Institute for Employment Research

The Research Institute of the Federal Employment Agency

IAB-Discussion Paper 7/2011

Articles on labour market issues

Income taxes, subsidies to education, and investments in human capital

Concetta Mendolicchio
Dimitri Paolini
Tito Pietra

Income taxes, subsidies to education, and investments in human capital

Concetta Mendolicchio (IAB)
Dimitri Paolini (DEIR and CRENoS, Università di Sassari)
Tito Pietra (DSE, Università di Bologna)

Mit der Reihe „IAB-Discussion Paper“ will das Forschungsinstitut der Bundesagentur für Arbeit den Dialog mit der externen Wissenschaft intensivieren. Durch die rasche Verbreitung von Forschungsergebnissen über das Internet soll noch vor Drucklegung Kritik angeregt und Qualität gesichert werden.

The “IAB Discussion Paper” is published by the research institute of the German Federal Employment Agency in order to intensify the dialogue with the scientific community. The prompt publication of the latest research results via the internet intends to stimulate criticism and to ensure research quality at an early stage before printing.

Contents

Abstract

We study a two-sector economy with investments in human and physical capital and imperfect labor markets. Human and physical capital are heterogeneous. Workers and firms endogenously select the sector they are active in and choose the amount of their sector-specific investments. To enter the high-skill sector, workers must pay a fixed cost that we interpret as a direct cost of education. Given the distribution of the agents across sectors, in equilibrium, in each sector there is underinvestment in both human and physical capital, due to non-contractibility of investments. A second source of inefficiency is related to the self-selection of the agents into the two sectors: typically too many workers invest in education. Under suitable restrictions on the parameters, the joint effect of the two distortions is that equilibria are characterized by too many people investing too little effort in the high skill sector. We also analyze the welfare properties of equilibria and study the effects of several tax policies on the total expected surplus. In particular, we consider the equilibrium associated with a flat labor income tax. Under suitable restrictions on the parameters, a revenue neutral progressive change in the marginal tax rates is welfare improving.

Zusammenfassung

Wir untersuchen eine Zwei-Sektoren-Ökonomie mit Investitionen in Human- und physisches Kapital und unvollkommenen Arbeitsmärkten. Human- und physisches Kapital sind heterogen. Arbeiter und Unternehmen wählen endogen sowohl den Sektor, in dem sie tätig werden, als auch die Menge der sektorspezifischen Investitionen. Für Arbeiter fallen Fixkosten an, falls sie im Sektor tätig werden wollen, der ausschließlich hoch qualifizierte Arbeiter beschäftigt. Für eine gegebene Verteilung der Agenten über die Sektoren ist das Gleichgewicht dieser Ökonomie durch Unterinvestition in Human- und physisches Kapital in beiden Sektoren gekennzeichnet. Ursächlich dafür ist die Annahme, dass Investitionen nicht vertragsfähig sind. Eine zweite Ursache von Ineffizienz ist die Selbstselektion von Agenten in die beiden Sektoren: typischerweise wählen zu viele Arbeiter den Sektor für Hochqualifizierte. Zusammen bewirken diese beiden Verzerrungen, dass im Gleichgewicht zu viele Arbeiter im Hochqualifizierten-Sektor tätig werden wollen, die dabei aber insgesamt zu wenig Bildungsanstrengung in Humankapital investieren. Weiter untersuchen wir die gleichgewichtigen Wohlfahrtswirkungen von Steuern. Es zeigt sich, dass für realistische Parameterrestriktionen eine budgetneutrale progressive Änderung der marginalen Steuersätze wohlfahrtssteigernd wirkt.

JEL classification: J24; H2

Keywords: Human capital; Efficiency; Labour income tax

Acknowledgements: This paper is part of the Ph.D. dissertation at IRES, Universitè Catholique de Louvain, of the first author. We thank for helpful comments the members of her dissertations committee (R. Boucchekine, V. Vandenberghe, M. Belot, B. Decreuse, B. Van der Linden), P. Pestieau, and participants of seminars at the University of Luxembourg, at the Università di Salerno, at ASSET 2009, at the 9^{th} Journees L.-A. Gerard Varet, at ESWC 2010 and at EEA 2010. We acknowledge the financial support of MIUR-PRIN 2006 and Fondazione Banco di Sardegna. The first author acknowledges the support of the “Programma Visiting Professor” of the Università di Sassari (Italy).

1. Introduction

In the last few decades, causes and consequences of investments in human capital have been a central field of research due to several motivations. Among them, the relevance of human capital externalities in growth theory, and the issues related to the dynamics of the wage premium and, more generally, to the evolution of income distribution. Still, the analysis of human capital externalities is far from settled from both the empirical and the theoretical viewpoints. Empirically, it is not obvious that there are significant, positive differences between social and private returns, at least at the level of subsidies prevailing in most Western countries.[1] From a theoretical viewpoint, the microeconomic mechanism generating the externality is not fully understood. A better understanding of its nature has policy relevance. This is true even if one is willing to take for granted that there are no significant, unexploited, positive externalities, because this is typically obtained with high subsidies to education.[2]

In this paper, we extend the microeconomic analysis of the distortions related to investments in human capital and derive some results on the welfare effects of different policies: fixed taxes/subsidies on the direct cost of the acquisition of high skill human capital, and taxes on labor income, or - equivalently in our set-up - on the investment in human capital.

We consider economies with three key features:

1. Ex-ante, workers are heterogeneous, while firms are identical,
2. Investments in human and physical capital are non-contractible,
3. There are two separate sectors employing different kinds of human and physical capital, so that an agent must choose both the level of his/her investment and its type.

The economy is basically a two-sector generalization, with sector specific inputs, of the model considered in Acemoglu (1996), which aims to provide an explicit equilibrium foundation for the existence of positive externalities related to human capital accumulation. In his framework, firms and workers choose the amount of their investments. Then, they are matched randomly (but preserving full employment), and income distribution is determined by a bargaining process. While a similar analysis could be carried out in several frameworks with the properties listed above, we focus the analysis on a Roy model of investments in human capital which is as close as possible to the one analyzed by Acemoglu. Indeed, after agents have chosen the sector they are going to be active in, i.e., the nature of their investment, our model reduces to a pair of separated Acemoglu's economies. In our set-up, income distribution takes place through bargaining, too. However, when workers are heterogeneous, the driving features of our results are asymmetric information on the workers' types and non-contractibility of the investments. The bargaining set-up is, of course, important, but it does not affect some key aspects of the welfare results.[3]

Our main departure from Acemoglu (1996) is that we adopt the notion of human capital put forth in Roy (1951): there are distinct markets for high skill and low skill labor, and we assume that they are perfectly non-substitutable. However, contrary to what is often assumed in Roy models, once a worker has selected the type of human capital she wants to acquire, she still has

[1] For the U.S.A., a negative conclusion is reached, for instance, by Heckman, Layne-Farrar, and Todd (1996) and by Acemoglu and Angrist (2001). For E.U. countries, the results in De la Fuente (2003) are also negative. See also Krueger and Lindhal (2000).

[2] In 2005, in the OECD average, 85.5% of the direct cost of education (all levels included) is financed by public sources (see OECD (2008, Table B3.1, p. 251)). The EU19 average is 90.5%. At the tertiary level, these percentages are, respectively, 73.1% and 82.5% (Table B3.2b, p. 253).

[3] Indeed, one can define economies with perfectly competitive spot labor markets, asymmetric information and lack of contractibility, where there is still a negative externality in human capital investments.

to decide how much effort to invest. Then, human capital translates one-to-one into efficiency units of high skill (low skill, respectively) labor.[4] Hence, each worker makes two separate choices, at the intensive and the extensive margin. Most of the recent literature takes a different point of view, adopting the efficiency units approach with homogeneous human capital, therefore ruling out, by assumption, all the consequences of self-selection of agents into different labor markets, which are, instead, relevant from both the theoretical and the empirical viewpoints.[5]

With imperfect markets and self-selection of workers into different labor markets, two distinct distortions are at work. Lack of contractibility of investments and the bargaining set-up generate a hold-up problem, inducing an inefficiently low level of investments, in human and physical capital of both types (hence, in each sector). This is the key mechanism at play in Acemoglu's paper. Secondly, given that workers are heterogeneous, when a subset of them switches from one sector to the other, there is an impact on the distribution of returns of the firms, hence on their optimal investments. In turn, this affects the optimal level of investments of workers. This second potential source of distortion is independent of the random matching set-up, and is at work even when spot labour markets are perfectly competitive (but lack of contractibility and asymmetric information hold).[6] This mechanism has been analyzed in the economics of education literature at least since Betts (1998).[7]

Therefore, in our set-up, public policies have two distinct effects on expected total surplus, our measure of welfare. The first is their impact on the level of the optimal investments of the agents acquiring a sector-specific skill: we will refer to it as incentive effect. The second is their impact on the agents' distribution across markets, i.e., the composition effect. In "pure" Roy models (with self-selection, but no choice of the investment effort) only the composition effect is at play. In "pure" efficiency-units models (without self-selection) only the incentive effect is at work. As usual, a hold-up problem on the returns on the investment in human capital induces underinvestment in education: less workers invest in education and each worker invests less effort than in the case of full appropriation of the marginal return of the investment. The impact on welfare of the composition effect is less obvious. An improvement of the conditionally expected level of human capital has always a positive effect on equilibrium utilities of all the workers and on the profits of the firms which remain active in the same sector. The expected producer's surplus of the firms which switch sector may actually decrease, but, under suitable restrictions on the parameters, the total effect is always positive. Bear in mind that, in our economy, there is always full employment and, therefore, the classical congestion externality, characterized by the violation of the Hosios condition, is absent.

We consider two separate sectors, using sector specific inputs (high/low skill human and physical capital). The crucial property is that human and physical capital are heterogeneous. To identify one type of capital with one sector somewhat simplifies the set-up and sharpens the welfare results. However, the two distinct distortions would be at work even with just one productive

[4] As usual, we can also interpret effort in the acquisition of human capital as elastic supply of labor of a given skill.

[5] A survey supporting this claim is in Sattinger (1993). A more recent discussions of the different empirical implications of efficiency units vs. Roy models is, for instance, in Carneiro, Heckman, and Vytlacil (2001). Investments in human capital in a two-sector economy with frictions due to random matching (but with perfectly inelastic supply of human and physical capital) have been studied in Sattinger (2003), Charlot and Decreuse (2005), and Mendolicchio, Paolini, and Pietra (2010).

[6] With perfectly competitive spot labor market, the hold-up problem disappears, and in each sector (taking as given the distribution af agents) investments are at their constrained efficient level. However, due to asymmetric information and lack of contractibility, the composition effect still induces constrained inefficiency of equilibria, which are always characterized by overinvestment in education.

[7] In the context of random matching models, it has been first studied in Charlot and Decreuse (2005).

sector employing both high and low skill labor, provided that there is a sufficient degree of substitutability. Bear in mind that whenever in what follows we mention the two sector structure of the economy, we implicitly mean that the two sectors use different kinds of human and physical capital.

Finally, in our full employment set-up, the elasticity of investments in physical capital plays a key role. Individual effort of workers depends upon the composition of the pool of workers in a sector only indirectly, because of its direct effect on the optimal level of investments in physical capital, which, in turn, is increasing in the conditional expectation of the effort of the workers active in a market. Hence, our model adopts, in a parsimonious way, the simplest structure of the economy which may deliver the basic insight.

We provide two sets of results concerning the efficiency properties of the equilibria. First, we show that an appropriate policy of subsidies to the investments and taxes on the direct costs of education can implement the constrained efficient allocation. Secondly, we consider the equilibrium associated with an arbitrary (but not too high) flat labor income tax and study the welfare effects of changes in the tax structure. This allows us to get some intuition concerning the relative magnitudes of incentive and composition effects. Some of the results in Acemoglu (1996) survive in our class of economies. For instance, in both cases, the human capital externality is related to its (sector-specific) average level. There are, on the other hand, sharp differences with respect to the policy prescriptions: in the one-sector model, subsidies to investments in human capital (or to labor supply) are unambiguously beneficial. This is because only the incentive effect is at play: a subsidy to the investments in human capital (or a reduction of the labor income tax rate) of any subset of agents increases them and, therefore, their expected value as a first order effect. This has a positive impact on the firms' investment decisions and, in turn, further increases the optimal investment of all the workers. This chain of positive feedbacks guarantees that this is welfare improving. To reformulate the point differently: in one-sector economies, there is a unique distortion induced by the hold-up problem which induces underinvestment for both firms and workers. Any policy increasing the investments of any subset of agents is welfare improving.

With two sectors, the incentive effect of a policy can be strengthened, weakened, or overturned, by its composition effect. Consider, for instance, a reduction in the marginal tax rate on low labor income (in our set-up: on the income of low-skill workers). If total factor productivities are sufficiently diverse across sectors and workers sufficiently heterogeneous, this always increases total surplus, because the positive effect on individual effort in the low-skill sector is strengthened by the composition effect, i.e., by the improvement of the expected human capital of the pool of workers in both markets. An increase in taxes on the direct costs of education also increases total surplus, just because of its composition effect. On the other hand, a decrease in the marginal income tax rate for high-skill workers has a (first order) positive incentive effect on their investments, but a negative composition effect. Hence, it always has a negative impact on the equilibrium utility of low-skill workers (and on the equilibrium profits of the firms active in that sector). The total effect for the agents active in the high-skill sector may be positive or negative, according to the magnitudes of the (positive) incentive effect and the (negative) composition effect. We provide a robust example where the effect of such a tax rate reduction on total surplus is negative. We conclude considering revenue neutral tax changes: the most interesting result is that, under our assumptions, a progressive change in the marginal labor income tax rates is welfare improving.

There is a large literature on the effects of subsidies to education and of labor income taxes on accumulation of human capital. The usual arguments favoring subsidies hinge either on their positive externality effects, or on the existence of liquidity constraints. Additionally, subsidies to education have been analyzed as one of the components of the optimal mix of redistributive policies (see Bovenberg and Jacobs (2005), Jacobs (2005, 2007), Jacobs and Bovenberg (2008), Jacobs, Schindler and Yang (2009), Schindler and Weigert (2008, 2009)). The last two aspects may be both empirically and theoretically important, but we abstract from them, focussing the analysis on the pure efficiency issue related to the presence of a hold-up problem and of self-selection. The classical analysis of the effects of labor income tax on investments in human capital started with the seminal papers by Ben-Porath (1970), Boskin (1975) and Heckman (1976).[8] A flat labor income tax has a negative impact on human capital accumulation just because of the non-deductibility of the direct costs of education. On the other hand, by depressing the net interest rate, in fully specified life-cycle models of consumer behavior, a tax on total income may actually have a positive effect. Eaton and Rosen (1980) extend the analysis to (uninsurable) multiplicative wage uncertainty, pointing out that a flat earning tax affects investments in human capital through its effects on their riskiness and (via an income effect) on the attitude toward risk (see, also, Anderberg and Andersson (2003), and Anderberg (2009)). Consider now a progressive income tax (compared with a revenue-neutral flat one). The canonical conclusion is that it discourages investments at the high skill level, while it may encourage them for the less skilled. While this literature provides us with many insights, it mostly deals with economies where there is no self-selection into different skills, so that one of the key mechanism at work in our economy is absent. Also, bear in mind that, in our set-up, at the equilibrium, workers face no uncertainty, so that the mechanism pointed out in Eaton and Rosen (1980) is absent.

A final remark: we consider investments in education as a benchmark case where heterogeneous agents make choices involving both the extensive and the intensive margins and where the composition effect matters. There are many other possible applications of the same basic framework, such as choices involving migration.

2. The Model

The economy is composed by two separate production sectors, denoted by $s \in \{ne, e\}$. Workers (denoted by a subscript i when we refer to individuals, I when we refer to their set) and firms (denoted by j and J, respectively) can choose to enter one of the two-sector, paying a fixed cost. Workers' costs, (c_I^{ne}, c_I^e), are exogenous, and can be interpreted as private, direct, fixed costs of education (tuitions and the like). We denote firms' costs (d_J^{ne}, d_J^e). They are endogenously determined, and will be discussed later on.

There are two intervals of equal length of workers and firms, $\Omega_I = \Omega_J \equiv [\underline{\theta}, \overline{\theta}] \subset \mathbb{R}_{++}$, both endowed with the Lebesgue measure. Each interval is partitioned into two sets, $\{\Omega_I^{ne}, \Omega_I^e\} \equiv \Omega_I^P$ and $\{\Omega_J^{ne}, \Omega_J^e\} \equiv \Omega_J^P$, determined endogenously. Let $\mu(\Omega_I^s)$ $(\mu(\Omega_J^s))$ denote the measure of the set Ω_I^s (Ω_J^s, respectively). In sector s, production requires a firm j (with physical capital k_j^s) and a worker i (with stock of human capital h_i^s). Once the partitions Ω_I^P and Ω_I^P are given, each sector of the economy reduces to the set-up studied in Acemoglu (1996). Firms are identical, and choose their investments in physical capital to maximize their expected profits. Workers choose their investments in human capital to maximize their expected utilities.

The economy lasts one period, divided into several subperiods. In subperiod 0, firms and workers enter one of the two sectors and carry out their investments. At 1, each firm active in sector s

[8] In our economy, one obtains substantially identical results considering direct (non-linear) subsidies to effort and subsidies to the direct costs of education. Previous, related work in this area includes Blankenau (2005), Blankenau and Camera (2006, 2009), Caucutt and Kumar (2003), Lloyd-Ellis (2000), Sahin (2004), and Su (2004).

is matched with a worker active in the same sector (we will be more precise on this issue later on). In the final subperiod, production takes place and the total output of each match is split according to the Nash bargaining solution with exogenous weights β and $(1-\beta)$.[9] Evidently, given that investments are carried out before matches take place, agents cannot contract with their partners a given level of investment. This is one of the key features of the economy.

For each worker active in sector s, the utility function is

$$U_i^s\left(C_i^s, h_i^s\right) = C_i^s - \frac{1}{\delta_i}\frac{h_i^{s(1+\Gamma)}}{1+\Gamma},$$

where C_i^s denotes consumption, h_i^s is the amount of human capital (or the labor supply). Let c_I^s be the (fixed) cost of the investment in sector s human capital. Then, in the absence of taxes and subsidies, if worker i is active in sector s and matched with firm j, C_i^s is given by labor income minus c_I^s. Workers are heterogeneous because of the parameter δ_i, indexing their marginal disutility of effort: ceteris paribus, larger values of δ_i are associated with higher values of the optimal choice of human capital. Without any essential loss of generality, we assume that $\delta_i = i$, and that δ_i is uniformly distributed on $[\underline{\theta}, \overline{\theta}]$, $\underline{\theta} > 0$. More general assumptions on the distribution of δ_i would not change any essential result.

Technologies are described by a pair of Cobb-Douglas production functions with constant returns to scale. When active in sector s, and matched with worker i with human capital h_i^s, firm j has production function

$$y_{ij}^s = A^s h_i^{s\alpha} k_j^{s(1-\alpha)},$$

with $A^e > A^{ne}$. Let p be the unit price of physical capital in both sectors. This implies some loss of generality, but simplifies notation and computations. Most important, similar results hold for $p^e \neq p^{ne}$.

As we will see, given any arbitrary partition of workers and firms (compatible with full employment), expected producers' surpluses are positive in both sectors and always larger in sector e. To avoid additional complications not really germane to our main issue, and to maintain the similarity with Acemoglu's model, we want to consider an economy with full employment at the equilibrium. This requires that, at the equilibrium, each agent is actually matched with a partner. We assume, as implicit in Acemoglu (1996), that the matching function guarantees with probability one a match to each agent, provided that $\mu\left(\Omega_I^s\right) = \mu\left(\Omega_J^s\right)$.[10] Given the focus of the paper, the partition Ω_I^P must be determined endogenously. Hence, to guarantee full employment, we need that, at each equilibrium, $\mu(\Omega_I^s) = \mu\left(\Omega_J^s\right)$. The easiest way to obtain this property is to introduce a feature of the economy such that equilibrium expected profits are always equal in the two sectors. One way to obtain this is to assume that the technology exploited in sector ne is free, while the one adopted in sector e is protected by a patent, owned by some outside agent.[11] The right to use the patent is auctioned off to firms before the firm-worker-match is obtained.[12] Given that, at an equilibrium, expected profits in both sectors must be identical, the equilibrium royalties must be equal to the (positive) difference between the expected producer's surpluses in the two sectors. Then, at each equilibrium, each firm is indifferent among sectors, so that we can choose Ω_J^P with $\mu\left(\Omega_I^s\right) = \mu\left(\Omega_J^s\right)$, the property we

[9] For a rationalization of this allocation rule in this context, see the Appendix in Acemoglu (1996). We assume that β is sector-invariant. Given that it is exogenous, to let it vary across sectors would just introduce more notation without providing any substantive additional insight.

[10] A commonly used function which delivers this property is $\pi_j^s = \frac{\min\left\{\mu(\Omega_I^s), \mu(\Omega_J^s)\right\}}{\mu(\Omega_J^s)}$, where π_j^s is the probability of a match for a firm active in sector s.

[11] Clearly, nothing would change if each technology were subject to a distinct patent.

[12] An auction delivering the result we need is based on closed envelope, first price bids by the firms. The royalty is allocated to each firm bidding the maximum price.

are looking for. Alternatively, one could adopt a structure based on a continuum of separate islands, with an identical "number" of firms and workers on each island, no mobility across islands and asymmetric information on the workers' types.[13]

Without any loss of generality, the prices of both kinds of output are set equal to 1 and, therefore, omitted.

Finally, notice that there are always three additional, trivial, equilibria: the ones where all the workers and the firms are in one of the two sectors, and the one where none is active in any sector. As usual, we ignore them.

3. Equilibrium

Later on, we will show that, at the equilibrium, it is always $\Omega_I^e = [\delta^F, \overline{\theta}]$ (or $\Omega_I^e = \emptyset$), where δ^F denote the equilibrium value of the threshold in the economy with frictions.[14] Hence, we can restrict the analysis to partitions Ω_I^P and Ω_J^P defined by an arbitrary level of the threshold, denoted $\widehat{\delta}$, and write $\Omega_J^s(\widehat{\delta})$ and $\Omega_I^s(\widehat{\delta})$.

For future reference, we determine the optimal amount of investments assuming that there is a public intervention defined by a pair of vectors $\xi^s \equiv (\tau^s, \zeta^s, \Delta c_I^s)$, $\xi \equiv (\xi^e, \xi^{ne})$, describing (possibly) sector specific subsidies and taxes. We assume that there are step-linear taxes on labor income (with rates τ^s, $s = ne, e$), and on the cost of the investments in physical capital (with rates ζ^s, $s = ne, e$), and fixed taxes, or subsidies, on the direct costs of education, Δc_I^s (we will always set $\Delta c_I^{ne} = 0$). We write the tax rates as sector specific just to simplify the notation: at equilibrium, this system of taxes is isomorphic to a system of step-linear taxes on labor income and on investments in physical capital.[15]

Pick an arbitrary threshold $\widehat{\delta}$. If active in sector s, firm j selects the value of k_j^s solving the expected profits maximization problem

$$\begin{aligned} \text{choose } k_j^s \quad &\in \quad \arg\max_{k_j^s} E_{\Omega_I^s(\widehat{\delta})} \left((1-\beta)\, A^s h_i^{s\alpha} k_j^{s(1-\alpha)} - p\,(1+\zeta^s)\, k_j^s \right) - d_J^s \\ &\equiv \quad (1-\beta)\, A^s E_{\Omega_I^s(\widehat{\delta})} \left(h_i^{s\alpha}\right) k_j^{s(1-\alpha)} - p\,(1+\zeta^s)\, k_j^s - d_J^s, \qquad (\Pi^s) \end{aligned}$$

where, given any random variable x^s, with $x^s : \Omega_I^s \to \mathbb{R}$, (or y^s, with $y^s : \Omega_J^s \to \mathbb{R}$), $E_{\Omega_I^s(\widehat{\delta})}(x_j^s) \equiv \frac{\int_{\Omega_I^s(\widehat{\delta})} x_i^s di}{\mu(\Omega_I^s(\widehat{\delta}))}$ (or $E_{\Omega_J^s(\widehat{\delta})}(y_i^s)$) denotes the conditional expectation of x_i^s over the set $\Omega_I^s(\widehat{\delta})$ (or of y_j^s over $\Omega_J^s(\widehat{\delta})$).

The pair of maps $K^s(\widehat{\delta};\xi)$, $s = ne, e$, defines the optimal investment in physical capital for the firms active in the two sectors. They are $j-$invariant because firms in each sector are identical, and depend upon the exogenous vector ξ, the arbitrary threshold $\widehat{\delta}$, and the conditional expectations $E_{\Omega_I^s(\widehat{\delta})}(h_i^{s\alpha})$. Let $\Pi^s(\delta_i, \widehat{\delta};\xi)$ be the surplus (because inclusive of d_J^s) of the firm matched with worker i in sector s.

[13] A third alternative would be to assume that firms cannot move across sectors. A non-null measure of firms is exogenously assigned to each sector. We then pick a matching function which always guarantees that each firm is matched with a worker (and conversely) for each non-trivial partition of the workers. As long as there is a continuum of agents in each sector, this can be done. Of course, this approach would break down if we had a finite number of agents and, anyhow, is based on a very ad hoc trick.

[14] Obviously, the worker with $\delta_i = \overline{\delta}$ is indifferent between the two sectors. For convenience, we assume that he/she enters sector e.

[15] Evidently, the same closed form could be obtained by using taxes (or subsidies) based on the effort in education, which, however, could not be directly observable.

The optimization problem of worker i (if active in s) is

$$\begin{aligned} &\text{choose } h_i^s \in \arg\max_{h_i^s} E_{\Omega_J^s(\widehat{\delta})}\left(U_i^s(.)\right) && (U^s) \\ \equiv \; & (1-\tau^s)\,\beta A^s h_i^{s\alpha} E_{\Omega_J^s(\widehat{\delta})}(k_j^{s(1-\alpha)}) - \frac{1}{\delta_i}\frac{h_i^{s(1+\Gamma)}}{1+\Gamma} - (c_I^s + \Delta c_I^s)\,. \end{aligned}$$

The pair of maps $H^s(\delta_i, \widehat{\delta}; \xi)$, $s = ne, e$, describes the optimal investments in human capital of the agents in each sector.

Let $V^s(\delta_i, \widehat{\delta}; \xi)$ be the associated level of utility of agent i, if active in sector s. Worker i enters sector e if and only if

$$F(\delta_i, \widehat{\delta}; \xi) \equiv V^e(\delta_i, \widehat{\delta}; \xi) - V^{ne}(\delta_i, \widehat{\delta}; \xi) \geq 0,$$

where $F(\delta_i, \widehat{\delta};\xi)$ is agent i's utility gain due to his investment in education. It is easy to check that, for each given $(\widehat{\delta}, \xi)$, $F(.)$ is strictly increasing in δ_i.

Definition 1. Given ξ, an equilibrium of the economy with frictions is a threshold value $\delta^F \in [\underline{\theta}, \overline{\theta}]$, and a royalty $d_J^{eF} \geq 0$, such that:
i. $K^s(\delta^F;\xi)$ solves (Π^s), $s = ne$ for each $j = i$ such that $\delta_i < \delta^F$, $s = e$ for each $j = i$ such that $\delta_i \geq \delta^F$;
ii. $H^s(\delta_i, \delta^F;\xi)$ solves (U^s), $s = ne$ for $\delta_i < \delta^F$, and $s = e$ for $\delta_i \geq \delta^F$;
iii. $E_{\Omega_I^e(\delta^F)}(\Pi^e(\delta_i, \delta^F;\xi)) - E_{\Omega_I^{ne}(\delta^F)}(\Pi^{ne}(\delta_i, \delta^F;\xi)) = d_J^{eF} > 0$;
iv. $F(\delta_i, \delta^F;\xi) \geq 0$ if and only if $\delta_i \geq \delta^F$.

First, observe that the conditional expectations $(E_{\Omega_J^s(\widehat{\delta})}(k_j^{s(1-\alpha)}), E_{\Omega_I^s(\widehat{\delta})}(h_i^{s\alpha}))$, $s = ne, e$, are computed making reference to the actual values $\{H^s(.), K^s(.)\}$, $s = ne, e$, so that we are imposing rational expectations. Conditions $(i - ii)$ impose individual optimality in the choice of the investment. Conditions $(iii - iv)$ impose individual optimality in the choice of the sector where an agent is active. By (iii), each firm is indifferent between being active in any of the two sectors, so that we can impose $\Omega_J^P = \Omega_I^P = \left\{[\underline{\theta}, \delta^F), [\delta^F, \overline{\theta}]\right\}$ (by iv).

The main results concerning existence of equilibria and their properties are summarized in Proposition 1. The proof is in the appendix. Here we just provide an outline of the argument. First, given an arbitrary $\widehat{\delta}$, we compute the values of $(\widetilde{H}^s(\delta_i, \widehat{\delta}; \xi), \widetilde{K}^s(\widehat{\delta}; \xi))$, $s = ne, e$, the demand functions for investment in human and physical capital obtained imposing that (conditional on $\widehat{\delta}$) expectations are fulfilled (see eqs. $(A3)$ and $(A4)$ in the appendix). Occasionally, we will refer to $(\widetilde{H}^s(\delta_i, \widehat{\delta};\xi), \widetilde{K}^s(\widehat{\delta};\xi))$ and the derived maps $\widetilde{V}^s(\delta_i, \widehat{\delta};\xi)$ and $\widetilde{\Pi}^s(\delta_i, \widehat{\delta};\xi)$ as the equilibrium maps conditional on $\widehat{\delta}$.

Let $\widetilde{F}(\delta_i, \widehat{\delta};\xi)$ be the analogous of map $F(.)$, obtained using $(\widetilde{H}^s(\delta_i, \widehat{\delta}; \xi), \widetilde{K}^s(\widehat{\delta}; \xi))$. Given that $\widetilde{F}(\delta_i, \widehat{\delta};\xi)$ is strictly increasing in δ_i, $\widetilde{F}(\delta_i, \widehat{\delta};\xi) = 0$ at $\delta_i = \widehat{\delta}$ gives us the equilibrium value of the threshold, i.e., $\delta^F(\xi)$. Hence, $\delta^F(\xi)$ is the solution to the equation

$$\widetilde{F}(\delta_i = \widehat{\delta}, \widehat{\delta}; \xi) \equiv f(\widehat{\delta}; \xi) - (c_I^e + \Delta c_I^e) = 0,$$

where, by direct computation (and using $(A3)$ and $(A4)$),

$$\begin{aligned} f(\widehat{\delta}; \xi) \;\equiv\; & \widehat{\delta}^{\frac{\alpha}{1+\Gamma-\alpha}} \left(A^e E_{\Omega_I^e(\widehat{\delta})}(\delta_i^{\frac{\alpha}{1+\Gamma-\alpha}})^{(1-\alpha)}\right)^{\frac{1+\Gamma}{\alpha\Gamma}} \chi^e(\xi) - \\ & \widehat{\delta}^{\frac{\alpha}{1+\Gamma-\alpha}} \left(A^{ne} E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\frac{\alpha}{1+\Gamma-\alpha}})^{(1-\alpha)}\right)^{\frac{1+\Gamma}{\alpha\Gamma}} \chi^{ne}(\xi)\,. \end{aligned} \qquad (1)$$

The variables $\chi^s(\xi), \chi^s(\xi) \equiv \frac{1+\Gamma-\alpha}{1+\Gamma}(1-\tau^s)^{\frac{1+\Gamma}{\Gamma}}\left(\alpha^{\frac{1}{\Gamma}}\beta^{\frac{1+\Gamma}{\Gamma}}\right)\left(\frac{(1-\alpha)(1-\beta)}{p(1+\zeta^s)}\right)^{\frac{(1+\Gamma)(1-\alpha)}{\alpha\Gamma}}$, are scalars depending upon the exogenous parameters.

Proposition 1. Fix (α,β) and let $\xi = (\tau,\tau,0,0,0)$. Given $(\Gamma, A^e, A^{ne};\xi)$, there are $\left\{\underline{C},\overline{C}\right\} >> 0$ such that, for almost every $c_I^e \in (\underline{C},\widetilde{C})$, there is an equilibrium with threshold value $\delta^F(\xi) \in \left(\underline{\theta},\overline{\theta}\right)$. Moreover, given (Γ, A^{ne}), there is $\underline{A}^e$ such that, for each $A^e > \underline{A}^e$, the equilibrium is unique and $\frac{\partial f(.)}{\partial\widehat{\delta}}|_{\widehat{\delta}=\delta^F} > 0$. Also, $\frac{\partial\delta^F(.)}{\partial\tau^e} > 0$, $\frac{\partial\delta^F(.)}{\partial\tau^{ne}} < 0$, $\frac{\partial\delta^F(.)}{\partial\Delta c_I^e} > 0$, $\frac{\partial\delta^F(.)}{\partial A^e} < 0$ and $\frac{\partial\delta^F(.)}{\partial A^{ne}} > 0$, where $\delta^F(A^e, A^{ne};\xi)$ is the function associating with the vector ξ the (unique) equilibrium threshold. The same results hold, given (A^e, A^{ne}), for Γ sufficiently small.

Proof. See the appendix. ∎

Given the focus of the paper, it is convenient to consider a vector ξ with the stated properties, just to simplify computations. Nothing relevant depends upon this restriction. In the following, we will mostly consider the leading case where $\frac{\partial f(.)}{\partial\widehat{\delta}}|_{\widehat{\delta}=\delta^F} > 0$ at each equilibrium threshold.[16] This restriction delivers two different properties for equilibria. First, $\frac{\partial f(.)}{\partial\widehat{\delta}} > 0$ at each equilibrium threshold implies its uniqueness. Secondly, by the implicit function theorem, the comparative statics properties depend upon the derivatives of the equilibrium conditions with respect to the exogenous parameters (A^e, A^{ne},ξ) and $\widehat{\delta}$. The signs of these derivatives with respect to (A^e, A^{ne},ξ) are always uniquely defined. Hence, the comparative statics of the equilibrium threshold just depends upon the sign of $\frac{\partial f(.)}{\partial\widehat{\delta}}|_{\widehat{\delta}=\delta^F}$, and to restrict the analysis to economies with $\frac{\partial f(.)}{\partial\widehat{\delta}}|_{\widehat{\delta}=\delta^F} > 0$ at each δ^F allows us to obtain well-defined results. Different sets of restrictions on the parameters would guarantee that $\frac{\partial f(.)}{\partial\widehat{\delta}}|_{\widehat{\delta}=\delta^F} > 0$. The ones proposed above seem fairly weak and natural. That some additional restrictions are necessary to obtain $\frac{\partial f(.)}{\partial\widehat{\delta}}|_{\widehat{\delta}=\delta^F} > 0$ at each equilibrium is shown in Example A1 (in the appendix). There we construct an economy with $\frac{\partial f(.)}{\partial\widehat{\delta}} > 0$ for $\widehat{\delta}$ sufficiently close to $\underline{\theta}$ and negative for $\widehat{\delta}$ large enough. Given that $\frac{\partial f(.)}{\partial\widehat{\delta}}$ is continuous on $[\underline{\theta},\overline{\theta}]$, for this economy $f(.)$ has at least one local maximum, $\overline{\delta}$. Hence, each economy with c_I^e such that $c_I^e < f(\overline{\delta})$, and close enough to $f(\overline{\delta})$, has at least two equilibria. The precise magnitude of the restriction on the ratio A^e/A^{ne} obviously depend upon the precise values of (α,β,Γ). Numerical simulation suggest that they are not overly restrictive.

Let's compare the equilibrium allocation of this economy to the one of the associated Walrasian economy (the one with perfect contractibility and competitive wages). There are three main results. Fix $\xi = 0$. Pareto inefficiency of equilibria is obvious, because, in the economy with frictions, a firm's investment does not depend upon the value of δ_i of the worker it is matched with, while it does at any Pareto efficient allocation. Second, the Walrasian equilibrium of the same economy dominates the equilibrium of this economy in terms of total expected surplus, but it is not necessarily Pareto superior. Indeed, at $\xi = 0$, and using $(A3)$ and $(A4)$ in the appendix, the physical/human capital ratio at the two allocations satisfy

$$\frac{\widetilde{K}^s(\delta^F)}{\widetilde{H}^s(\delta_i,\delta^F)} = \left(\frac{(1-\beta)^{\frac{1}{\alpha}}E_{\Omega_I^s(\delta^F)}(\delta_i^{\frac{a}{1+\Gamma-\alpha}})^{\frac{1}{\alpha}}}{\delta_i^{\frac{1}{1+\Gamma-\alpha}}}\right)\frac{K^{Ws}(\delta_i)}{H^{Ws}(\delta_i)},$$

where the superscript "W" denotes the equilibrium values at the Walrasian allocation. If δ^F is large enough, compared to $\underline{\theta}$, and for sufficiently small δ_i, in sector ne the term in brackets is always greater than one, so that $\frac{\widetilde{K}^{ne}(\delta^F)}{\widetilde{H}^{ne}(\delta_i,\delta^F)} > \frac{K^{Wne}(\delta_i)}{H^{Wne}(\delta_i)}$. This immediately implies that agents with a sufficiently low δ_i are better off at the equilibrium of the frictional economy. A third

[16] To avoid misunderstandings: $\widetilde{F}(\delta_i,\widehat{\delta},\xi)$ is always strictly increasing in δ_i. The function $f(\widehat{\delta},\xi)$ is obtained setting $\delta_i = \widehat{\delta}$ and it does not necessarily have this property.

observation is that the threshold value δ^F can be either lower or higher than its value in the Walrasian economy. For instance, let $\xi = 0$, set $[\underline{\theta}, \overline{\theta}] = [1, 2]$, $A^e = 2$, $A^{ne} = 1$, $\alpha = \beta = 1/2$, and $\Gamma = 2$. By direct computation, one can verify that, for $c_I^e < 0.7$, $\delta^F < \delta^W$, while the opposite occurs for $c_I^e > 0.71$. Hence, lack of contractibility always induces Pareto inefficiency because of lower than optimal investments, while it has an ambiguous effect on the size of the set of people investing in education. From this viewpoint, therefore, it does not induce unambiguously overeducation (or undereducation).

Finally, consider the asymptotic behavior of the equilibrium allocation along any sequence $\{A^{ev}\}_{v=1}^{v=\infty}$ with $A^{ev} \to A^{ne}$. Let $f(\widehat{\delta}, A^e, A^{ne})$ be the function obtained from $f(\widehat{\delta};\xi)$ setting $\xi = 0$ and making explicit its dependence on (A^e, A^{ne}) (similarly for $\delta^F(A^e, A^{ne})$). It is easy to check that $\lim_{\widehat{\delta} \to \overline{\theta}} f(\widehat{\delta}, A^e, A^{ne}) > 0$, for each $A^e \geq A^{ne}$. Hence, there is an interval of values of c_I^e such that the associated equilibrium threshold is strictly smaller than $\overline{\theta}$ even if $A^e = A^{ne}$.[17] Hence, the equilibrium investments in high skill human capital is positive even when this skill is completely useless, from the technological viewpoint. When $A^e = A^{ne}$, the two sectors are essentially identical, while to operate in sector e, requires the use of costlier skills. Therefore, Pareto efficiency requires us to shut down this sector. This is similar to what happens in signalling models.

The main purpose of the paper is to analyze the policy implications of workers' self-selection into distinct labor markets. However, it is interesting to consider the comparative statics of equilibria, also because the welfare effects of different policies comes through their impact on the equilibrium values of the endogenous variables.

Let $\phi \equiv (\xi, A^e, A^{ne})$. Let $w^s(\delta_i, \delta^F, \phi)$ be worker i's wage in sector s. The standard deviation, $\sigma_{\Omega_I^s(\delta^F)}(\delta^F, \phi)$, measures the variability of wages within sector s. $WP_{\Omega_I^e(\delta^F)}(\delta^F, \phi)$ is the wage premium.[18]

Proposition 2. Fix $(\Gamma, c_I^e, \alpha, \beta)$. Assume that $\frac{\partial f(.)}{\partial \widehat{\delta}}|_{\widehat{\delta} = \delta^F(\xi)} > 0$. At $\xi = 0$, the following sign restrictions are satisfied:[19]

$$\begin{bmatrix} & d\tau^e & d\tau^{ne} & \Delta c_I^e & dA^e & dA^{ne} \\ E_{\Omega_I^e(\delta^F(\xi))}(\widetilde{H}^e(.)) & ? & - & + & ? & + \\ E_{\Omega_I^{ne}(\delta^F(\xi))}(\widetilde{H}^{ne}(.)) & + & - & + & - & + \\ \widetilde{K}^e(.) & ? & - & + & ? & + \\ \widetilde{K}^{ne}(.) & + & - & + & - & + \\ E_{\Omega_I^e(\delta^F(\xi))}(w^e(.)) & ? & - & + & ? & + \\ E_{\Omega_I^{ne}(\delta^F(\xi))}(w^{ne}(.)) & + & - & + & - & + \\ \sigma_{\Omega_I^{ne}(\delta^F(\xi))}(.) & + & - & + & - & + \\ WP_{\Omega_I^e(\delta^F(\xi))}(.) & - & + & - & + & - \end{bmatrix}.$$

[17] Depending upon the values of the other parameters, we may have (at least) two equilibria with different thresholds, or a unique equilibrium. What is relevant is that there is always some level of investment in "high skills".

[18] In general, there are three different notions of wage premium: $\frac{w^e(.)}{w^{ne}(.)}$, $E_{\Omega_I^e(\delta^F)}\left(\frac{w^e(.)}{w^{ne}(.)}\right)$ and $E_{\Omega_I^{ne}(\delta^F)}\left(\frac{w^e(.)}{w^{ne}(.)}\right)$. Due to linearity of the wage function with respect to $(\delta_i^{\frac{\alpha}{1+\Gamma-\alpha}})$ in each sector, here they coincide.

[19] Each cell reports the sign of the derivative of the function on the row with respect to the variable on its column. We omit the standard deviation of the wages of skilled workers. For this variable, it is impossible to reach any well-defined, general result. For reasonable values of the parameters, $\alpha = \frac{2}{3}$ and $\Gamma > \frac{1}{2}$, some numerical simulations show that the composition effect has the sign opposite to the one of $\frac{\partial \delta^F}{\partial \phi}$. Therefore, $\frac{\partial \sigma_{\Omega_I^e(\delta^F)}(.)}{\partial \phi}$ is positive for $\phi' \in \{\tau^e, A^e\}$, negative for $\phi' \in \{\tau^{ne}, \Delta c_I^e, A^{ne}\}$.

All the results follow by tedious, but straightforward, computations. The intuition behind them is based on the interaction of the incentive and the composition effect. For instance, consider an increase in τ^{ne}, i.e., in the marginal tax rate on the labor income of the ne workers. As a pure incentive effect, $d\tau^{ne} > 0$ reduces their effort in education, and pushes down the threshold δ^F. Hence, because of the composition effect, it reduces the (conditional) expected human capital of both low and high skilled workers. This, in turn, reduces investments in physical capital in both sectors. This negative feed-back strengthens the initial impacts. Hence, the effects on expected human and physical capitals and on wages are negative in both sectors. For the wage premium, by direct computation, it turns out that both direct and composition effects are positive. The standard deviation of wages of unskilled workers decreases because both effects are negative. On the other hand, $d\tau^e > 0$ has unambiguously a positive effect on the level of human (and, consequently, of physical) capital and on the wages in the ne sector, because it increases the value of $\delta^F(\xi)$ (composition effect). Indeed, given that the expected human capital of the pool of ne workers increases because of the increase in τ^e, physical capital also increases, stimulating these workers' optimal investments. The impact in the e sector is ambiguous because the incentive effect reduces the optimal investment in human capital. However, the composition effect acts in the opposite direction, because it induces workers with a (relatively) low value of δ_i to switch to the low skill sector. This has a positive impact on the expected level of human capital in the e sector and, therefore, on the investments in physical capital, inducing a positive feed-back. The effect of exogenous changes in technology, (dA^e, dA^{ne}), can be explained basically in the same way. In particular, in this set-up, "skill biased" technical change $(dA^e > 0, dA^{ne} = 0)$ has a negative impact on the expected human, and physical, capital and on the wages in the "low skill" sector, an ambiguous impact in the "high skill" one, and a positive effect on the wage premium.

4. Efficiency properties of equilibria

We have already argued that the equilibria of the economy with frictions are Pareto inefficient. We will now show that they do not satisfy either a weaker criterion of constrained optimality (CO in the sequel) which takes into account the imperfections which characterize the economy. Most interesting is the analysis of their inefficiency in terms of the amount, and type, of investments. In the sequel, we will mainly refer to investments in human capital. Similar considerations hold for the ones in physical capital.

In our set-up, inefficiencies can be of two different types. First, an individual can choose an amount of investment different from the CO one, given the partition Ω_I^P associated with the CO allocation. We will refer to this possible source of inefficiency as underinvestment (or over-investment) in educational effort. Secondly, an agent can choose to invest in a type of education different from the one assigned to her at the CO allocation. We will say that there is underinvestment in educational level when agent i invests in education ne, while, at the CO allocation, she should invest in education level e.

In the one-sector model, equilibria are unambiguously characterized by underinvestment in educational effort. Here, the same effect is at work: in each sector, given any arbitrary $\widehat{\delta}$, an increase in the investments of firms and workers leads to a Pareto improvement. Once we consider an arbitrarily fixed threshold $\widehat{\delta}$, the argument is identical to the one in Acemoglu (1996). Set $\xi = 0$ (and omit it, for notational convenience). Fix $\widehat{\delta}$, so that each sector is identical to the economy analyzed there, and consider a small change in the investment of each agent. The changes in utilities and producers' surplus evaluated at the equilibrium (conditional on $\widehat{\delta}$) pair $(\widetilde{H}^s(\delta_i, \widehat{\delta}), \widetilde{K}^s(\widehat{\delta}))$ (and taking into account that investments in physical capital are $j-$invariant) are given by

$$\left(\alpha\beta A^s\left[\frac{\widetilde{K}^s(\widehat{\delta})}{\widetilde{H}^s(\delta_i,\widehat{\delta})}\right]^{1-\alpha}-\frac{1}{\delta_i}\widetilde{H}^s(\delta_i,\widehat{\delta})^{\Gamma}\right)dh+\left((1-\alpha)\,\beta A^s\left(\frac{\widetilde{H}^s(\delta_i,\widehat{\delta})}{\widetilde{K}^s(\widehat{\delta})}\right)^{\alpha}\right)dk>0 \quad (2)$$

and

$$\left((1-\alpha)\,(1-\beta)\,A^s\frac{E_{\Omega_I^s(\widehat{\delta})}(\widetilde{H}^s(\delta_i,\widehat{\delta})^{\alpha})}{\widetilde{K}^s(\widehat{\delta})^{\alpha}}-p\right)dk+\left(\alpha\,(1-\beta)\,A^s\frac{\widetilde{K}^s(\widehat{\delta})^{(1-\alpha)}}{E_{\Omega_I^s(\widehat{\delta})}(\widetilde{H}^s(\delta_i,\widehat{\delta})^{(1-\alpha)})}\right)dh>0, \quad (3)$$

respectively. The inequalities hold because the first terms in parentheses in (2) and (3) are zero, at the optimal solutions of (Π^s) and (U^s), while the second terms are positive. Hence, given any $\widehat{\delta}$, there is underinvestment in educational effort and physical capital, in each sector. This establishes, in a more direct way, the Pareto inefficiency of the equilibria of our economy.

In the two-sector case, there is a second potential source of inefficiency, because changes in the value of $\widehat{\delta}$ may also entail Pareto improvements. An increase in the threshold value $\widehat{\delta}$ increases the conditional expected amount of human capital in both sectors at the same time and, consequently, induces an increase in the amount of physical investments of firms in both sectors. Indeed, given that $\delta_i^{\frac{\alpha}{1+\Gamma-\alpha}}$ is a strictly monotonically increasing function,

$$\frac{\partial E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\frac{\alpha}{1+\Gamma-\alpha}})}{\partial\widehat{\delta}}>0, \text{ for each } s \text{ and } \widehat{\delta}, \quad (4)$$

and, consequently, using $(A3)$ and $(A4)$, $\frac{\partial\widetilde{H}^s(\delta_i,\widehat{\delta})}{\partial\widehat{\delta}}>0$ and $\frac{\partial\widetilde{K}^s(\widehat{\delta})}{\partial\widehat{\delta}}>0$, for each s and $\widehat{\delta}$. More relevant, from $(A5)$, $(A6)$ and (4), for each i and $\widehat{\delta}$, $\frac{\partial\widetilde{V}^s(\delta_i,\widehat{\delta})}{\partial\widehat{\delta}}>0$ and $\frac{\partial E_{\Omega_I^s(\widehat{\delta})}(\widetilde{\Pi}^s(\delta_i,\widehat{\delta}))}{\partial\widehat{\delta}}>0$.

These properties do not suffice to establish our claim, because a change in the threshold induces a jump in the producer's surplus for the firms shifting from one sector to the other. We will get back to this issue later on.

To analyze the welfare properties of equilibria, it is convenient to introduce an explicit notion of (constrained) efficiency. As usual in economies with imperfect markets, we consider the metaphor of a benevolent planner choosing an allocation while facing constraints aiming to capture the ones the agents face in the decentralized economy. We provide two results. First, we show that there are constrained optimal allocations (Proposition 3), and that they can be attained with an appropriate system of taxes and subsidies (Corollary 1). The amount of subsidies and taxes is entirely dictated by the features of the CO allocation, and they can be quite large. That's why, in Prop. 4 and 5, we study the effects of small changes in taxes and subsidies on total surplus evaluated at the market equilibrium, taking as given the actual demand and supply functions of the agents. Proposition 4 considers generic changes in taxes and subsidies. In Proposition 5, we consider revenue neutral changes.

Bear in mind that, in the following, we always consider changes in total surplus. We are not concerned with actual Pareto improvements. However, given that utility functions are quasi-linear, an increase in total surplus immediately translates (modulo an appropriate - and $i-$contingent - system of lump-sum taxes and transfers) into a Pareto improvement.

4.1. Constrained optimal allocations

The objective function of the planner is $P\left(h_i^s, k_j^s, \Omega_I^s, \Omega_J^s\right)$, the sum of the expected utilities and producers' surpluses of the agents. His policy instruments are the partitions Ω_I^P and Ω_J^P and a pair of maps $(H^{COs}(\delta_i, \widehat{\delta}), K^{COs}(\widehat{\delta}))$. We restrict the partitions to have the structure $\Omega_I^e(\widehat{\delta}) = \Omega_J^e(\widehat{\delta}) = [\widehat{\delta}, \overline{\theta}]$. Given that firms are ex-ante identical, the informational constraints embedded into the definition of $P(.)$, and the properties of the (implicit) matching function, to impose this structure on Ω_I^P and Ω_J^P does not entail any loss of generality. Also, observe that, given that firms are identical, expected total surplus and realized total surplus coincide.

We define an allocation Constrained Optimal (or CO) if and only if it solves the planner's optimization problem. Let δ^{CO} be the level of the threshold associated with the CO allocation.

Proposition 3. Under the maintained assumptions, there is a CO allocation. Equilibrium allocations are never CO, and are characterized by underinvestment in the amount of physical capital and in educational effort. $\delta^F > \delta^{CO}$ and $\delta^F < \delta^{CO}$ can both occur.

Proof. See the appendix. ∎

The source of inefficiency considered by Acemoglu (1996) reappears in our set-up, because, given any threshold level $\widehat{\delta}$, $H^{COs}(\delta_i, \widehat{\delta}) > \widetilde{H}^s(\delta_i, \widehat{\delta})$, for each δ_i, and $K^{COs}(\widehat{\delta}) > \widetilde{K}^s(\widehat{\delta})$. On the other hand, the relation between δ^{CO} and δ^F is not univocal. In the proof in the appendix, we provide an example of an economy such that $\delta^F < \delta^{CO}$ if the direct costs of education are sufficiently low, while $\delta^F > \delta^{CO}$ for sufficiently high values of c_I^e. In interpreting this result, bear in mind that, in computing δ^F and δ^{CO}, we use different investment functions: $(\widetilde{H}^s(.), \widetilde{K}^s(.))$ and $\left(H^{COs}(.), K^{COs}(.)\right)$, respectively. On the other hand, in Corollary 1, we show that, once the optimal subsidies (τ, ζ) are introduced, to implement the CO allocation we always need $\Delta c_I^e > 0$. Thus, given the optimal taxes, CO always requires us to shrink the set of agents investing in the high skill sector.

It is easy to see that the CO distribution of investments in human and physical capital can be attained with an appropriate system of taxes and subsidies. Given that preferences are quasi-linear, the system of tax and subsidies can be balanced using uniform lump-sum taxes on workers (in the absence of positive endowments of consumption goods, this could entail negative consumption for some subset of agents).

Corollary 1. There is a system of taxes and subsidies ξ, with $\Delta c_I^e > 0$, such that the associated equilibrium allocation is CO.

Proof. See the appendix. ∎

In our set-up (as well as in Acemoglu (1996)), equilibria of the economy with frictions are constrained inefficient for each value of β, because, at $\xi = 0$, even if $\delta^{CO} = \delta^F$, for each β,

$$\frac{\widetilde{H}^s(\delta_i, \delta^{CO})}{H^{COs}(\delta_i, \delta^{CO})} = (1-\beta)^{\frac{1-\alpha}{\alpha\Gamma}} \beta^{\frac{1}{\Gamma}} \neq 1, \text{ for each } s \text{ and } i,$$

and

$$\frac{\widetilde{K}^s(\delta^{CO})}{K^{COs}(\delta^{CO})} = (1-\beta)^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} \beta^{\frac{1}{\Gamma}} \neq 1, \text{ for each } s.$$

In the usual random matching model, constrained efficiency is obtained when the Hosios-condition is satisfied, i.e., when β is equal to the absolute value of the elasticity of the matching function. On the contrary, in our economy, given any threshold $\widehat{\delta}$, as observed in Acemoglu (1996, p. 789), the externalities are related to "the value of the future matches and are always positive". Moreover, the distribution of workers across sectors may fail to be optimal, but this induces a market failure different from the one due to the congestion externality characterizing the usual economies with random matching. This is why here the Hosios-condition has no connection with efficiency.

4.2. Welfare improving tax policies

We conclude considering the welfare effects of alternative tax schemes. Tax changes have two distinct effects. First, they may change the marginal return of the investment in effort, given the type of skill an individual acquires. This is a direct incentive effect. Second, they affect the distribution of workers in the two sectors. This is the composition effect, which, in turn, changes the marginal return of the investment in effort because of its effect on the optimal level of investment in physical capital. The role of the "composition effect" in our economy has peculiar features. A marginal change in the threshold $\delta(\xi)$ has no direct effect on total workers' surplus, because, by definition of equilibrium, $V^{ne}(\delta_i = \delta^F(\xi), \delta^F(\xi)) = V^e(\delta_i = \delta^F(\xi), \delta^F(\xi))$. Similarly, given that a firm's expected profits are equal across sectors, the "direct" composition effect on total expected profits is zero. Unfortunately, in our set-up, due to the (strictly positive) royalties, typically, at the margin, the producer's surplus in the high skill sector is strictly larger than in the other sector. This follows from the particular structure of our economy, that we have justified above. In the welfare analysis, this makes it harder for our main results to hold, because the indirect impact of the composition effect has to be sufficiently large to compensate its (negative) direct impact on welfare. This indirect effect goes first through the positive impact of the increase in the (conditional) expectation of the level of human capital in each sector on the firms' investments in physical capital. Generally speaking, the mechanism at work here holds true in a stronger form in any economy with self-selection of agents in distinct sectors and where there is some positive feed-back between the variables of interest and the conditional expectation of some feature of the pool of agents self-selecting in one market.[20]

Therefore, in the proof of the two final Propositions, we need additional restrictions on the exogenous parameters, sufficient to guarantee that the composition effect on welfare of an increase in the threshold is positive. They are formulated implicitly, as an upper bound on the value of the equilibrium threshold (i.e., on the direct cost of education, given the other parameters). They do not appear unreasonable. For instance, fix, as usual, $\alpha = \frac{2}{3}$, $\beta = \frac{1}{10}$ and $\xi = 0$. Set $\Omega_I = [1, 4]$. For $\Gamma \leq 1$, i.e., given an elastic effort supply, and $\frac{A^{ne}}{A^e} = 0.9$, the composition effect is always positive for $\delta^F \leq 1.7$. The interval of values of δ^F such that it is positive is decreasing in Γ.

Consider as a starting point an economy with a flat labor income tax. An increase of taxes on the direct cost of education (Δc_I^e) has a pure composition effect, due to quasi-linearity of the utility functions. Changes in the marginal tax rates have both incentive and composition effects: As obvious, an increase in the marginal rate on high income individuals (i.e., in our set up, on the high skill workers) has a negative direct incentive effect, but a positive composition effect. Changes in the marginal tax rate on the low income workers have negative incentive

[20] Charlot and Decreuse (2005) consider a two-sectors, dynamic random search model. There, an increase in the value of the threshold improves the conditional expectation of the productivities of the workers in both sectors. This makes it profitable for firms to create new vacancies and, therefore, leads to a decrease in unemployment in each sector. This may entail a welfare improvement. In a way, in their model creation of vacancies has a role similar to the one played here by the increase in physical investments. A similar mechanism is also at play in the model with islands and perfectly competitive spot labor markets outlined at the end of Section 2.

and composition effects. Hence, they have, unambiguously, a negative impact on welfare. The intuition is fairly simple, also given Prop. 2 above. For instance, $d\tau^{ne} > 0$ has a direct, negative incentive effect on effort in this sector. It also makes convenient for some subset of workers to move to sector e, so that it moves down the value of the threshold. Hence, it has a negative composition effect on effort and, in turn, on investments in physical capital, in both sectors. To the contrary, in the case of $d\tau^{e} > 0$, the impacts on total surplus of incentive and composition effects have opposite signs and, under suitable conditions, the second can actually dominate, so that we can obtain a welfare improvement by moving from a flat income tax to a progressive one.

Let's make formal the heuristic argument above. Given ξ, workers and firms choose their individually optimal behavior. Let $S(\xi)$ be the expected total surplus at the equilibrium associated with the vector ξ of policy instruments. Let $R(\xi)$ be the total tax revenue. Then,

$$S(\delta^F(\xi);\xi) \equiv \sum_s \left(\int_{\Omega_J^s(\delta^F(\xi))} E_{\Omega_I^s(\delta^F(\xi))}(\widetilde{\Pi}^s(.))dj + \int_{\Omega_I^s(\delta^F(\xi))} \widetilde{V}^s(.)di \right) + R(\delta^F(\xi);\xi)).$$

The first set of results concerns the effects of a change of one of the tax rates.

Proposition 4. Consider an equilibrium associated with an arbitrary $\overline{\xi} = (\overline{\tau}, \overline{\tau}, \Delta\overline{c}_I^e)$, $\overline{\tau} > 0$ and sufficiently small, and satisfying $\frac{\partial f(.)}{\partial \widehat{\delta}}|_{\widehat{\delta}=\delta^F(\xi)} > 0$. Then, for $[\underline{\theta}, \overline{\theta}]$ large enough and $\delta^F(\xi)$ sufficiently close to $\underline{\theta}$,

i. $d\Delta c_I^e > 0$, and sufficiently small, increases total surplus,

ii. $d\tau^{ne} < 0$, and sufficiently small, increases total surplus,

iii. $d\tau^{e} < 0$, and sufficiently small, may decrease total surplus.

The proofs of (i, ii) are in the appendix, where we also establish that the welfare effect of a change of τ^e is, in general, indeterminate. The third statement is shown in Example A3, also in the appendix. The two assumptions on the support $[\underline{\theta}, \overline{\theta}]$ and the value of $\delta^F(\xi)$ guarantee that the composition effect $\frac{\partial S(\delta^F(\xi);\xi)}{\partial \delta^F(\xi)}$ is positive.

Changes in expected surplus are our measure of welfare gains and losses. However, different policy instruments have different implications also in terms of individual welfare. Under the maintained assumptions, a decrease in the value of τ^{ne} (or an increase of Δc_I^e) has a positive impact on the utility level of all the workers and on the expected profits of each firm (the effect on the equilibrium level of the royalties is, however, indeterminate, in general). On the contrary, a decrease in τ^e has always a negative impact on the utility of all the workers in sector ne (and on the expected surplus of all the firms active in this sector). It may have a positive or negative impact on utility and surplus of agents active in sector e, and on the equilibrium royalties.

Consider now policies where reductions in the income taxes are financed through taxes on the direct costs of education, or by revenue neutral changes $(d\tau^e, d\tau^{ne})$.

Proposition 5. Let $\xi = (\overline{\tau}, \overline{\tau}, \Delta\overline{c}_I^e)$, with $\Delta\overline{c}_I^e \leq 0$. Consider balanced budget policies $(d\tau^e, d\tau^{ne})$ and $(d\tau^s, d\Delta c_I^e)$. Under the assumptions of Prop. 1, $(d\tau^e, -d\tau^{ne}) >> 0$ and $(-d\tau^{ne}, d\Delta c_I^e) >> (0,0)$ increase expected total surplus, $(-d\tau^e, d\Delta c_I^e) >> (0,0)$ may decrease it.

The proof is in the appendix. The first result implies that some (small) degree of progressiveness in the labor income taxation is welfare improving. Under the assumptions of Prop. 5, an increase in the value of the equilibrium threshold is welfare improving. In the proof, we show that a revenue neutral policy $(d\tau^e, -d\tau^{ne}) >> 0$ always has a positive effect on $\delta^F(\xi)$. Hence, the composition effect increases welfare. It also turns out that, for $\overline{\tau}$ sufficiently small (we need to be on the increasing part of the Laffer curve), the direct incentive effect of a revenue neutral tax change is also positive. Therefore, this policy change is unambiguously welfare improving.

The second result can be explained along the same lines. On the other hand, a decrease in τ^e, balanced by an increase in Δc_I^e has an ambiguous effect on welfare. The pure incentive effect of the policy is welfare improving. The difference with respect to the previous case is that now the total effect of the policy on the equilibrium threshold depends in a non-trivial way upon the parameters, because $d\tau^e < 0$ makes investment in education more appealing, while $d\Delta c_I^e > 0$ acts in the opposite direction. The total effect on welfare is, therefore, indeterminate. However, in general, the revenue neutral policy $(-d\tau^e, \Delta c_I^e) >> 0$ has a larger positive (or a smaller negative) effect on welfare than a pure reduction of the marginal labor income tax rate.

Finally, we have been considering a sector-contingent vector of subsidy rates (τ^e, τ^{ne}). This is certainly an unusual feature of the policy. However, let $w^s(\delta_i, \delta^F)$ be agent i's labor income in sector s. It is easy to check that

$$\max_{\Omega_I^{ne}(\delta^F)} w^{ne}(\delta_i, \delta^F) \leq w^{ne}(\delta^F, \delta^F) < w^e(\delta^F, \delta^F) \leq \min_{\Omega_I^e(\delta^F)} w^e(\delta_i, \delta^F).$$

Hence, given the properties of the utility functions, the same results can be obtained with a standard system of step-linear taxes or subsidies.

5. Conclusions

The paper considers a class of economies where we model both extensive and intensive margins of investment choices. The main conclusion is that the results typically obtained in an efficiency unit set-up (which considers only the intensive margin) can fail to be robust to its natural extension to a Roy's model with optimal choice of investments in human and physical capital. The efficiency unit framework rules out, by assumption, all the phenomena induced by the self-selection of the agents into different labor markets and, therefore, all the welfare consequences related to the composition effect.

Our analysis is carried out for a simple, parametric class of economies. This allows us to compute explicitly the equilibria and the welfare effects of different policies, and to compare directly our results with the ones of Acemoglu (1996). Evidently, to consider quasi-linear utility function is restrictive, in particular in the analysis of the welfare impact of the various policies. However, first, an extension of the analysis to a richer environment is possible, but at a high cost in terms of analytical tractability. Secondly, all the results are "open", so that they certainly survive in environments where income effects are sufficiently small. What matters most, the basic intuition behind the welfare results is strong, and they should be robust to many possible extensions of the basic set-up.

There are two main messages of the paper: in environments characterized by lack of contractibility, irreversibility of the investments in human capital generates a hold-up problem. This tends to depress investments below their optimal level, so that a pecuniary externality in human capital is generated. However, if workers self-select into distinct labor markets by investing in different types of human capital, a second distortion arises whenever wages are an increasing function of the conditional expectation of the level of human capital of workers active in a market. In our model this is induced by the positive effect of this expectation on the level of the investments in physical capital. This second externality may induce overinvestment in education at the extensive margin. While both phenomena have been previously discussed in the literature, we provide a relatively simple set-up where we can analyze their joint effect on welfare. From a strictly theoretical viewpoint, our result is in the same spirit of Acemoglu (1996): There is a positive externality induced by the conditional expectations of the level of human capital investments in each sector. It acts primarily via their effects on investments in physical capital. The policy implications are, however, sharply different. Without self-selection, any policy which provides a positive incentive to the workers' choice on the intensive margin is welfare improving. With self-selection, the welfare improving effect of such a policy must be evaluated also taking into account its, possibly welfare worsening, impact on the choices at the extensive margin. For instance: we have established that, starting with a flat labor income tax rate, a small, revenue neutral, change in the marginal tax rates, making the tax system progressive, can be welfare improving. It is straightforward to check that this can never happen in the one sector version of the model.

6. Appendix

3A. Equilibrium

We start with an arbitrary threshold $\widehat{\delta}$. The first order conditions (FOCs in the sequel) of problem (Π^s) imply

$$K_j^s(E_{\Omega_I^s(\widehat{\delta})}(h_i^{s\alpha});\xi) = \left[\frac{(1-\beta)(1-\alpha)A^s E_{\Omega_I^s(\widehat{\delta})}(h_i^{s\alpha})}{p(1+\zeta^s)}\right]^{\frac{1}{\alpha}} \qquad (A1)$$

The ones of optimization problem (U^s) imply

$$H_i^s(E_{\Omega_J^s(\widehat{\delta})}(k_j^{s1-\alpha});\xi) = \left[\delta_i\alpha\beta(1-\tau^s)A^s E_{\Omega_J^s(\widehat{\delta})}(k_j^{s1-\alpha})\right]^{\frac{1}{1+\Gamma-\alpha}}. \qquad (A2)$$

Given that firms in sector s are, ex-ante, identical, $K_j^s(.) = K^s(.)$, so that $E_{\Omega_J^s(\widehat{\delta})}(K_j^s(.)^{1-\alpha}) = K^s(.)^{1-\alpha}$.

Let $\gamma \equiv \frac{1+\Gamma}{1+\Gamma-\alpha}$, so that $(\gamma-1) \equiv \frac{\alpha}{1+\Gamma-\alpha}$.

Solving $(A1)$ and $(A2)$, by imposing that expectations are fulfilled, we obtain

$$\begin{aligned}\widetilde{K}^s(\widehat{\delta};\xi) \;&=\; \left[\frac{(1-\alpha)(1-\beta)}{p(1+\zeta^s)}E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1})\right]^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} \\ &\quad \times((1-\tau^s)\alpha\beta)^{\frac{1}{\Gamma}}A^{s\frac{1+\Gamma}{\alpha\Gamma}},\end{aligned} \qquad (A3)$$

and

$$\begin{aligned}\widetilde{H}^s(\delta_i,\widehat{\delta};\xi) \;&=\; \left[\frac{(1-\alpha)(1-\beta)}{p(1+\zeta^s)}E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1})\right]^{\frac{1-\alpha}{\alpha\Gamma}} \\ &\quad \times\delta_i^{\frac{1}{1+\Gamma-\alpha}}((1-\tau^s)\alpha\beta)^{\frac{1}{\Gamma}}A^{s\frac{1}{\alpha\Gamma}}.\end{aligned} \qquad (A4)$$

Using these functions, agent i's utility, at the $\widehat{\delta}-$conditional equilibrium and if active in sector s, is

$$\begin{aligned}\widetilde{V}^s(\delta_i,\widehat{\delta};\xi) \;&\equiv\; U_i^s(\widetilde{H}^s(\delta_i,\widehat{\delta};\xi),\widetilde{K}^s(\widehat{\delta};\xi)) = -(c_I^s+\Delta c_I^s) \\ &\quad + \left[\frac{(1-\alpha)(1-\beta)}{p(1+\zeta^s)}E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1})\right]^{\frac{(1+\Gamma)(1-\alpha)}{\alpha\Gamma}}\end{aligned} \qquad (A5)$$

$$\times \delta_i^{\gamma-1} \left((1-\tau^s) \beta \right)^{\frac{1+\Gamma}{\Gamma}} A^{s\frac{1+\Gamma}{\alpha\Gamma}} \alpha^{\frac{1}{\Gamma}} \frac{1}{\gamma}.$$

Similarly, given an arbitrary $\widehat{\delta}$, firm j (ex-post) surplus, if active in sector s and matched with worker i, is

$$\widetilde{\Pi}^s(\delta_i, \widehat{\delta}; \xi) = (1-\beta) A^{s\frac{1+\Gamma}{\alpha\Gamma}} \left((1-\tau^s) \alpha\beta \right)^{\frac{1}{\Gamma}} \left(\delta_i^{\gamma-1} - (1-\alpha) E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1}) \right) \times \left(\frac{(1-\alpha)(1-\beta)}{p(1+\zeta^s)} E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1}) \right)^{\frac{(1+\Gamma)(1-\alpha)}{\alpha\Gamma}}. \qquad (A6)$$

Its expected value is

$$E_{\Omega_I^s(\widehat{\delta})}(\widetilde{\Pi}^s(\delta_i, \widehat{\delta}; \xi)) = \left[\frac{(1-\alpha)(1-\beta)}{p(1+\zeta^s)} E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1}) \right]^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} \times \frac{p(1+\zeta^s)\alpha\left((1-\tau^s)\alpha\beta \right)^{\frac{1}{\Gamma}} A^{s\frac{1+\Gamma}{\alpha\Gamma}}}{(1-\alpha)}. \qquad (A7)$$

Proof of Prop. 1. Consider an arbitrary vector ξ, with $\tau^e = \tau^{ne}$ and $\zeta = 0$, so that $\chi^e(\xi) = \chi^{ne}(\xi)$.

Pick the partition $\Omega_I^P(\widehat{\delta})$ induced by any arbitrary $\widehat{\delta}$. Assume that there is an agent i' such that $\delta_{i'} = \widehat{\delta}$ at $\widehat{\delta}$ solving $F(\widehat{\delta}, \widehat{\delta}) = 0$. Evidently, $\widetilde{F}(\delta_i, \widehat{\delta}) \geq 0$ if and only if $\delta_i \geq \widehat{\delta}$. Hence, each equilibrium partition Ω_I^P such that $\Omega_I^s \neq \emptyset$, each s, satisfies $\Omega_I^e(\delta^F) = [\delta^F, \overline{\theta}]$, as claimed in the text.

For each threshold $\widehat{\delta}$, and each s, $E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1})$ is the conditional expectation of a strictly increasing function, hence it is strictly increasing in $\widehat{\delta}$ and well-defined on $\left[\underline{\theta}, \overline{\theta}\right]$. It follows that $f(\widehat{\delta})$ is continuous and strictly positive for each $\widehat{\delta} \in \left[\underline{\theta}, \overline{\theta}\right]$, including the boundary points. Let $\min_{[\underline{\theta},\overline{\theta}]} f(\widehat{\delta}) \equiv \underline{C} \geq 0$ and $\overline{C} \equiv \max_{[\underline{\theta},\overline{\theta}]} f(\widehat{\delta}) > \underline{C}$, because $f(\widehat{\delta})$ is clearly not constant over $\left[\underline{\theta}, \overline{\theta}\right]$. Then, by the intermediate value theorem, for each c_I^e such that $c_I^e \in \left[\underline{C}, \overline{C}\right]$, there is $\delta^F(\xi) \in \left[\underline{\theta}, \overline{\theta}\right]$ such that $f(\delta^F(\xi), \delta^F(\xi)) - c_I^e = 0$. Evidently, for most of the values of c_I^e, $\delta^F(\xi) \in \left(\underline{\theta}, \overline{\theta}\right)$.

Using $(A7)$, and given that $E_{\Omega_I^e(\widehat{\delta})}(\delta_i^{\gamma-1}) > E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\gamma-1})$, and $A^e > A^{ne}$,

$$d_J^{eF} = \left[E_{\Omega_I^e(\delta^F(\xi))}(\widetilde{\Pi}^e(\delta_i, \delta^F(\xi))) - E_{\Omega_I^{ne}(\delta^F(\xi))}(\widetilde{\Pi}^{ne}(\delta_i, \delta^F(\xi))) \right] > 0.$$

Hence, all the equilibrium conditions are satisfied at $\delta^F(\xi)$. This establishes the first part of the Proposition.

We now proceed to study uniqueness of equilibrium and its comparative statics properties. Clearly, $\frac{\partial \widetilde{F}(.)}{\partial \Delta c_I^e} = -1 < 0$, and

$$\frac{\partial f(.)}{\partial \tau^s} = (-1)^{\varphi(s)} \delta^{F\gamma-1} (A^s E_{\Omega_I^s(\delta^F)}(\delta_i^{\gamma-1})^{(1-\alpha)})^{\frac{1+\Gamma}{\alpha\Gamma}} \frac{\chi^s(\xi)}{(1-\tau^s)} \frac{1+\Gamma}{\Gamma},$$

with $\varphi(e) = 1$ and $\varphi(ne) = 2$, so that $\frac{\partial f(.)}{\partial \tau^e} < 0$ and $\frac{\partial f(.)}{\partial \tau^{ne}} > 0$. Also, $\frac{\partial f(.)}{\partial A^e} > 0$ and $\frac{\partial f(.)}{\partial A^{ne}} < 0$. Hence, the signs of the comparative statics properties, and uniqueness of equilibrium, could be immediately established if the sign of $\frac{\partial f(.)}{\partial \widehat{\delta}}|_{\widehat{\delta}=\delta^F(\xi)}$ were uniquely defined. Unfortunately, this is not the case. As established in Example A1 below, there are economies with multiple equilibria and where, obviously, $sign|\frac{\partial f(.)}{\partial \widehat{\delta}}|_{\widehat{\delta}=\delta^{F}(\xi)}$ varies across them. Hence, to establish the second part of Prop. 1, we need to impose additional restrictions on the parameter space. By direct computation,

$$\begin{aligned}\frac{\partial f(.)}{\partial \widehat{\delta}} &= (\gamma-1)\frac{1}{\widehat{\delta}}f(.)+\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}\frac{\widehat{\delta}^{\gamma-1}}{\widehat{\delta}}\chi(\xi)\\ &\quad [A^{e\frac{1+\Gamma}{\alpha\Gamma}}E_{\Omega_I^e(\widehat{\delta})}(\delta_i^{\gamma-1})^{\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}}\eta^e(\widehat{\delta})-\chi(\xi)A^{ne\frac{1+\Gamma}{\alpha\Gamma}}E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\gamma-1})^{\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}}\eta^{ne}(\widehat{\delta})],\end{aligned}$$

where $\eta^s(\widehat{\delta})$ is the elasticity of $E_{\Omega_I^s(\widehat{\delta})}(\delta_i^{\gamma-1})$ with respect to $\widehat{\delta}$,

$$\eta^e(\widehat{\delta})\equiv\frac{-\gamma\widehat{\delta}^{\gamma}(\overline{\theta}-\widehat{\delta})+\widehat{\delta}(\overline{\theta}^{\gamma}-\widehat{\delta}^{\gamma})}{(\overline{\theta}-\widehat{\delta})(\overline{\theta}^{\gamma}-\widehat{\delta}^{\gamma})}\text{ and }\eta^{ne}(\widehat{\delta})\equiv\frac{\gamma\widehat{\delta}^{\gamma}(\widehat{\delta}-\underline{\theta})-\widehat{\delta}(\widehat{\delta}^{\gamma}-\underline{\theta}^{\gamma})}{(\widehat{\delta}-\underline{\theta})(\widehat{\delta}^{\gamma}-\underline{\theta}^{\gamma})}.$$

With a straightforward manipulation, we obtain

$$sign\frac{\partial f(.)}{\partial \widehat{\delta}}=sign\left(\begin{array}{c}(\gamma-1)\left(1-\left(\frac{A^{ne}}{A^e}\right)^{\frac{1+\Gamma}{\alpha\Gamma}}\left(\frac{E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\gamma-1})}{E_{\Omega_I^{e}(\widehat{\delta})}(\delta_i^{\gamma-1})}\right)^{\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}}\right)\\ +\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}\left(\eta^e(\widehat{\delta})-\left(\frac{A^{ne}}{A^e}\right)^{\frac{1+\Gamma}{\alpha\Gamma}}\left(\frac{E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\gamma-1})}{E_{\Omega_I^{e}(\widehat{\delta})}(\delta_i^{\gamma-1})}\right)^{\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}}\eta^{ne}(\widehat{\delta})\right)\end{array}\right).$$

The first term is strictly positive.

Define $G^e(\delta)\equiv\delta(\overline{\theta}^{\gamma}-\delta^{\gamma})-\gamma\delta^{\gamma}\left(\overline{\theta}-\delta\right)$ as the numerator of $\eta^e(\overline{\delta})$. Geometrically, it is easy to see that $G^e(\underline{\theta})>0$. Also, $G^e(\overline{\theta})=0$. If there is $\widetilde{\delta}$ such that $G^e(\widetilde{\delta})<0$, there must also be $\overline{\delta}$ such that $G^e(\overline{\delta})<0$ and $\frac{\partial G^e(\delta)}{\partial\delta}|_{\overline{\delta}=\delta}>0$. However,

$$\begin{aligned}\frac{\partial G^e(\delta)}{\partial\delta} &= \overline{\theta}^{\gamma}-(\gamma+1)\delta^{\gamma}-\gamma^2\overline{\theta}\delta^{\gamma-1}+\gamma(\gamma+1)\delta^{\gamma}\\ &= (\overline{\theta}^{\gamma}-\delta^{\gamma})-\gamma^2\delta^{\gamma-1}(\overline{\theta}-\widehat{\delta})<0\end{aligned}$$

at each $\overline{\delta}$ such that $G^e(\overline{\delta})<0$, because $\gamma>1$. The contradiction implies that $G^e(\delta)\geq 0$, for each $\delta\in[\underline{\theta},\overline{\theta}]$. Hence, $\eta^e(\delta)\geq 0$, for each $\delta\in[\underline{\theta},\overline{\theta}]$.

Consider now $\eta^{ne}(\widehat{\delta})$. Iterated applications of de L'Hôpital's rule show that $\lim_{\widehat{\delta}\to\underline{\theta}}\eta^e(\widehat{\delta})=\frac{\gamma-1}{2}>0$ and $\eta^{ne}(\widehat{\delta})$ is clearly bounded. Evidently, $\frac{E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\gamma-1})}{E_{\Omega_I^{e}(\widehat{\delta})}(\delta_i^{\gamma-1})}<1$. Hence, given (α,Γ), for $\left(\frac{A^{ne}}{A^e}\right)$ sufficiently small, $\frac{\partial f(.)}{\partial\widehat{\delta}}>0$.

Alternatively, fix $\frac{A^{ne}}{A^e}\leq 1$. For any sequence $\Gamma^v\to 0$, the associated sequence $\left(\frac{A^{ne}}{A^e}\right)^{\frac{1+\Gamma^v}{\alpha\Gamma^v}}\left(\frac{E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\gamma^v-1})}{E_{\Omega_I^{e}(\widehat{\delta})}(\delta_i^{\gamma^v-1})}\right)^{\frac{(1-\alpha)(1+\Gamma^v)}{\alpha\Gamma^v}}$ also converges to zero, while (γ^v-1) converges to $\frac{\alpha}{1-\alpha}>0$. For each Γ, $\lim_{\widehat{\delta}\to\underline{\theta}}\eta^{ne}(\widehat{\delta})=\frac{\gamma-1}{2}$, which converges to $\frac{\alpha}{2(1-\alpha)}$ for $\Gamma\to 0$. Also, for each $\widehat{\delta}>\underline{\theta}$, $\lim_{\Gamma\to 0}\eta^{ne}(\widehat{\delta})$ is uniformly bounded above. It follows that, given $\frac{A^{ne}}{A^e}\leq 1$ and $\alpha\in(0,1)$, $\frac{\partial f(.)}{\partial\widehat{\delta}}>0$ at each $\widehat{\delta}\in\left[\underline{\theta},\overline{\theta}\right]$, for Γ sufficiently small.

EXAMPLE A1. Consider an economy with $[\underline{\theta}, \overline{\theta}] = [1, 2]$, $\alpha = \frac{2}{3}$, $\Gamma = 3$. Fix $\tau^e = \tau^{ne}$, and choose (A^e, A^{ne}) such that $\left(\frac{A^{ne}}{A^e}\right)^2 = 0.984$. Then,

$$\left(\frac{1}{A^e}\right)^2 \frac{f(\widehat{\delta})}{\chi(\xi)} = \widehat{\delta}^{\frac{1}{5}} \left(\frac{5}{6} \frac{2^{\frac{6}{5}} - \widehat{\delta}^{\frac{6}{5}}}{2 - \widehat{\delta}}\right)^{\frac{2}{3}} - 0.984 \times \widehat{\delta}^{\frac{1}{5}} \left(\frac{5}{6} \frac{\widehat{\delta}^{\frac{6}{5}} - 1}{\widehat{\delta} - 1}\right)^{\frac{2}{3}}.$$

By direct computation, $\lim_{\widehat{\delta} \to 1} \frac{f(\widehat{\delta})}{\chi(\xi)} = 0.06940$, $\frac{f(1.3)}{\chi(\xi)} = 0.06953$ and $\lim_{\widehat{\delta} \to 2} \frac{f(\widehat{\delta})}{\chi(\xi)} = 0.06924$. Therefore, $\frac{\partial f(\widehat{\delta})}{\partial \widehat{\delta}} > 0$ at some $\widehat{\delta} \in (1, 1.3)$ and $\frac{\partial f(\widehat{\delta})}{\partial \widehat{\delta}} < 0$ at some $\widehat{\delta}' \in (1.3, 2)$.

Also, set $\left(\frac{A^{ne}}{A^e}\right) = 1$. By numerical computation, for Γ large, say $\Gamma = 3$, $\frac{\partial f(\widehat{\delta})}{\partial \widehat{\delta}} < 0$. For $\Gamma \in (2, 2.1)$, it is (inverted) U-shaped, for $\Gamma < 1.9$, $\frac{\partial f(\widehat{\delta})}{\partial \widehat{\delta}} > 0$.

4A. Efficiency properties of equilibria

4.1A. Constrained optimal allocations

The planner's objective function is

$$\begin{aligned} P\left(h_i^s, k_j^s, \Omega_I^s, \Omega_J^s\right) \equiv & \sum_s \int_{\Omega_I^s(\widehat{\delta})} \left[\beta E_{\Omega_J^s(\widehat{\delta})} (A^s h_i^{s\alpha} k_j^{s(1-\alpha)}) - \frac{1}{\delta_i} \frac{h_i^{s(1+\Gamma)}}{1+\Gamma} - c_I^s\right] di \\ & + \sum_s \int_{\Omega_J^s(\widehat{\delta})} \left[(1-\beta) E_{\Omega_I^s(\widehat{\delta})} (A^s h_i^{s\alpha} k_j^{s(1-\alpha)}) - p k_j^s\right] dj. \end{aligned}$$

Given that the optimal choice k_j^s is $j-$invariant and that $\mu(\Omega_I^s(\widehat{\delta})) = \mu(\Omega_J^s(\widehat{\delta}))$, it can be rewritten as

$$P(h_i^s, k^s, \widehat{\delta}) = \sum_s \int_{\Omega_I^s(\widehat{\delta})} \left(A^s h_i^{s\alpha} k^{s(1-\alpha)} - \frac{1}{\delta_i} \frac{h_i^{s(1+\Gamma)}}{1+\Gamma}\right) di - \sum_s (c_I^s + p k^s) \mu(\Omega_I^s(\widehat{\delta})).$$

The planner's optimization problem is

$$\max_{(h_i^s, k^s, \widehat{\delta})} P(h_i^s, k^s, \widehat{\delta}).$$

It is convenient to decompose it into three problems. First, given an arbitrary value $\widehat{\delta}$, we determine the maps $(H^{COs}(\delta_i, \widehat{\delta}), K^{COs}(\widehat{\delta}))$ solving, for each s, the optimization problem

$$\begin{aligned} \max_{(h_i^s, k^s)} P_{\widehat{\delta}}^s(h_i^s, k^s) \equiv & \int_{\Omega_I^s(\widehat{\delta})} \left[A^s h_i^{s\alpha} k^{s(1-\alpha)} - \frac{1}{\delta_i} \frac{h_i^{s(1+\Gamma)}}{1+\Gamma}\right] di \\ & - (c_I^s + p k^s) \mu(\Omega_I^s(\widehat{\delta})). \end{aligned} \qquad (P_{\delta^*}^s)$$

Next, given the value functions $P^s(\widehat{\delta})$ of the two problems $(P_{\widehat{\delta}}^s)$, $s = ne, e$, we recast problem (P) as

$$\max_{\widehat{\delta}} \overline{P}(\widehat{\delta}) \equiv P^e(\widehat{\delta}) + P^{ne}(\widehat{\delta}), \qquad (\overline{P})$$

finding the optimal value of $\widehat{\delta}$, δ^{CO}.

Proof of Prop. 3. Given that optimization problem $(P_{\widehat{\delta}}^s)$ is concave, each s, its solution is completely characterized by the FOCs:

$i.\quad \frac{\partial P^s_{\widehat{\delta}}(h^s_i,k^s)}{\partial h_i} = \alpha A^s k^{s(1-\alpha)} h_i^{s(\alpha-1)} - \frac{1}{\delta_i} h_i^{s\Gamma} = 0,$

$ii.\quad \frac{\partial P^s_{\widehat{\delta}}(h^s_i,k^s)}{\partial k} = (1-\alpha) A^s k^{s(-\alpha)} \int_{\Omega^s_I(\widehat{\delta})} h_i^{s\alpha} di - p \int_{\Omega^s_I(\widehat{\delta})} di = 0,$

which imply

$a.\quad K^{COs}(\widehat{\delta}) = A^{s\frac{1+\Gamma}{\alpha\Gamma}} \alpha^{\frac{1}{\Gamma}} \left(\frac{1-\alpha}{p} E_{\Omega^s_I(\widehat{\delta})}(\delta_i^{\gamma-1})\right)^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}},$

$b.\quad H^{COs}(\delta_i,\widehat{\delta}) = \delta_i^{\frac{1}{1+\Gamma-\alpha}} \alpha^{\frac{1}{\Gamma}} A^{s\frac{1}{\alpha\Gamma}} \left(\frac{1-\alpha}{p} E_{\Omega^s_I(\widehat{\delta})}(\delta_i^{\gamma-1})\right)^{\frac{1-\alpha}{\alpha\Gamma}}.$

Comparing $a-b$ to $(A3)-(A4)$, $K^{COs}(\widehat{\delta}) > K^s(\widehat{\delta})$ and $H^{COs}(\delta_i,\widehat{\delta}) > H^s(\delta_i,\widehat{\delta})$, for each $\widehat{\delta}$, δ_i and s. Therefore, equilibria are always characterized by underinvestment in physical capital and in the effort in education. Demand and supply functions are clearly well-defined and continuous at each $\widehat{\delta} \in [\underline{\theta},\overline{\theta}]$. By substituting in the objective function the optimal values $(K^{COs}(\widehat{\delta}), H^{COs}(\delta_i,\widehat{\delta}))$, we obtain

$$b\overline{P}(\widehat{\delta}) \equiv \frac{\alpha\Gamma}{1+\Gamma} \sum_s \mu(\Omega^e_I(\widehat{\delta})) A^{s\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^s_I(\widehat{\delta})}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} - \mu(\Omega^e_I(\widehat{\delta})) bc^e_I,$$

where $\frac{1}{b} \equiv \alpha^{\frac{1}{\Gamma}} \left(\frac{1-\alpha}{p}\right)^{\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}}$. Given that $\overline{P}(\widehat{\delta})$ is a continuous function, problem $(\overline{P})$ has a solution, either internal or at one of the boundary points, and, therefore, CO allocations exist.

Compare a market allocation and any CO allocation. If $\delta^{CO} = \delta^F = \widehat{\delta}$, $K^{COs}(\widehat{\delta}) \neq K^s(\widehat{\delta})$ and the market allocation is not CO. Otherwise, $\delta^{CO} \neq \delta^F$ and constrained inefficiency follows immediately.

To establish the second part of Prop. 3, observe that, by direct computation and rearranging terms, the (necessary) FOC of problem $(\overline{P})$ can be written as

$$\begin{aligned} \gamma b \frac{\partial \overline{P}(\widehat{\delta})}{\partial \widehat{\delta}} &= -\widehat{\delta}^{\gamma-1} \left(A^{e\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^e_I(\widehat{\delta})}(\delta_i^{\gamma-1})^{\frac{(1+\Gamma)(1-\alpha)}{\alpha\Gamma}} - A^{ne\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^{ne}_I(\widehat{\delta})}(\delta_i^{\gamma-1})^{\frac{(1+\Gamma)(1-\alpha)}{\alpha\Gamma}} \right) \\ &\quad + \frac{(1-\alpha)(1+\Gamma)}{1+\Gamma-\alpha} \left(A^{e\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^e_I(\widehat{\delta})}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} - A^{ne\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^{ne}_I(\widehat{\delta})}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} \right) \\ &\quad + \gamma b c^e_I. \end{aligned} \qquad (A8)$$

Given $\xi = 0$, suppose $\delta^F > \underline{\theta}$. The condition defining the equilibrium value δ^F (i.e., the solution to $\widetilde{F}(\widehat{\delta},\widehat{\delta}) = 0$) can be recasted as

$$\begin{aligned} \frac{\gamma b c^e_I}{(\beta(1-\beta)^{\frac{1-\alpha}{\alpha}})^{\frac{1+\Gamma}{\Gamma}}} &= \left(A^{e\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^e_I(\delta^F)}(\delta_i^{\gamma-1})^{\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}} - A^{ne\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^{ne}_I(\delta^F)}(\delta_i^{\gamma-1})^{\frac{(1-\alpha)(1+\Gamma)}{\alpha\Gamma}} \right) \\ &\quad \times \delta^{F\gamma-1}. \end{aligned}$$

Hence, at δ^F,

$$\begin{aligned} \gamma b \frac{\partial \overline{P}(\widehat{\delta})}{\partial \widehat{\delta}}|_{\widehat{\delta}=\delta^F} &= -\gamma b c^e_I \left(\frac{1-(\beta(1-\beta)^{\frac{1-\alpha}{\alpha}})^{\frac{1+\Gamma}{\Gamma}}}{(\beta(1-\beta)^{\frac{1-\alpha}{\alpha}})^{\frac{1+\Gamma}{\Gamma}}} \right) + \frac{(1-\alpha)(1+\Gamma)}{1+\Gamma-\alpha} \\ &\quad \times \left(A^{e\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^e_I(\delta^F)}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} - A^{ne\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega^{ne}_I(\delta^F)}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} \right). \end{aligned}$$

The first term in brackets is positive. The second is positive for each value of $\widehat{\delta}$, and is bounded away from zero for each $\widehat{\delta} < \overline{\theta}$. This suggests that the sign of $\frac{\partial \overline{P}(\widehat{\delta})}{\partial \widehat{\delta}}|_{\widehat{\delta}=\delta^F}$ is indeterminate, as shown in the following example.

Let $\Gamma = p = A^e = 1$, $\alpha = \frac{2}{3}, \beta = \frac{1}{10}, [\underline{\theta},\overline{\theta}] = [1,4]$, $A^{ne} = 0.9$ and $\Gamma = 0.8$. Using the expressions above, one can compute the values of δ^F and δ^{CO} (in the example, $\overline{P}(\widehat{\delta})$ is a concave, monotonically increasing function, so that the FOC above are necessary and sufficient for an optimal interior solution). For c_I^e sufficiently small, $\delta^F < \delta^{CO}$. For higher values of c_I^e, the opposite occurs.

Proof of Corollary 1. Fix $\overline{\zeta}^e = \overline{\zeta}^{ne} = -\beta$ and $\overline{\tau}^e = \overline{\tau}^{ne} = -\frac{1-\beta}{\beta}$. Comparing $(A3 - A4)$ with $(a-b)$ in the last proof, given any threshold value $\widehat{\delta}$, the FOCs of the individual optimization problem in the actual economy imply that the FOCs of the (constrained) planner's optimization problem are satisfied, so that the optimal choices in the economy with taxes coincide with their constrained optimal values. To conclude, we need to find the value of Δc_I^e such that $\delta^F(\xi) = \delta^{CO}$.

In the economy with the optimal tax rates, using the definition of b introduced in the proof of Prop. 3, the condition defining the equilibrium value $\delta^F(\overline{\tau}, \overline{\zeta})$ is

$$\begin{aligned} & \widehat{\delta}^{\gamma-1}\left(A^e E_{\Omega_I^e(\widehat{\delta})}(\delta_i^{\gamma-1})^{(1-\alpha)}\right)^{\frac{1+\Gamma}{\alpha\Gamma}} - \widehat{\delta}^{\frac{\alpha}{1+\Gamma-\alpha}}\left(A^{ne} E_{\Omega_I^{ne}(\widehat{\delta})}(\delta_i^{\gamma-1})^{(1-\alpha)}\right)^{\frac{1+\Gamma}{\alpha\Gamma}} \\ = \; & \gamma b\left(c_I^e + \Delta c_I^e\right). \end{aligned} \tag{A9}$$

Set

$$\overline{\Delta c_I^e} = \frac{(1-\alpha)}{b\gamma^2}\left(A^{e\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega_I^e(\delta^{CO})}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}} - A^{ne\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega_I^{ne}(\delta^{CO})}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}}\right) > 0.$$

Then, at $\overline{\xi}$, $(A9)$ coincides with $(A8)$ above, so that $\delta^F(\overline{\xi}) = \delta^{CO}$. We are implicitly assuming that the solution is unique, which is necessarily true if $\frac{\partial f(.)}{\partial \widehat{\delta}} > 0$ at each $\widehat{\delta}$. Otherwise, $\delta^F(\overline{\xi}) = \delta^{CO}$ is one of the equilibrium thresholds associated with $\overline{\xi}$.

4.2A Welfare improving tax policies

Proof of Proposition 4. Let $Y^s(\delta^F(\xi);\xi)$ be the equilibrium level of the aggregate output in sector s,

$$\begin{aligned} Y^s(\delta^F(\xi);\xi) \; = \; & \mu(\Omega_I^s(\delta^F(\xi)))\,(\alpha\beta)^{\frac{1}{\Gamma}}\left(\frac{(1-\alpha)(1-\beta)}{p}\right)^{\frac{(1+\Gamma)(1-\alpha)}{\alpha\Gamma}} \\ & \times (1-\tau^s)^{\frac{1}{\Gamma}}\, A^{s\frac{1+\Gamma}{\alpha\Gamma}} E_{\Omega_I^s(\delta^F(\xi))}(\delta_i^{\gamma-1})^{\frac{1+\Gamma-\alpha}{\alpha\Gamma}}. \end{aligned}$$

The aggregate surplus is

$$S(\delta^F(\xi);\xi) = \sum_s Y^s(\delta^F(\xi);\xi)\left(\alpha + \beta - \alpha\beta\frac{2+\Gamma-\tau^s}{1+\Gamma}\right) - \mu(\Omega_I^s(\delta^F(\xi)))c_I^e.$$

Evidently,

$$\frac{\partial Y^e(.)}{\partial \delta^F(\xi)} = \frac{1+\Gamma}{\alpha\Gamma}\frac{Y^e(\delta^F(\xi);\xi)}{\mu(\Omega_I^s(\delta^F(\xi)))}\left(1-\alpha-\frac{\delta^F(\xi)^{\gamma-1}(\overline{\theta}-\delta^F(\xi))}{\overline{\theta}^\gamma - \delta^F(\xi)^\gamma}\right)$$

and

$$\frac{\partial Y^{ne}(.)}{\partial \delta^F(\xi)} = \frac{1+\Gamma}{\alpha\Gamma}\frac{Y^{ne}(\delta^F(\xi);\xi)}{\mu(\Omega_I^s(\delta^F(\xi)))}\left(\frac{\delta^F(\xi)^{\gamma-1}(\delta^F(\xi)-\underline{\theta})}{\delta^F(\xi)^\gamma - \underline{\theta}} - (1-\alpha)\right).$$

Hence,

$$\frac{\frac{\partial S}{\partial \delta^F(\xi)}}{D(.)} \; = \; \frac{c_I^e}{D(.)} + \left((1-\alpha) - \frac{\delta^F(\xi)^{\gamma-1}(\overline{\theta}-\delta^F(\xi))}{\overline{\theta}^\gamma - \delta^F(\xi)^\gamma}\right)$$

$$+\left(\frac{A^{ne}}{A^{e}}\right)^{\frac{1+\Gamma}{\alpha\times\Gamma}}\left(\frac{E_{\Omega_I^{ne}(\delta^F(\xi))}(\delta_i^{\gamma-1})}{E_{\Omega_I^{e}(\delta^F(\xi))}(\delta_i^{\gamma-1})}\right)^{\frac{1+\Gamma-\alpha}{\alpha\times\Gamma}}\left(\frac{\delta^F(\xi)^{\gamma-1}(\delta^F(\xi)-\underline{\theta})}{\delta^F(\xi)^{\gamma}-\underline{\theta}^{\gamma}}-(1-\alpha)\right),$$

for $D(.)\equiv\frac{1+\Gamma}{\alpha\Gamma}(\alpha+\beta-\alpha\beta\frac{2+\Gamma-\tau}{1+\Gamma})\frac{Y^e(.)}{\mu(\Omega_I^s(\delta^F(\xi)))}$.

We claim that, for $[\underline{\theta},\overline{\theta}]$ large enough and $\delta^F(\xi)$ sufficiently small, $\frac{\partial S}{\partial\delta^F(\xi)}>0$.

The first term is bounded away from zero, for each finite $\overline{\theta}$. The second converges to $(1-\alpha)$ for any sequence $\{\delta^v\}_{v=1}^{v=\infty}$, $\delta^v\to\underline{\theta}$ and any divergent sequence $\left\{\overline{\theta}^v\right\}_{v=1}^{v=\infty}$. For any sequence $\{\delta^v\}_{v=1}^{v=\infty}$, $\delta^v\to\underline{\theta}$, the last term in brackets has limit $\frac{\alpha\Gamma}{1+\Gamma}>0$. Hence, for $\overline{\theta}$ large and δ close enough to $\underline{\theta}$, $\frac{\partial S}{\partial\delta^F(\xi)}>0$.

Evidently, $\frac{\partial S(.)}{\partial\Delta c_I^e}=\frac{\partial S(.)}{\partial\delta^F(\xi)}\frac{\partial\delta^F(\xi)}{\partial\Delta c_I^e}$, with $\frac{\partial\delta^F(\xi)}{\partial\Delta c_I^e}>0$, hence $\frac{\partial S(.)}{\partial\Delta c_I^e}>0$. Also,

$$\frac{\partial S(.)}{\partial\tau^s}=\frac{\partial S(.)}{\partial\tau^s}|_{\delta^F(\xi)}+\frac{\partial S(.)}{\partial\delta^F(\xi)}\frac{\partial\delta^F(\xi)}{\partial\tau^s}$$

where

$$\frac{\partial S(.)}{\partial\tau^s}|_{\delta^F(\xi)}=-\frac{\alpha+\beta-\alpha\beta(2-\tau^s)}{(1-\tau^s)\Gamma}Y^s(.)<0.$$

Given that $\frac{\partial\delta^F(\xi)}{\partial\tau^{ne}}<0$, $\frac{\partial S(.)}{\partial\tau^{ne}}<0$. Finally, given that $\frac{\partial\delta^F(\xi)}{\partial\tau^e}>0$, the sign of $\frac{\partial S(.)}{\partial\tau^e}$ is indeterminate.

EXAMPLE A3: Welfare-improving effect of an increase in the highest marginal tax rate τ^e.

Fix $\Omega_I=[1,6]$, $\alpha=\frac{2}{3}$, $\beta=\frac{1}{10}$, $\Gamma=0.5$ and $\frac{A^{ne}}{A^e}=0.9$. Also, set $\tau^e=0.3$, $\tau^{ne}=0$. By numerical computation, one can check that $\frac{\partial\delta^F(\xi)}{\partial\tau^e}>0$, on $[1,6]$, which implies $\frac{\partial f}{\partial\delta^F(\xi)}>0$. Also, $\frac{\partial S(.)}{\partial\delta^F(\xi)}>0$, for each $\delta^F(\xi)$. Moreover, $\frac{\partial S(.)}{\partial\tau^e}<0$ for $\delta^F(\xi)\in[1,3.5]$, $\frac{\partial S(.)}{\partial\tau^e}>0$ for $\delta^F(\xi)\in[4,6]$. Hence, if the share of highly educated workers is below 40%, an increase in the tax progressiveness is welfare improving.

Proof of Prop. 5. Define tax revenues as

$$R(.)=\tau^e\beta Y^e+\tau^{ne}\beta Y^{ne}+\mu(\Omega_I^e(\delta^F(\xi)))\Delta c_I^e,$$

with

$$\begin{aligned}\frac{\partial R}{\partial\tau^s}&=\beta Y^s(.)+\tau^s\beta\frac{\partial Y^s}{\partial\tau^s}|_{\delta^F(\xi)}+\beta\left[\sum_s\tau^s\frac{\partial Y^s}{\partial\delta^F(\xi)}\right]\frac{\partial\delta^F(\xi)}{\partial\tau^s}\\&=\beta\frac{\Gamma-(1+\Gamma)\tau^s}{(1-\tau^s)\Gamma}Y^s(.)+\beta\left[\sum_s\tau^s\frac{\partial Y^s}{\partial\delta^F(\xi)}\right]\frac{\partial\delta^F(\xi)}{\partial\tau^s},\end{aligned}$$

and

$$\frac{\partial R}{\partial\Delta c_I^e}=\mu\left(\Omega_I^e(\delta^F(\xi))\right)-\Delta c_I^e+\left[\sum_s\tau^s\beta\frac{\partial Y^s}{\partial\delta^F(\xi)}\right]\frac{\partial\delta^F(\xi)}{\partial\Delta c_I^e}.$$

By the implicit function theorem,

$$\frac{\partial\tau^{ne}}{\partial\tau^{e}}|_{R=0} = -\frac{\frac{\Gamma-(1+\Gamma)\tau^{e}}{(1-\tau^{e})\Gamma}Y^{e}(.) + \left[\sum_{s}\tau^{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right]\frac{\partial\delta^{F}(\xi)}{\partial\tau^{e}}}{\frac{\Gamma-(1+\Gamma)\tau^{ne}}{(1-\tau^{ne})\Gamma}Y^{ne}(.) + \left[\sum_{s}\tau^{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right]\frac{\partial\delta^{F}(\xi)}{\partial\tau^{ne}}}.$$

Assume that τ^{s} satisfies $\frac{\Gamma}{1+\Gamma} > \tau^{s}$, so that $\frac{\partial R}{\partial\tau^{s}}|_{\delta^{F}(\xi)} > 0$. Consider a revenue neutral increase in τ^{e}, starting from a flat income tax, τ. By direct computation, given that $\frac{\partial\tau^{ne}}{\partial\tau^{e}}|_{R} < 0$,

$$\begin{aligned}\frac{\partial S(.)}{\partial\tau^{e}} &= -\frac{\alpha+\beta-\alpha\beta\,(2-\tau)}{(1-\tau)\,\Gamma}\left(Y^{e}\,(.) - Y^{ne}\,(.)\left|\frac{\partial\tau^{ne}}{\partial\tau^{e}}|_{R}\right|\right)\\ &\quad+\left(\alpha+\beta-\alpha\beta\frac{2+\Gamma-\tau}{1+\Gamma}\right)\left(\sum_{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right)\left(\sum_{s}\frac{\partial\delta^{F}(\xi)}{\partial\tau^{e}}\right).\end{aligned}$$

For τ sufficiently small, $\frac{\partial R}{\partial\tau^{e}} > 0$ and $\frac{\partial R}{\partial\tau^{ne}} > 0$. Hence,

$$\left|\frac{\partial\tau^{ne}}{\partial\tau^{e}}|_{R}\right| = \frac{\frac{\Gamma-(1+\Gamma)\tau^{e}}{(1-\tau^{e})\Gamma}Y^{e}(.) + \tau\left[\sum_{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right]\frac{\partial\delta^{F}(\xi)}{\partial\tau^{e}}}{\frac{\Gamma-(1+\Gamma)\tau^{ne}}{(1-\tau^{ne})\Gamma}Y^{ne} + \tau\left[\sum_{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right]\frac{\partial\delta^{F}(\xi)}{\partial\tau^{ne}}} > \frac{Y^{e}(.)}{Y^{ne}} > 0.$$

The inequality holds because $\frac{\partial\delta^{F}(\xi)}{\partial\tau^{e}} > 0$, $\frac{\partial\delta^{F}(\xi)}{\partial\tau^{ne}} < 0$, while, under the maintained assumptions, $\sum_{s}\frac{\partial Y^{e}}{\partial\delta^{F}(\xi)} > 0$. Hence,

$$\left(-Y^{e}\,(.) + Y^{ne}\,(.)\left|\frac{\partial\tau^{ne}}{\partial\tau^{e}}|_{R}\right|\right) > \left(-Y^{e}\,(.) + Y^{ne}\,(.)\,\frac{Y^{e}(.)}{Y^{ne}}\right) = 0.$$

Consequently,

$$\begin{aligned}\frac{\partial S(.)}{\partial d\tau^{e}} &> \left(\alpha+\beta-\alpha\beta\frac{2+\Gamma-\tau}{1+\Gamma}\right)\left(\sum_{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right)\\ &\quad\times\left(\frac{\partial\delta^{F}(\xi)}{\partial\tau^{e}} + \frac{\partial\delta^{F}(\xi)}{\partial\tau^{ne}}\frac{d\tau^{ne}}{d\tau^{e}}|_{R}\right) > 0.\end{aligned}$$

This establishes the first result.

By the implicit function theorem,

$$\frac{\partial\Delta c_{I}^{e}}{\partial\tau^{s}}|_{R=0} = -\frac{\beta\frac{\Gamma-(1+\Gamma)\tau^{s}}{(1-\tau^{s})\Gamma}Y^{s}(.) + \beta\left(\sum_{s}\tau^{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right)\frac{\partial\delta^{F}(\xi)}{\partial\tau^{s}}}{(\bar{\theta}-\delta^{F}\,(\xi)) - \Delta c_{I}^{e} + \beta\left(\sum_{s}\tau^{s}\frac{\partial Y^{s}}{\partial\delta^{F}(\xi)}\right)\frac{\partial\delta^{F}(\xi)}{\partial\Delta c_{I}^{e}}},$$

and $\frac{\partial\Delta c_{I}^{e}}{\partial\tau^{s}}|_{R=0} < 0$ for τ sufficiently small and $\Delta c_{I}^{e} < 0$.

Consider now an increase in labor income taxes to finance fixed subsidies to education. Evidently,

$$\begin{aligned}\frac{\partial S}{\partial d\tau^{s}} &= -\left(\frac{\alpha+\beta-\alpha\beta\,(2-\tau^{s})}{(1-\tau^{s})\,\Gamma}\right)Y^{s}(.) + \left(\sum_{s}\left(\alpha+\beta-\alpha\beta\frac{2+\Gamma-\tau^{s}}{1+\Gamma}\right)\frac{\partial Y^{s}}{\partial\delta^{F}\,(\xi)}\right)\\ &\quad\times\left(\frac{\partial\delta^{F}\,(\xi)}{\partial\tau^{s}} + \frac{\partial\delta^{F}\,(\xi)}{\partial\Delta c_{I}^{e}}\frac{\partial\Delta c_{I}^{e}}{\partial\tau^{s}}|_{R=0}\right).\end{aligned}$$

An increase in τ^{ne} has an unambiguous negative impact on total surplus, because the last term in brackets is negative. This shows (*ii.*).

In the case of changes in τ^e, $\frac{\partial \delta^F(\xi)}{\partial \tau^e} > 0$, while $\frac{\partial \delta^F(\xi)}{\partial \Delta c_I^e} > 0$ and $\frac{\partial \Delta c_I^e}{\partial \tau^e}|_{R=0} < 0$, so that the sign of the last term in brackets (and, consequently, of $\frac{\partial S}{\partial d\tau^e}$) is indeterminate.

7. References

Acemoglu, D.T., 1996, A microfoundation for social increasing returns in human capital accumulation, Quarterly Journal of Economics, 111, 779-804.

Acemoglu, D., and J. Angrist, 2001, How large are the social returns to education? Evidence from compulsory schooling laws, in B. Bernanke and K. Rogoff (eds.), NBER Macroeconomic Annual 2000, Cambridge MA, MIT Press, 9-58.

Anderberg, D., 2009, Optimal Policy and the Risk-Properties of Human Capital Reconsidered, Journal of Public Economics, 93, 1017-1026.

Anderberg, D., and F. Andersson, 2003, Investments in human capital, wage uncertainty, and public policy, Journal of Public Economics, 87, 1521-1537.

Ben-Porath, Y., 1970, The production of human capital over time, in W. Hansen (ed.), Education, Income and Human Capital, NBER, 129-154.

Betts, J.R., 1998, The Impact of Educational Standards on the Level and Distribution of Earnings, American Economic Review, 88, 266-275

Blankenau, W., 2005, Public schooling, college subsidies and growth, Journal of Economic Dynamics and Control, 29, 487–507.

Blankenau, W., and G. Camera, 2006, A simple economic theory of skill accumulation and schooling decisions, Review of Economic Dynamics 9, 93–115.

Blankenau, W., and G. Camera, 2009, Public spending on education and the incentives to student achievement, Economica, 76, 505–527.

Boskin, M.J., 1975, Notes on the tax treatment of human capital, NBER WP 116.

Bovenberg, A.L., and B. Jacobs, 2005, Redistribution and education subsidies are Siamese twins, Journal of Public Economics, 89, 2005-2035.

Carneiro, P., J.J. Heckman, and E. Vytlacil, 2001, Estimating the return to education when it varies among individuals, mimeo.

Caucutt, E.M., and K.B. Kumar, 2003, Higher education subsidies and heterogeneity. A dynamic analysis, Journal of Economic Dynamics and Control, 27, 1459-1502.

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

Da Costa, C.E., and Maestri, L.J., 2007, The Risk Properties of Human Capital and the Design of Government Policies, European Economic Review, 51, 695-713.

De la Fuente, A., 2003, Human Capital in a Global and Knowledge-Based Economy. Part II: Assessment at the EU Country Level. Final Report, European Commission, Directorate General for Employment and Social Affairs.

Eaton, J., and H.S. Rosen, 1980, Taxation, human capital, and uncertainty, American Economic Review, 70, 705-715.

Heckman, J.J., 1976, A life-cycle model of earnings, learning, and consumption, Journal of Political Economy, 84, S9-S44.

Heckman, J.J., A. Layne-Farrar, and P. Todd, 1996, Human capital pricing equations with an application to estimating the effect of schooling quality on earnings, Review of Economics and Statistics, 78, 562-610.

Jacobs, B., 2005, Optimal income taxation with endogenous human capital, Journal of Public Economic Theory, 7, 295-315.

Jacobs, B., 2007, Optimal redistributive tax and education policies in general equilibrium, CESIFO WP 2162.

Jacobs, B., and A.L. Bovenberg, 2008, Optimal taxation of human capital and the earnings function, CESIFO WP. 2250.

Jacobs, B., D. Schindler, and H. Yang, H., 2009, Optimal taxation of risky human capital, CESIFO W.P. NO. 2529.

Krueger, A.B., and M. Lindhal, 2000, Education for growth: Why and for whom?, Journal of Economics Literature, 39, 1111-1136.

Lloyd-Ellis, H., 2000, Public education, occupational choice, and the growth inequality relationship, International Economic Review, 41, 171-202.

Lucas, R.E., 1988, On the mechanics of economic development, Journal of Monetary Economics, 22, 3-42.

Mendolicchio, C., D. Paolini and T. Pietra, 2010, Investments in education in a random matching model, DSE Working paper 702, Università di Bologna.

OECD, 2008, Education at a Glance 2008. OECD Indicators, OECD, Paris.

Roy, A., 1951, Some thoughts on the distribution of earnings, Oxford Economic Papers, 3, 135-146.

Sahin, A., 2004, The incentive effects of higher education subsidies on student effort, Federal Reserve Bank of New York, Staff Report 192.

Sattinger, M., 1993, Assignment models of the distribution of earnings, Journal of Economic Literature, 31, 831-880.

Sattinger, M., 2003, A search version of the Roy model, mimeo.

Schindler, D., and B. Weigert, 2008, Insuring Educational Risk: Opportunities versus Income, CESifo WP 2348.

Schindler, D., and B. Weigert, 2009. Educational and Wage Risk: Social Insurance vs. Quality of Education, CESifo WP 2513.

Su, X., 2004, The allocation of public funds in a hierarchical educational system, Journal of Economic Dynamics and Control, 28, 2485-2510.

Recently published

No.	Author(s)	Title	Date
13/2010	Klinger, S. Rothe, T.	The impact of labour market reforms and economic performance on the matching of short-term and long-term unemployed	7/10
14/2010	Stegmaier, J.	Effects of workplace representation on firm-provided further training in Germany	8/10
15/2010	Dauth, W.	The mysteries of the trade: Employment effects of urban interindustry spillovers	8/10
16/2010	Stüber, H. Beissinger, T.	Does downward nominal wage rigidity dampen wage increases?	8/10
17/2010	Schneider, J. Beblo, M.	Health an work – Indicators and determinnats: A revised literature and data review for Germany	9/10
18/2010	Konle-Seidl, R.	Make further vocational training pay: Possibilities and limits of promoting transitions by Public Employment Services (PES)	10/10
19/2010	Heineck, G. Wölfel, O.	Parental risk attitudes and children's secondary school track choice	10/10
20/2010	Dürnberger, A. Drasch , K. Matthes, B.	Kontextgestützte Abfrage in Retrospektiverhebungen: Ein kognitiver Pretest zu Erinnerungsprozessen bei Weiterbildungsereignissen	10/10
21/2010	Hohmeyer, K. Wolff, J.	Direct job creation in Germany revisited: Is it effective for welfare recipients and does it matter whether participants receive a wage?	10/10
22/2010	Schlitte, Friso	Local human capital, segregation by skill, and skill-specific employment growth	11/10
1/2011	Grün, C. Mahringer, H. Rhein, Th.	Low-wage jobs: A means for employment integration oft he unemployed? Evidence from administrative data in Germany and Austria	1/11
2/2011	Achatz, J. Trappmann, M.	Arbeitsmarktvermittelte Abgänge aus der Grundsicherung: Der Einfluss von personen- und haushaltsgebundenen Arbeitsmarktbarrieren	2/11
3/2011	Patuelli, R. Schanne, N. Griffith, D.A. Nijkamp, P.	Persistence of regional unemployment: Application of a spatial filtering approach to local labour markets in Germany	2/11
4/2011	Riphahn, R.T. Trübswetter, P.	The intergenerational transmission of educational attainment in East and West Germany	2/11
5/2011	Koller, L.	Lohnmobilität alleinstehender SGB-II-Leistungsempfänger	2/11
6/2011	Brück-Kling berg, A. Burkert, C. Garloff, A. Seibert, H. Wapler, R.	Does higher education help immigrants find a job? A survival analysis	3/11

As per: 14.03.2011

For a full list, consult the IAB website
http://www.iab.de/de/publikationen/discussionpaper.aspx

Imprint

IAB-Discussion Paper 7/2011

Editorial address
Institute for Employment Research
of the Federal Employment Agency
Regensburger Str. 104
D-90478 Nuremberg

Editorial staff
Regina Stoll, Jutta Palm-Nowak

Technical completion
Jutta Sebald

All rights reserved
Reproduction and distribution in any form, also in parts,
requires the permission of IAB Nuremberg

Website
http://www.iab.de

Download of this Discussion Paper
http://doku.iab.de/discussionpapers/2011/dp0711.pdf

For further inquiries contact the author:

Concetta Mendolicchio
Phone +49.911.179 2330
E-mail concetta.mendolicchio@iab.de