mu_i}{\mu_i + \delta} n_i \quad (2.8) \]

Aggregate output in sector \( i \) is \( Y_i = A_i l_i \bar{a}_i \), where \( \bar{a}_i \equiv \mathbb{E}[a \mid (a, c) \in \Omega_i] \) is the average ability among workers in sector \( i \). Finally, the aggregate schooling cost is \( C \equiv \delta \int \int_{\Omega_i} \phi(a, c) \, dcda \).

### 2.2. Decentralized economy

**Agents’ gains.** In the decentralized economy, the allocation of workers between sectors is driven by self-selection in education, while sectoral job creation depends on firms’ incentives to enter the search market. The size of rents accruing to each party when a match occurs is determined by ex-post Nash bargaining.

Let \( U_{ia} \) and \( W_{ia} \) be the respective values of being unemployed and employed in sector \( i \) for a type \( a \) worker. Let also \( V_i \) and \( J_{ia} \) denote the values of a sector \( i \)-vacant job and of a sector \( i \)-filled job with a type \( a \) worker.

Unemployed receive no income. Denoting the wage rate by \( w_{ia} \), we have:

\[ rU_{ia} = \mu_i [W_{ia} - U_{ia}] \quad (2.9) \]
\[ rW_{ia} = w_{ia} \quad (2.10) \]
\[ rJ_{ia} = y_{ia} - w_{ia} + \delta V_i \quad (2.11) \]
\[ \rho V_i = -\gamma + \eta_i [\mathbb{E} (J_{ia} | (a, c) \in \Omega_i) - V_i] \quad (2.12) \]

As it is standard, each value can be considered an asset value. The return to the asset is equal to the flow benefit (or loss), plus the expected gain (loss) resulting from a potential change of state. Random matching in each sector implies the type \( a \) of the incoming worker is a priori unknown to the employer. This justifies the conditional mean operator \( \mathbb{E} \) in equation (2.12).

Wages are determined by Nash bargaining:

\[ \beta [J_{ia} - V_i] = (1 - \beta) [W_{ia} - U_{ia}] \quad (2.13) \]

where \( \beta \in (0, 1) \) is workers’ exogenous bargaining power.\(^7\)

\(^7\)When \( \beta = 0 \), workers do not get any return from their education and none would educate. When \( \beta = 1 \), job creation is nil and the return to education would also be zero. We rule out those cases here.
Schooling, non-participation and tightness. We present how (i) non-participation, (ii) self-selection in education and (iii) sector-specific market tightness are determined.

(i) Non-participants get a negative return from their participation to each sector. Formally, an agent endowed with ability \( a \) and schooling cost \( c \) does not participate in the labor market if and only if

\[
\max \langle U_{la}, U_{ha} - c \rangle < 0
\]  

(2.14)

(ii) Each agent compares her utility if she decided to get education (net of the schooling cost) to that she would get if she decided to remain uneducated. Therefore, an agent endowed with ability \( a \) and schooling cost \( c \) becomes educated if and only if

\[
U_{ha} - c \geq \max \langle U_{la}, 0 \rangle
\]  

(2.15)

(iii) Firms enter the search market until the exhaustion of all rents. Free entry implies

\[
V_{v}^{u} \leq 0
\]  

(2.16)

In the rest of the paper, we focus on the case where \( \rho \) tends to 0, a usual assumption in the search literature, which simplifies the welfare analysis.

2.3. Centralized economy

The social planner chooses the number of vacancies in each sector as well as the assignment of workers to sectors – the allocation of talents. The planner takes the joint distribution of schooling cost and ability as given\(^8\) and maximizes the discounted path of aggregate consumption. When the rate of time preference \( \rho \) tends to 0, we have:

\[
S = \max_{\Omega, \theta_i} \left\langle \sum_{i} (Y_i - \gamma v_i) - C \right\rangle
\]  

(2.17)

This is the sum of sector-specific output net of search costs, minus aggregate schooling costs. Using the definitions above, and noting that \( v_i = \theta_i u_i \), we get

\[
S = \max_{\Omega, \theta_i} \left\langle \sum_{i} \int_{\Omega_i} \int_{\Omega_h} \phi (a, c) (1 - u_i) \left[ A_i a - \frac{\gamma}{\eta_i} \right] d\sigma d\eta - \delta \int_{\Omega_i} \int_{\Omega_h} \phi (a, c) d\sigma d\eta \right\rangle
\]  

(2.18)

subject to \( u_i = \delta / (\delta + \mu (\theta_i)) \), \( \Omega_i \subset \mathbb{R}_+ \times \mathbb{R}_+ \), \( \Omega_i \cap \Omega_j = \emptyset \) for all \( i \neq j \), and \( \cup_i \Omega_i = \mathbb{R}_+ \times \mathbb{R}_+ \).

As a first step, we turn to the frictionless environment.

---

\(^8\)The planner could actually alter the pecuniary elements included in schooling costs. For instance, if heterogenous schooling costs resulted from credit constraints, the planner could reduce aggregate schooling costs simply by reallocating financial resources among agents.
3. Walrasian allocation

In this section, we analyze the Walrasian environment. A Walrasian environment is an environment where frictions are negligible. The scale parameter of the matching technology $m_0$ tends to infinity. We show the assignment of workers to jobs through self-selection in education is socially efficient.

3.1. Decentralized economy

When $m_0$ tends to infinity, the value of search tends to the value of employment. From the Nash bargaining equation (2.13), the value of a filled job for the firm tends to 0. Consequently, the wage is equal to the output flow:

$$w_{ia} = y_{ia}$$

(3.1)

As $y_{ia} \geq 0$, all agents participate in the labor market and $\Omega^*_e = \emptyset$. Self-selection in education is given by (2.15), which can be written as

$$(a, c) \in \Omega^*_e \Leftrightarrow A_ha/\delta \geq A_la/\delta + c$$

(3.2)

From this equation, we can draw an indifference locus in the $(a, c)$ plane, that is the set of characteristics of agents indifferent between acquiring education or remaining uneducated:

$$c = R^*a$$

(3.3)

where $R^* \equiv (A_h - A_l)/\delta$ is the return to schooling, that is the differential return to ability in the high and low productivity sectors. It is increasing in the productivity differential $A_h - A_l$, and it is independent of the workers’ ability.

Figure 3.1 depicts the assignment of agents to sectors. Individuals whose characteristics are below the indifference locus get education. Conversely, those whose characteristics are above the locus remain uneducated.

This self-selection mechanism refers to Roy (1951), where a continuum of (heterogeneous) workers select themselves between a discrete number of occupations. Each individual is characterized by a bundle of sector-specific skills, and chooses which sector to enter on the basis of comparative advantage. In our model, there is a unique labor market ability common to all sectors. This implies the returns to ability in the low- and high-skill sectors are perfectly correlated. However, the cost of entry to the high-skill sector differs across individuals. If one considers utilities obtained in each case instead of abilities, that is $U_h = A_ha/\delta - c$ and $U_l = A_la/\delta$, then Roy’s setup and ours are equivalent in terms of the payoffs the agents can expect.

Moscarini (2001) builds on and extends Roy’s setup to an environment with frictions, but does not consider the normative implications of his model. The decentralized economy of our model and Moscarini’s converge towards (formally) similar walrasian allocations as search frictions die down.
Figure 3.1: The allocation of talents in the Walrasian economy. Agents whose characteristics are below the line get educated, while the others remain uneducated.

One may note that part (iii) of Assumption 1 implies that \( \mathbb{E}(a \mid c < R^* a) > \mathbb{E}(a) \). The mean ability among educated individuals is larger than the unconditional mean ability. This standard property has no implications for the returns to schooling in the Walrasian environment. As we shall see, the results are deeply different in the frictional environment.

### 3.2. Centralized economy

When \( m_0 \) tends to infinity, the social criterion is the same as in (2.18), with \( u_i = 0 \) and \( \eta(\theta_i) \) tends to infinity. In the absence of externality, the fact that \( y_{ia} \geq 0 \) implies (like the decentralized economy) \( \Omega^*_{e} = \emptyset \). A type \( a \) worker gets education if and only if her net contribution to the social criterion is non-negative, that is

\[
(a, c) \in \Omega^*_h \Leftrightarrow A_h a \geq A_l a + \delta c
\]  

It follows that the social return to schooling \( R^* = R^* \) and \( \Omega^*_{h} = \Omega^*_h \). Consequently, self-selection is socially efficient.

The introduction of matching frictions dramatically alters this result, as we demonstrate below.
4. Frictional environment

In this section, we show search frictions create two important externalities implying the private return differs from the social return to schooling. On the one hand, self-selection in education leads to a composition externality that relates sector-specific wages and employment opportunities to sector-specific mean ability through job creation in each sector. Due to this externality, poorly able but rich individuals crowd out abler but poorer individuals from schooling. On the other hand, ex-post Nash bargaining generally fails to internalize standard congestion externalities. It follows the returns to each type of occupation are below the social returns. As a consequence, there may be over-education at all schooling costs.

4.1. Decentralized economy

A worker endowed with ability $a$ working in sector $i$ is paid the wage $w_{ia}$ such that

$$w_{ia} = \beta \frac{\delta + \mu_i}{\delta + \beta \mu_i} y_{ia}$$  \hspace{1cm} (4.1)

The worker gets a share $\beta (\delta + \mu_i) / (\delta + \beta \mu_i)$ of the output flow $y_{ia}$. This share increases with tightness and bargaining power and decreases with $\delta$ (see Pissarides, 2000). The wage also increases with the output flow $y_{ia}$, which raises match surplus. The value of search is

$$U_{ia} = \frac{\beta \mu_i}{\delta + \beta \mu_i} y_{ia}$$  \hspace{1cm} (4.2)

Like in the Walrasian case, $y_{ia} \geq 0$ implies $U_{ia} \geq 0$, and therefore all agents are willing to participate in the labor market in the decentralized outcome. Hence $\Omega_e = \emptyset$. According to the self-selection rule (2.15), self-selection in education is now given by

$$(a, c) \in \Omega_h \iff \frac{\beta \mu_h}{\delta + \beta \mu_h} A_h a \geq \frac{\beta \mu_l}{\delta + \beta \mu_l} A_l a + c$$  \hspace{1cm} (4.3)

The main difference with the Walrasian case is that the discounted value of output in each sector is weighted by a term lower than one reflecting employment opportunities and bargaining power in each sector. This term tends to one as $\mu_i$ tends to infinity.

As in the Walrasian case, we can define the set of individuals indifferent between being educated or remaining uneducated:

$$c = \left[ \frac{\beta \mu_h}{\delta + \beta \mu_h} A_h - \frac{\beta \mu_l}{\delta + \beta \mu_l} A_l \right] a$$  \hspace{1cm} (4.4)

Like the Walrasian case, equation (4.4) defines a linear relationship between schooling cost and ability. The slope of the curve is given by the return to schooling $R^*$, which is formally defined below.
**Definition 1 A characterization of the decentralized allocation**

A decentralized allocation is a vector \((\theta^*_i, \theta^*_h, \alpha^*_i, \alpha^*_h, R^*)\) satisfying the following conditions

(i) free entry
\[
\frac{\gamma}{\eta(\theta^*_i)} = \frac{(1 - \beta) A \alpha^*_i}{\delta + \beta \mu (\theta^*_i)}, \quad i = h, l
\]  
(4.5)

(ii) self-selection
\[
\alpha^*_i = \mathbb{E}[a | c > R^*a] \quad \text{and} \quad \alpha^*_h = \mathbb{E}[a | c \leq R^*a]
\]  
(4.6)

(iii) return to schooling
\[
R^* = \frac{\beta \mu (\theta^*_h) A_h}{\delta + \beta \mu (\theta^*_h)} - \frac{\beta \mu (\theta^*_l) A_l}{\delta + \beta \mu (\theta^*_l)}
\]  
(4.7)

Individuals self-select in education on the basis of the return to schooling \(R^*\). Equation (4.7) shows that it is increasing in high-skill sector market tightness and decreasing in low-skill sector market tightness. Equation (4.6) indicates that, like the Walrasian case, self-selection alters the mean ability in each sector. However, and unlike the Walrasian case, changes in mean ability affect the return to schooling through job creation decisions.

In equation (4.5), the mean search cost \(\gamma / \eta(\theta_i)\) incurred by firms in each sector is equal to the expected value of a filled job. Rent-sharing and free entry imply that sector-specific market tightness increases with the mean ability in each sector.

An equilibrium is determined by means of a fixed-point argument. According to (4.5), tightness is an increasing function of sector-specific ability, i.e., \(\theta_i \equiv \Theta_i(\alpha_i)\). In turn, the return to schooling \(R\) is an increasing function of \(\theta_h\), and a decreasing function of \(\theta_l\). These properties imply \(R \equiv \mathcal{R}(\Theta_h(\alpha_h), \Theta_l(\alpha_l))\). Solving reduces to determine the fixed-point of

\[
\alpha^*_i = \mathbb{E}[a | c > \mathcal{R}(\Theta_h(\alpha^*_h), \Theta_l(\alpha^*_l)) a]
\]  
(4.8)

\[
\alpha^*_h = \mathbb{E}[a | c \leq \mathcal{R}(\Theta_h(\alpha^*_h), \Theta_l(\alpha^*_l)) a]
\]  
(4.9)

**Proposition 1 Existence of an interior equilibrium**

There exists an interior equilibrium with \(\alpha^*_l < \mathbb{E}(a) < \alpha^*_h\)

A few comments are in order. First, the existence of an interior equilibrium relates to our assumptions on the joint distribution of ability and schooling cost. Due to the effect of self-selection on the returns to schooling, it is not obvious that some agents choose to become educated in equilibrium. For instance, suppose that high ability, low schooling costs agents decide to become educated. Firms create many high-skill jobs. In turn,
this attracts additional agents in education, including some low ability workers. If those workers are sufficiently numerous, the mean ability falls below the unconditional mean ability, firms create fewer jobs, and search prospects may become lower in the high-skill sector than in the low-skill sector. Assumption 1 avoids this non-standard situation and allows us to apply the fixed-point theorem. If high ability workers decide to become educated, this must be because the equilibrium return to schooling is positive. In turn, when the equilibrium return to schooling is positive, the mean ability among educated individuals is larger than the unconditional mean ability. This confirms the fact that the equilibrium return to schooling is positive.

Second, Proposition 1 only requires that Assumption 1 holds for $d = 0$. On the contrary, the propositions gathered in the next subsection require Assumption 1 holds for all $d$, i.e. even if the distribution is truncated. This stronger requirement is needed because, as shown in the next subsection, some agents do not participate in the labor market in the efficient allocation, while they would be willing to participate in the decentralized equilibrium. Assumption 1 then boils down to assuming that we still have $\bar{a}_t < \bar{a}_h$ when the planner removes the least able agents from the unskilled market.

Third, Proposition 1 holds for all $A_l < A_h$. Actually, it also holds when $A_h$ is slightly lower than $A_l$ that is when the high-skill technology is (slightly) less efficient than the technology in the low-skill sector. High ability workers are willing to pay the schooling cost and to use a less efficient technology in order to benefit from the kind of sectoral peer effect that is induced by the combination of search frictions and endogenous job creation. In equation (4.7), the return to schooling can be positive even though $A_h < A_l$ provided that $\theta^s_h$ is sufficiently larger than $\theta^s_l$. From equation (4.5), this implies that $\bar{a}^s_h$ is sufficiently larger than $\bar{a}^s_l$. This property will not hold in the efficient allocation, as the planner will tend to allocate low ability workers to the high-skill sector in a similar situation.

Finally, Assumption 1 not only matters for the existence of an interior equilibrium but also reduces the scope for multiple equilibria. Consider for instance the following multiplier effect: as average ability in the high-skill sector increases, so does firms’ incentives to create jobs in this sector, and as more jobs are created, it increases the incentives for more people to get educated. If these marginal people are of the high ability, high-cost type, this may cause a further rise in the average ability in the high-skill sector, thus perpetuating the process. Under Assumption 1, the increase in the proportion of educated workers involves a decrease in the average ability of educated individuals and this multiplier effect does not take place.
4.2. Centralized economy

The characterization of the efficient allocation is derived from the first-order conditions to the maximization program (2.18). The computation is detailed in the Appendix.

**Definition 2** A characterization of the efficient allocation

An efficient allocation consists in a vector \((\theta^*_i, \theta^*_h, \bar{\alpha}^*_i, \bar{\alpha}^*_h)\) and three functions \(P^*_h, P^*_l, R^*\) mapping \(\mathbb{R}_+\) into \(\mathbb{R}\) such that

(i) job creation

\[
\gamma(\theta^*_i) = \alpha^*_i \frac{A_i}{\delta + (1 - \alpha^*_i) \mu(\theta^*_i)} \bar{\alpha}_i, \quad i = h, l
\]  

(ii) allocation of talents

\[
\overline{\alpha}^*_i = \mathbb{E}[a \mid c > R^* (a) a, P^*_h (a) \geq 0] \quad \text{and} \quad \overline{\alpha}^*_h = \mathbb{E}[a \mid c \leq R^* (a) a, P^*_h (a) \geq 0]
\]

(iii) return to schooling

\[
R^* (a) = P^*_h (a) - P^*_l (a)
\]

where

\[
P^*_i (a) = \frac{(1 - \alpha^*_i) \mu(\theta^*_i)}{\delta + (1 - \alpha^*_i) \mu(\theta^*_i)} \frac{A_i}{\delta} u^*_i + \frac{\alpha^*_i \mu(\theta^*_i) A_i a - \bar{\alpha}_i}{\delta}
\]

and \(\alpha^*_i \equiv \theta^*_i \mu'(\theta^*_i) / \mu(\theta^*_i), \quad u^*_i = \delta / (\delta + \mu(\theta^*_i))\).

Definition 2 states three properties. First, the return to ability in each sector depends on individual ability. Equation (4.13) can be written as follows:

\[
P^*_i (a) = P^*_i (\overline{\alpha}^*_i) + u^*_i \frac{\alpha^*_i}{1 - \alpha^*_i} P^*_i (\overline{\alpha}^*_i) a - \overline{\alpha}_i
\]

The first term is sector-specific and denotes the individual contribution to aggregate welfare at given sector-specific tightness. The second term is individual-specific and takes into account the impact of the agent on sector-specific tightness through congestion effects. It depends on ability: it is negative (positive) for individuals whose ability is below (above) the mean.

Second, the efficient allocation does not necessarily assign the best workers to the high-skill sector. The return to schooling \(R^* (a)\) equals the differential return to ability \(P^*_h (a) - P^*_l (a)\). When it is increasing in \(a\), the planner tends to assign high ability workers to education as the intuition suggests. However, the return to schooling may be decreasing in ability. In that case, the planner selects low ability workers into education. In the decentralized economy, high ability workers pay a price (the schooling cost) to benefit
from the presence of other high ability agents. This can be very expensive, especially when high ability workers tend to have a high schooling cost. The planner has another possibility: she can pay for bad workers and allocate them to the high-skill sector. This cannot happen in the decentralized economy as no one will be willing to pay to belong to a group of low ability persons.

Third, unlike the decentralized allocation, the efficient allocation is characterized by a set of non-participating agents, that is \( \Omega^* \neq \emptyset \). Non-participants are those whose social return to participation is negative in the low-skill sector, and lower than the schooling cost in the high-skill sector. Non-participants have low ability yielding insufficient output compared to the expected search spending. The discounted output produced by a worker in sector \( i \) is \( A_i a / \delta \). The expected recruitment cost is \( \gamma / \eta(\theta_i) \). Therefore, workers whose discounted output is lower than the recruitment cost should not search for jobs.

We now characterize the properties of \( R^*(a) \) and \( P^*_i(a) \).

**Proposition 2 Assortative matching in the efficient allocation**

If \( A_l \) is sufficiently lower than \( A_h \), the efficient allocation is such that (i) \( R^*(a) \) and \( R^*(a) a \) are strictly increasing in \( a \), (ii) \( P^*_i(a) \geq 0 \) when \( R^*(a) \geq 0 \), and (iii) \( \bar{a}_l^* < \bar{a}_h^* \).

When the technology in use in the high-skill sector is sufficiently better than the technology in use in the low-skill sector, the social return to schooling is increasing in ability. Points (i) and (ii) imply the efficient allocation can be depicted by Figure 4.1.

There are two lines. The set of non-participating agents \( \Omega^*_c \) is on the left of the participation line. The set of educated individuals \( \Omega^*_h \) is below the (positively sloped) indifference line. The set of uneducated participating agents \( \Omega^*_l \) is located between the two lines. The high-skill search market is tighter than the low-skill one, i.e. \( \theta^*_l < \theta^*_h \), the return to ability is larger in the high-skill sector, i.e. \( P^*_h(a) \geq P^*_i(a) \), and, finally, the return to schooling is positively related to ability.

Point (iii) means that the efficient allocation features positive assortative matching. The mean ability in the high-skill sector is larger than the mean ability in the low-skill sector, that is \( \bar{a}_l^* < \bar{a}_h^* \). This property is reasonable. However, as explained below Definition 2, the planner may well use education as a way to separate low ability workers from high ability individuals. Assumption 1 together with the requirement that the high-skill technology is sufficiently better than the low-skill technology guarantee that this phenomenon does not arise.

How much does the decentralized allocation differ from the efficient allocation? This issue is investigated in the next sub-section.
Figure 4.1: The efficient allocation of talents with search frictions. Agents whose characteristics are located on the left of the participation line do not participate. Agents whose characteristics are below the indifference line get educated. The others participate but remain uneducated.

4.3. The inefficiency of the decentralized outcome

We start the discussion by comparing private and social returns to schooling, assuming the Hosios condition $\beta = 1 - \alpha_i$ is satisfied in each sector, in order to focus on the externalities stemming from individual heterogeneity rather than those resulting from standard congestion externalities.

Proposition 3 Efficient vs equilibrium returns to schooling

Assume that $A_l$ is sufficiently lower than $A_h$, and that the Hosios condition holds. Then, there exists a unique $a_{\text{lim}} > 0$ such that $R^*(a) \leq R^*$ if and only if $a \leq a_{\text{lim}}$.

Unlike the social return to schooling, the private return does not depend on individual ability. People endowed with a low ability over-estimate the return to their schooling, while on the contrary, abler individuals underestimate their’s.

Proposition 3 allows us to discuss the efficiency of the decentralized allocation of talents across sectors when the Hosios condition holds. The situation is depicted by Figure 4.2, which confronts the decentralized allocation to the efficient one. This figure displays the following properties. First, workers endowed with a very low ability are excluded from the labor market in the efficient allocation, while they are not in the decentralized case.
Second, the slope of the social return to schooling is higher than the slope of the private one. This implies under-education for high schooling cost individuals – the poor – and over-education for low schooling cost persons – the rich. The threshold schooling cost separating rich from poor individuals is $c_{\lim} = R^* a_{\lim}$.

Over-education among the rich and under-education among the poor explain why the slope of the differential return to ability is larger in the efficient economy. On the one hand, the poorly talented but rich individuals who get education deteriorate the mean ability among the educated. This in turn reduces search prospects in the high-skill sector. On the other hand, highly talented but poor individuals who stay uneducated improve the mean ability among the uneducated. This raises wage and employment opportunities in the low-skill sector. Both effects lower the differential return to ability.

A deviation from the Hosios condition may alter the previous results. When the Hosios condition is not satisfied, the return to participation is inefficiently low in both sectors. The magnitude of such an inefficiency may differ across sectors, and, in turn, it may affect the private return to schooling. In many cases, the qualitative properties featured by Figure 4.2 would not be changed. Rich individuals would still crowd out the poor. However, in some pathological cases, the return to participation in the high-skill sector may be much less affected than the return to participation in the low-skill sector. The private return to schooling may then become larger than the efficient return at all costs.
schooling costs. Graphically, the slope of the private indifference line \( c = R^* a \) would be larger than the slope of the efficient line \( c = R^s (a) a \).

Generalized over-education, if it happens, is related to standard congestion externalities rather than to skill heterogeneity. To see this, suppose the distribution of ability is degenerate and consider the private and social returns to schooling:

\[
R^* = U_h (\beta) - U_l (\beta) \quad \text{and} \quad R^s = U_h (1 - \alpha_h) - U_l (1 - \alpha_l) \tag{4.15}
\]

where the equilibrium dependence vis-à-vis \( \beta \) has been highlighted. When the Hosios condition holds, job creation is efficient in each sector; so is the private return to search in each sector and, as a consequence, the private return to schooling equals the social return. However, when the Hosios condition is not satisfied, the return to search in the two sectors are generally too low, as \( U_i (\beta) < U_i (1 - \alpha_i) \). As a consequence, the private return to schooling generally differs from the social return. Overeducation may then take place whenever \( U_h (\beta) - U_h (1 - \alpha_h) > U_l (\beta) - U_l (1 - \alpha_l) \). For instance, this is the case when \( \beta = 1 - \alpha_h \neq 1 - \alpha_l \).

Those results generalize those in Charlot and Decreuse (2005). Compared to the current paper, our previous contribution corresponds to a case where the joint distribution of ability and schooling cost is degenerate, so that all agents are endowed with the same schooling cost irrespective of their ability. In this case, we show that overeducation prevails. In the current paper, the distribution of ability and schooling cost is not degenerate by assumption and the two models cannot be directly compared. However, we can arbitrarily increase the density \( \phi (a, c) \) around some \( \hat{c} \) so that \( \mathbb{E} (a \mid a > bc + d) \approx \mathbb{E} (a \mid a > bc + d, c = \hat{c}) \). The density \( \phi \) thus satisfies Assumption 1, since the conditional expectancy \( \mathbb{E} (a \mid a > bc + d) \) is increasing in \( b \hat{c} + d \). In that case, \( c_{\text{lim}} \) is larger than \( \hat{c} \) in Figure 4.2. Individuals whose schooling cost is \( \hat{c} \) always belong to the group of rich people, and some of them inefficiently choose to acquire education.\(^{11}\)

5. Education policy

In this section, we study the optimal education policy in the case where ability is not observable, whereas schooling costs are. The education policy consists of two instruments:

\(^{10}\)Bargaining powers may also differ across sectors. We did not consider this possibility for two reasons. First, wages are bargained individually in the model, and not collectively. The bargaining power is more likely individual-specific rather than sector-specific. Second, this would raise additional difficulties while stating the existence of a decentralized equilibrium.

\(^{11}\)Suppose \( \hat{c} > c_{\text{lim}} \) and let \((a^s, a^*)\) be such that \( R^i (a) a = \hat{c}, i = s, * \). Then, \( a^s < a^* \), which has two implications. First, \( R^s (a^*) > R^* \). Second, \( \sigma^s_h < \sigma^*_h \). These two implications are not compatible with each other. When \( A_l \) is sufficiently lower than \( A_h, \sigma^*_h < \sigma^*_h \) means that \( R^s (a^*) < R^* \), a contradiction.
a lump-sum fee granted to all individuals willing to educate, and a subsidy proportional to their schooling costs. We focus on the case where the Hosios condition is met, and show that on average education should be taxed more than it should be subsidized.

Specific assumptions. We investigate the possibility to alter the decentralized outcome through taxes and subsidies, so as to replicate the efficient allocation. For this question to make sense, we consider the realistic environment in which the planner cannot observe individual ability, while the schooling cost is observable: in several papers (see among others Heckman, 2000, Cameron and Heckman, 2001, Carneiro and Heckman, 2001), it is argued that children’s scholastic ability is strongly correlated to parental income. From this perspective, the individual schooling cost is observable as far as parents’ income is observable.

To simplify, we also assume the social planner does not address the market failure originating from the participation of very low ability agents to the low-skill sector. This can be so either because the minimum of the support of the skill distribution is larger than the (socially) optimal minimum ability required to participate in the low-skill sector, or because there are huge costs associated to the exclusion of a fraction of the population from all economic activities – at least higher than the gain induced by the nonparticipation of the least able.

Decentralizing the first best allocation. The planner has two sets of policy instruments. First, \( t_i, i = h, l \), is the sector-specific tax rate on output. Second, \( \tau \) is the proportion of private schooling cost financed by the State, while \( \tilde{\tau} \) is a lump-sum tuition fee on education. Hence, the actual cost of schooling is worth \((1 - \tau)(c + \tilde{\tau})\).\(^{12}\)

**Proposition 4** First best in the decentralized economy

Let \( \alpha_i \equiv \alpha(\theta_i) \), \( i = h, l \). The efficient allocation is decentralized iff

\[
(i) \quad 1 - t_i = \frac{\alpha_i}{1 - \alpha_i} \frac{u_i + \beta (1 - u_i)}{u_i + (1 - \alpha_i)(1 - u_i)}, \quad i = h, l
\]

\[
(ii) \quad \tilde{\tau} = \frac{\alpha_h}{1 - \alpha_h} u_h P^s_h (\bar{\alpha}^s_h) \bar{\alpha}^s_h - \frac{\alpha_l}{1 - \alpha_l} u_l P^s_l (\bar{\alpha}^s_l) \bar{\alpha}^s_l
\]

\[
(iii) \quad 1 - \tau = \beta \left[ \frac{\alpha_h}{1 - \alpha_h} P^e_h (\pi^e_h) - \frac{\alpha_l}{1 - \alpha_l} P^e_l (\pi^e_l) \right]
\]

Sector-specific taxes on output allow standard search externalities to be internalized, and compensate inadequate bargaining powers. Hence, the tax rate is positive if and only if \( \beta < 1 - \alpha_i \), it is nil when \( \beta = 1 - \alpha_l = 1 - \alpha_h \), and it is negative otherwise. The lump-sum tuition cost is equal to the social differential cost of participation to each sector; it deters the low ability agents with small schooling costs to invest in education. Finally, the proportion of schooling costs financed by the State \( \tau \) is smaller than 1.

---

\(^{12}\)The proceed of all taxes is redistributed through equal and non-distorsive lump-sum transfers. If negative, we assume all agents are endowed with enough wealth to finance the lump-sum tax.
When the Hosios condition is met. We now assume $\beta = 1 - \alpha_i^s = 1 - \alpha_h^s = 1 - \alpha$. To understand the optimal policy in this case, consider Figure 4.2. To make the efficient and the decentralized allocations coincide, it is necessary to shift the decentralized indifference line by means of a clockwise rotation to the left. This can be reached by means of a schooling fee, deterring the low ability and low schooling costs individuals to become educated, and a voucher increasing in schooling cost to attract abler workers with large schooling costs on the high-skill market.\textsuperscript{13} Hence,

\[
\bar{e} = \frac{\alpha}{1 - \alpha} \left[ u_h P_h^s (\bar{\pi}_h) \bar{\pi}_h^s - u_l P_l^s (\bar{\pi}_l) \bar{\pi}_l^s \right] \quad (5.1)
\]

\[
1 - \tau = \frac{P_h^s (\bar{\pi}_h^s) - P_l^s (\bar{\pi}_l^s)}{(1 - u_h) A_h / \delta - (1 - u_l) A_l / \delta} \quad (5.2)
\]

The education policy is redistributive: low schooling cost individuals are taxed, while high schooling costs individuals are subsidized. This raises the following question: On average, should education be taxed more than it should be subsidized? Formally, each individual pays an additional $(1 - \tau) \bar{e}$, but gets $\tau c$ back. Hence, $\tau c - (1 - \tau) \bar{e}$ is the voucher received by such an individual, and $\Delta = \int \int_{\Omega_h} [\tau c - (1 - \tau) \bar{e}] \phi(a, c) dadc$ is the net subsidy to education.

\textbf{Proposition 5} \textbf{The optimal policy under the Hosios condition}

\textit{Let $A_l$ be sufficiently lower than $A_h$ and assume that the Hosios condition holds. Then,}

\[
\Delta = \int \int_{\Omega_h} [\tau c - (1 - \tau) \bar{e}] \phi(a, c) dadc < 0 \quad (5.3)
\]

The optimal policy is to make individuals pay more on average than they would pay in the laissez-faire economy. The intuition for this result hinges on the social return to schooling given by equations (4.12) and (4.13) in Definition 2. It is strictly increasing in ability. However, it is also strictly concave. The social gain achieved by putting high ability individuals in the pool of educated workers is lower than the social gain obtained by deterring low ability workers to become educated.

This result generalizes Charlot and Decreuse (2005). When agents have the same schooling cost, they do not self-select enough on the basis of ability. The optimal education policy is to set a tax on education to deter too low ability agents from entering the educated sector. This result survives the consideration of schooling cost heterogeneity and the need to subsidize the education of high schooling cost and high ability individuals.

\textsuperscript{13}The optimal policy described here differs from what would result with credit market imperfections. In that latter case, income redistribution is a primary goal of education policy to decrease the aggregate schooling cost.
6. Conclusion

This paper studies the efficiency of educational choices in a two sector/two schooling level matching model of the labor market where a continuum of heterogenous workers allocates itself between sectors depending on their decision to invest in education. Individuals differ in ability and schooling cost, the search market is segmented by education, and there is free entry of new firms in each sector. Self-selection in education creates composition effects in the distribution of skills across sectors, as the distribution of ability among those who get education differs from the distribution of ability in the whole population. This in turn alters the intensity of job creation, implying the private and social returns to schooling always differ. Provided that ability and schooling cost are not too positively correlated, agents with large schooling costs – the ‘poor’ – underinvest in education, while there is overinvestment among the low schooling cost individuals – the ‘rich’. Efficiency can be restored by an education policy involving two instruments: a lump-sum tuition cost to deter low ability workers from schooling, and a subsidy increasing in private schooling costs to encourage more talented individuals with large schooling costs to get education. In addition, we show that education must be more taxed than subsidized in the case where the Hosios condition holds.
APPENDIX

A. Proofs

Characterization of the decentralized allocation (4.1) results from asset equations (2.9) to (2.11) combined with the bargaining rule (2.13), and imposing the free-entry condition \( V_i = 0 \). (4.2) results from (2.9) and (2.10) combined with the wage equation (4.1). (4.3) is then deduced from (4.2) and (2.15). The free-entry equation (4.5) results from \( V_i = 0 \); asset equation (2.11) and the wage equation (4.1).

Proof of proposition 1 We proceed in five steps. We start by the free-entry equations (4.5) determining sector-specific tightness.

Step 1. (i) At given \( \bar{a}_i > 0 \), there exists a unique \( \theta_i \equiv \Theta_i(\bar{a}_i) \) solving (4.5), \( i = h, l \).

(ii) \( \Theta_i \) is strictly increasing, with \( \lim_{a \to 0} \Theta_i(a) = 0 \) and \( \lim_{a \to \infty} \Theta_i(a) = \infty \).

(iii) \( \Theta_h(\bar{a}_h) > \Theta_l(\bar{a}_l) \) iff \( A_h\bar{a}_h > A_l\bar{a}_l \)

Proof. (i) Existence and uniqueness follow from Assumption 2. Claim (ii) results from the implicit function theorem, the fact the matching function satisfies the Inada conditions, and from its boundary properties. Claim (iii) follows directly from equations (4.5), \( i = h, l \).

We now consider the schooling return, rewritten here for convenience:

\[
R \equiv R(\theta_t, \theta_h) = \frac{\beta \mu_h(\theta_h)}{\delta + \beta \mu_h(\theta_h)} \frac{A_h}{\delta} - \frac{\beta \mu_l(\theta_l)}{\delta + \beta \mu_l(\theta_l)} \frac{A_l}{\delta} \tag{A.1}
\]

Step 2. (i) The function \( R \) is strictly decreasing in \( \theta_t \) and strictly increasing in \( \theta_h \)

(ii) \( R(\theta_t, \theta_h) \in [-A_l/\delta, A_h/\delta] \)

(iii) \( R(\theta_t, \theta_h) > 0 \) if \( \theta_h \geq \theta_t \)

Proof. (i) follows from the fact that the function \( \mu \) is strictly increasing in \( \theta_i \). Claim (ii) results from the boundary properties of the function \( \mu \). Claim (iii) is induced by \( A_h > A_l \).

Step 3. An interior equilibrium solves

\[
\bar{a}_i^* = \mathbb{E} [ a \mid c > R(\Theta_l(\bar{a}_l^*), \Theta_h(\bar{a}_h^*)) ] a \tag{A.2}
\]

\[
\bar{a}_h^* = \mathbb{E} [ a \mid c \leq R(\Theta_l(\bar{a}_l^*), \Theta_h(\bar{a}_h^*)) ] a \tag{A.3}
\]

with \( \theta_i^* = \Theta_i(\bar{a}_i^*), i = h, l \), and \( R(\Theta_l(\bar{a}_l^*), \Theta_h(\bar{a}_h^*)) > 0 \)

Proof. This is implied by Definition 1, step 1 and step 2.
We now proceed to solve the fixed point problem stated by equation (A.2) and (A.3). We define the function $\Psi : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+ \times \mathbb{R}_+$ such that $\Psi (a_l, a_h) = (\Psi_h (a_l, a_h), \Psi_l (a_l, a_h))$, where

$$
\Psi_l (a_l, a_h) = \mathbb{E} [a | c > R (\Theta_l (a_l), \Theta_h (a_h)) a]
$$
(A.4)

$$
\Psi_h (a_l, a_h) = \mathbb{E} [a | c \leq R (\Theta_l (a_l), \Theta_h (a_h)) a]
$$
(A.5)

**Step 4.** Let $\alpha_{\text{max}} = \max_{R \in [0, A_h/\delta]} \mathbb{E} [a | c \leq Ra]$ and $\Lambda = [0, \mathbb{E} (a)] \times [\mathbb{E} (a), \alpha_{\text{max}}]$. We have $\Psi (\Lambda) \subset \Lambda$

Proof. Take any pair $(a_l, a_h) \in \Lambda$. As $a_h \geq a_l$, step 1 implies $\Theta_h (a_h) > \Theta_l (a_l)$. Then, $R (\Theta_l (a_l), \Theta_h (a_h)) > 0$ from step 2. Assumption 1 and the definition of $\alpha_{\text{max}}$ imply

$$
\alpha_{\text{max}} \geq \mathbb{E} [a | c \leq R (\Theta_l (a_l), \Theta_h (a_h)) a] > \mathbb{E} (a)
$$
(A.6)

But, for all $b > 0$,

$$
x_h \mathbb{E} [a | c \leq ba] + (1 - x_h) \mathbb{E} [a | c > ba] = \mathbb{E} (a)
$$
(A.7)

where

$$
x_h = \int_a \int_{c \leq ba} \phi (a, c) dc da < 1
$$
(A.8)

Therefore, Assumption 1 also implies

$$
0 < \mathbb{E} [a | c > R (\Theta_l (a_l), \Theta_h (a_h)) a] < \mathbb{E} (a)
$$
(A.9)

It follows from inequalities (A.6) and (A.9) that $\Psi (a_l, a_h) \in \Lambda$, which establishes the claim.

**Step 5. (conclusion)** There exists an interior equilibrium with $\pi_i^* < \mathbb{E} (a) < \pi_h^*$

Proof. The set $\Lambda$ is compact and the function $\Psi$ is continuous. From the fixed-point theorem, there exists $(\pi_l^*, \pi_h^*) \in \Lambda$ such that $(\pi_l^*, \pi_h^*) = \Psi (\pi_l^*, \pi_h^*)$. To close the proof, note that inequalities (A.6) and (A.9) hold for all $(a_l, a_h) \in \Lambda$. Therefore, $\pi_l^* < \mathbb{E} (a) < \pi_h^*$.

**Characterization of the efficient allocation.** We derive the planner’s commands from (2.18). Let

$$
P_i^s (a) = (1 - u_i) \left[ \frac{A_i a}{\delta} - \frac{\gamma}{\eta_i} \right] \frac{1}{a}
$$
(A.10)

be the social return to participation in sector $i$ and

$$
R^s (a) = P_h^s (a) - P_i^s (a)
$$
(A.11)
be the social return to schooling. Also let \( p(a, c) \) be the proportion among workers of ability \( a \) and schooling cost \( c - (a, c) \)-workers – who participate in either one of the two search markets. Let also \( \pi_i (a, c) \) be the proportion of \( (a, c) \)-workers affected to the subset \( \Omega_i \). The planner’s objective (2.18) writes

\[
S = \max_{p, \pi, \theta} \left\{ \int_a^c p(a, c) \phi(a, c) \left( \sum_i \pi_i (a, c) P_i^s(a) - \pi_h (a, c) \right) dcda \right\} \quad (A.12)
\]

The maximization problem is subject to \( u_i = \mu(\theta_i) / (\delta + \mu(\theta_i)) \), \( i = h, l \), \( p(a, c) \in [0, 1] \) and \( \pi_i (a, c) = 1 - \pi_h (a, c) \in [0, 1] \).

For all \( (a, c) \in \Omega \), the first-order conditions write down

\[
p^s(a, c) = 1 \Leftrightarrow i \sum_i \pi_i^s (a, c) P_i^s(a) a - \pi_h^s (a, c) c \geq 0 \quad (A.13)
\]

\[
\pi^s(a, c) = 1 \Leftrightarrow R^s(a) a = \left[ P_h^s(a) - P_i^s(a) \right] a \geq c \quad (A.14)
\]

\[
\frac{\gamma}{\eta(\theta_i^s)} = \frac{\alpha(\theta_i^s)}{\delta + (1 - \alpha(\theta_i^s))\mu(\theta_i^s)}; i = h, l
\]

with

\[
\bar{\alpha}_i^s = \int_a^c \int_c \phi(a, c) \pi_i^s (a, c) p^s(a, c) adadc
\]

\[
\int_a^c \int_c \phi(a, c) \pi_i^s (a, c) p^s(a, c) dadc
\]

We impose \( p^s(a, c) = 1 \) when (A.13) holds with equality. Similarly, \( \pi_h^s (a, c) = 1 \) when \( R^s(a) a = c \). We can do so as the joint distribution over \( a \) and \( c \) has no mass point by assumption.

Finally, from (A.15) and (A.10), \( P_i^s(a) \) writes alternatively

\[
P_i^s(a) = (1 - u_i) \left[ \frac{A_i}{\delta} - \frac{\alpha(\theta_i^s) A_i}{\delta + (1 - \alpha(\theta_i^s))\mu(\theta_i^s) a} \right] \quad (A.17)
\]

\[
= \frac{(1 - \alpha_i^s) \mu(\theta_i^s) A_i}{\delta + (1 - \alpha_i^s) \mu(\theta_i^s) \delta} + u_i^s \frac{\alpha_i^s \mu(\theta_i^s)}{\delta + (1 - \alpha_i^s) \mu(\theta_i^s) \delta} A_i a - \bar{\alpha}_i
\]

\[
= P_i^s(\bar{\alpha}_i^s) + u_i^s \frac{\alpha_i^s}{1 - \alpha_i^s} P_i^s(\bar{\alpha}_i^s) \frac{a - \bar{\alpha}_i}{a} \quad (A.18)
\]

as appears in Section 4.

**Proof of proposition 2** We first prove that (i) and (ii) imply (iii). Then, we prove that (i) and (ii) hold when \( A_l \) is sufficiently lower than \( A_h \).

**Step 1. (i) and (ii) imply (iii).** (i) and (ii) characterize the allocation depicted by Figure 4.1. We have to show this allocation implies \( \bar{\alpha}_i^s > \bar{\alpha}_i^s \). But,

\[
\bar{\alpha}_i^s = \mathbb{E} \left[ a \mid c > R^s(a) , a \geq a_0 \right] \quad (A.20)
\]

\[
\bar{\alpha}_h^s = \mathbb{E} \left[ a \mid c \leq R^s(a) \right] = \mathbb{E} \left[ a \mid c \leq R^s(a) , a \geq a_0 \right] \quad (A.21)
\]
Assumption 1 implies

\[ \mathbb{E}[a \mid c > R^s(a), a \geq a_0] < \mathbb{E}[a \mid a \geq a_0] < \mathbb{E}[a \mid c \leq R^s(a), a \geq a_0] \]  

(A.22)

This establishes the claim.

Step 2. \( A_l \) sufficiently lower than \( A_h \) implies (i) and (ii). Let \( \Theta_i(a) \equiv \Theta(\beta, A_i a) \) and \( u_i \equiv \delta / (\delta + \mu(\Theta(\beta, A_i a_i))) \) to highlight the dependence vis-à-vis \( \beta \) and \( A_i \). The function \( \Theta \) is strictly increasing in its second argument from 0 to infinity. Therefore, \( (1 - u_i) A_i \) and \( (1 - u_i) \gamma / \eta(\Theta(\beta, A_i a)) \) go from 0 to infinity with \( A_i \) provided \( a_i > 0 \). Hence as far as \( \sigma^*_h > 0, R^s(a) \) and \( R^s(a) a \) are both strictly increasing in a provided \( A_i \) is sufficiently lower than \( A_h \). Now, let \( a_0 \) such that \( P^*_i(a_0) = 0 \). We have \( R^s(a_0) > 0 \) if \( A_i \) is sufficiently low. Thus, \( R^s(a) \geq 0 \) implies \( P^*_i(a) \geq 0 \). Now, we have to prove \( \sigma^*_h = 0 \) cannot be a solution to the maximization program. The (marginal) contribution of workers whose ability is close to \( a_0 \) and schooling cost close to 0 is \( \phi(a_0, 0) P^*_h(a_0) a_0 > 0 \) if assigned to the pool of educated workers. This closes the proof.

**Proof of proposition 3** We know from Proposition 2 that \( R^s \) is strictly increasing in \( a \). Therefore, we have to show that \( R^s(0) < 0 \) and \( R^s(\infty) > R^s \). As \( (1 - u_h) \gamma / \eta(\theta_h) > (1 - u_i) \gamma / \eta(\theta_i) \), we have \( \lim R^s(a) = -\infty \). Now, suppose \( R^s(\infty) \leq R^s \). Then \( \sigma^*_h > \sigma^*_h \). But \( R^s(\infty) > P^*_h(\sigma^*_h) - P^*_i(\sigma^*_i) \). As \( P^*_i \) is strictly increasing in its argument, and \( P^*_i \) is strictly increasing in \( A_i \) with \( \lim P^*_i(a) = 0 \), we have \( R^s(\infty) > P^*_i(\sigma^*_h) - P^*_i(\sigma^*_i) = R^* \) when \( A_i \) is sufficiently low. This contradicts the initial assumption. Therefore \( R^s(\infty) > R^s \).

**Proof of proposition 4** The policy is obtained by making the centralized and decentralized allocations coincide.

**Proof of proposition 5** Let \( P_i \equiv P^*_i(\sigma^*_i) \). If \( \beta = 1 - \alpha_i, i = h, l \), then \( t_i = 0 \) and (5.1) and (5.2) hold. Hence,

\[ \frac{1 - \tau}{\tau} \tilde{c} = \frac{P_h - P_l}{u_h P_h - u_l P_l} [u_h P_h \tilde{a}_h - u_l P_l \tilde{a}_l] \]
\[ > [P_h - P_l] \tilde{a}_h \equiv R^s(\sigma^*_h) \tilde{a}_h \]

Therefore,

\[ \Delta < \tau \int \int_{\Omega^*_h} [c - R^s(\sigma^*_h) \tilde{a}_h] \phi(a, c) da dc \]

From the f.o.c, \( (a, c) \in \Omega^*_h \) iff

\[ c \leq R^s(a) a = (P_h - P_l) a + \frac{\alpha}{1 - \alpha} [u_h P_h (a - \tilde{a}_h) - u_l P_l (a - \tilde{a}_l)] \]
Integrating over $\Omega_h^*$ both sides of the inequality yields

$$\int \int_{\Omega_h^*} c \phi(a, c) \, dcda < \int \int_{\Omega_h^*} (P_h - P_l) a \phi(a, c) \, dcda - \int \int_{\Omega_h^*} \frac{\alpha}{1 - \alpha} u_l P_l (a - \bar{a}_l) \phi(a, c) \, dcda$$

The result follows.
References


