Wage Premia, Education Race, and Supply of Educated Workers

Luca G. DEIDDA
Dimitri PAOLINI
We model a labor market in which workers’ level of education might be a signal of skills. We show that whenever the wage premium for education increases over time – as it might happen under skill biased technological progress – the investment in education needed to sustain a separating equilibrium in which skilled workers perfectly signal their type, also increases. Hence, an increase in the education wage premium induces an education race. If the borrowing capacity of poor workers is lower than that of rich ones due to capital market imperfections, poor-skilled workers will finally fall behind in this race – and pool together with some unskilled ones – as the investment they would have to undertake to signal their type eventually becomes unaffordable to them. Such mechanism supports a supply side explanation for the joint long run trends of (i) the education wage premia, and (ii) the relative supply, of postgraduates and college graduates in the US labor market, which complements the demand based explanation for wage skill premia based on skill bias technological change hypothesis.

Keywords: education race, skills, signaling, supply of educated workers, wage education premium.

JEL Classification: D4, D8, L15

1 Università di Sassari, CRENoS, Italy; and SOAS, University of London, U.K. E-mail: deidda@uniss.it
2 Università di Sassari, CRENoS, Italy; Université catholique de Louvain, CORE and IRES, B-1348 Louvain-la-Neuve, Belgium.

We thank Arpad Abraham, Fabrizio Adriani, Luisito Bertinelli, Sergio Currarini, Leo Ferraris, Axel Gautier, Piero Gottardi, Andreas Irmen, Francesco Lippi, Ramon Marimon, Massimo Morelli, Matthew Neidell, Pierre Picard, for their insights and suggestions at various stages of this project, and seminar participants at EUI (Florence), CREA (University of Luxembourg), DSE (Università di Firenze), ASSET2012, III Workshop CRENoS-DSE on Institutions, Individual Behavior and Economic Outcomes - Argentiera 2012, for useful comments. We also thank David Autor for making available on his webpage the data used in Autor and Acemoglu (2011), along with the relevant STATA codes. Part of this paper has been written while Dimitri Paolini author was visiting CREA (University of Luxembourg), and Luca Deidda was visiting the department of Economics of European University Institute, Firenze. Dimitri Paolini gratefully acknowledges the financial support from the Luxemburgish Fond National de la Recherche (FNR; AM2c measure). Both authors gratefully acknowledge financial support by the Italian Ministero dell’ Università (PRIN), by the Regione Autonoma della Sardegna (Legge n. 7), and Fondazione Banco di Sardegna.
1 Introduction

There is a large literature on the long-run dynamics of wage inequality across education groups and labor supply by education level in the US.\(^1\) Over the period 1963-2008, such inequality has widened: More educated workers have been generally gaining on less-educated ones, at all levels of education (see figure 1). Over the same time period, the percentage of postgraduates (PGs), college graduates (CGs) and workers with some college (SCs) in the overall workforce has increased, while the percentage of high school graduates (HSGs) and High school dropouts (HSDs) have been decreasing (see figure 2).

A well established explanation for the observed trends in wage premia for education is that skill bias technical change – possibly induced by greater availability of educated workers – has resulted in a stronger demand for skills and – therefore – a higher wage skill premium.\(^2\)

Focusing on higher levels of education, we observe that, over the time period considered (see figure 3):

i. The wage premium experienced by PGs has grown substantially more than that of CGs;\(^3\)

ii. The relative supply of PGs has increased less than that of CGs.

That is, in the long run, the relative supply of PGs has grown less than that of CGs even if the wage premium for education has been increasing in favor of PGs. This is in spite of the fact that – at least for the 1989-2008 period, for which we could find online official data from the US National Center for Education Statistics (NCES) – the average growth rate of the annual tuition and fees for PG degrees has been just 0.262% higher than that of CGs degrees, which is significantly lower than the average growth rate of wage differential in favor of PGs relative to CGs over the same period, which is around 10%.\(^4\)


\(^2\)Goldin and Katz, 1998, Autor, Katz and Krueger, 1998 provide a comprehensive analysis of the empirical evidence on long term skill bias technological change, and education wage premia. Autor and Acemoglu, 2011, review what they call the canonical model of labor supply and demand used within the extensive literature on returns to skills and wage inequality under skill biased technological change.

\(^3\)Barrow and Rouse (2005) calculate that the hourly wage gap between college and non college educated workers which had grown by 25%, grew only by 10% in the 1990s.

\(^4\)Calculations are based on data taken from the NCES Digest of Education Statistics, 2011, tables 349 and 352. See also
Figure 1: Composition-Adjusted, real, log weekly wages for full-time full-year workers 1963-2008.

Source: March CPS dataset used in Acemoglu and Autor, 2011, made available – together with the relevant STATA codes – by David Autor at http://economics.mit.edu/faculty/dautor/data/acemoglu. The real log weekly wage for each education group is computed following Autor and Acemoglu, 2011.

If the return to postgraduate education has grown more than the return to college education, why has the long-run relative supply of PGs increased less than that of CGs?

There is definite evidence that cognitive and noncognitive skills developed early in life – $S_0$ in the terminology of Heckman et al. (2006) – play a significant role in explaining success in schooling and earnings (wages) profile.\(^5\) For this reason, in trying and answer the above question, we abstract from the effects that education has on skill formation, and focus on a model à la Spence (1974) in which education plays a signaling role. We construct a baseline model to convey the main results, which we then explore within the context of a model that accounts for exogenous skill biased technological progress according to the standard approach based on a CES technology.

We consider a labor market populated by competitive firms and heterogeneous workers. Workers are

\(^5\)Cunha, Heckman, Lockner, and Masterov, 2006 provide an exhaustive discussion of the literature on the determinants of skills.
heterogeneous along two dimensions: Financial wealth and skills. While information about individual skills is private, worker’s education is observable. Marginal productivity of labor at firm level depends positively on workers’ skills. As firms hire workers without observing their individual skills, equilibrium wages equal hired workers’ expected marginal productivity, conditional on the available information on the distribution of skills across workers.

Workers use their time endowment to study and work. Studying yields a lower disutility to skilled workers than to unskilled ones. This leads to the possibility of a skill separating equilibrium (SSE) in which only skilled workers educate themselves, thereby earning the associated wage premium.

We assume that investing in education has a cost in terms of financial resources, which increases in the amount of education, and that workers’ have no access to borrowing. Hence, the SSE described

---

6 Alternative, one could assume that skilled workers benefit more from education than unskilled ones.

7 We solve the model under such extreme case of borrowing constraints, because given our stylized setup, it simplifies the exposition without loss of generality.
above can exist only if the amount of education that sustains it does not exceed what poor-skilled workers can afford to self-finance. Assume this is initially the case, and suppose that due to exogenous factors, such as exogenous skill biased technological progress, the wage premium between educated (skilled) and uneducated (unskilled) workers associated with the SSE goes up. Unskilled workers would then have a stronger motive to mimic skilled workers by investing in education, as the reward to such investment has increased. Accordingly, the level of investment in education by skilled workers necessary to sustain the SSE must increase. That is, as the wage premium for education increases, skilled workers engage in an education race.

Crucially, if the wage premium for education associated with the above SSE grows large enough, the investment in education necessary to sustain such equilibrium becomes greater than the maximum amount that poor-skilled workers can self-finance. At this stage, since workers cannot borrow to finance education, the SSE cannot exist any longer. We show that – under these circumstances – the robust
equilibrium is a skill pooling equilibrium (SPE) in which only rich-skilled workers are able to perfectly signal their high skills by choosing a level of education higher than before, while poor-skilled are pooled with some of the unskilled ones (possibly all of them) at some lower level of education. That is, poor-skilled workers finally fall behind in the education race.

The distribution of education levels across workers who are heterogeneous in wealth and skills changes with the equilibrium. Rich-skilled workers educate themselves more in any SPE than they do in the SSE. Therefore, the relative supply of workers who choose a level of education equal to the highest level of education associated with any SPE, increases from zero to a positive value as we move from the SSE to a SPE. Moreover, in the SSE only skilled workers invest in education. Differently, in an SPE we could have that both skilled and unskilled invest in education, as poor-skilled pool might with some of the unskilled ones at some intermediate, positive, level of education. Hence, as we move from the SSE to an SPE, the highest level of education achieved by workers increases, and we could also observe more levels of education. Indeed, casual observation suggests that in the old days there were very few PGs (hardly none), relative to CGs, and, in general, less variety in the education composition of the workforce, compared to what we observe today.

As we move from an SSE to an SPE, the fraction of workers who acquire the highest level of education, which is now higher than before, could be lower than the increase in relative supply of workers who get lower – but positive – education. This, in spite of the fact that the wage premium for the highest level of education has increased more than that for lower levels of education. Hence, the model rationalizes the evidence according to which the relative supply of PGs has increased less than that of CGs in spite of the fact that the wage premium has grown in favor of PGs. Accordingly, we see our model as complementary to well-established demand-driven explanations of the observed trends in wage premia for education and supply of educated workers based upon skill-biased technological progress. Actually, according to the model, whether enough poor and skill workers fall behind in the race so that the relative supply of PGs grows less than that of CGs it depends on the distribution of skills and wealth across workers. The lower
is the fraction of skilled workers, the lower is the fraction of poor workers in the workforce population required for the result to hold.\textsuperscript{8}

Goldin and Katz, 2008 (ch. 8), show that a theoretical framework based upon the race between workers’ education (which they use to measure the supply of skills) and skill-biased technological change (demand for skills) is capable of explaining the long-run trends in wage premia for education and education attainments of the workforce for the US economy.\textsuperscript{9} Our race setup differs from theirs, in that we model education choices explicitly, in the context of a game where education has a signaling role. As a result, in our case, the average level of workers’ skills is, conditional on education, endogenous, and evolves as a byproduct of the race.\textsuperscript{10} Consider our model and suppose the economy is initially in a SSE where all skilled workers are CGs, and assume this is the maximum level of education that poor-skilled workers can afford. In the presence of skill-biased technological progress, skilled workers will have to study more and more in order to signal their skills as the wage premium for education goes up. This finally leads to an SPE in which, rich-skilled workers become PGs, while poor-skilled workers pool with some of the unskilled ones at some lower level of education, possibly staying CGs, as before. If so, the quality of CGs will have gone down along with the transition from SSE to SPE due to the education race triggered by exogenous technical progress. This result is consistent with the direct evidence on the decline in the quality of college graduates provided by Carneiro and Lee, 2011 (see section V of their paper).\textsuperscript{11}

The mechanism we propose relies on the idea that – in the presence of borrowing constraints – poor

\textsuperscript{8}There is some concern that family background in the US has deteriorated over time (see Heckman, 2006), which hinders the formation of early skills. If early skills become more scarce, then, according to our model, wealth inequality matters more for the effectiveness of education as a signal of skills, and, therefore, for the evolution of education levels of the workforce population.

\textsuperscript{9}Acemoglu and Autor, 2012, provide a comprehensive critical assessment of such approach.

\textsuperscript{10}Mendolicchio, Paolini and Pietra, 2012, analyze a setup where the average level of skills of workers with a given level of education is endogenously determined due to composition effects.

\textsuperscript{11}Related to that, Lindsay and Machin, 2011, show that – in the US and Great Britain, where relative salaries of PGs have grown more than those of CGs, PGs and CGs appear to be imperfect substitutes in production, and there have been "[...] trend increases over time in the relative demand for postgraduate vis-à-vis college only workers [...], which are " [...] significantly correlated with technical change as measured by changes in industry computer usage and investment [...]", and, [...] " the skills sets possessed by postgraduates and the occupations in which they are employed are significantly different to those of college only graduates [...]", (taken from their paper’s abstract).
workers have a lower ability (propensity) to invest in education, irrespectively of their skills. In the empirical literature, one can find various arguments that could offer support to the idea that credit constraints might be important.\textsuperscript{12} There is, for instance, a widespread evidence on the relationship between family income and college attendance.\textsuperscript{13} As Carneiro and Heckman, 2002, suggest, one could interpret such correlation as the result of (1) short run credit constraints faced by teenagers, or; (2) Some sort of long-run credit constraint, namely ”[...] the inability of the child to buy the parental environment and genes that form the cognitive and noncognitive abilities required for success in school [...]” (page, 706). Although the two interpretations are not mutually exclusive, Carneiro and Heckman argue that – according to the US data they analyze – interpretation (2) is far the more important, as parental income plays a minor role once they control for other long term family factors.\textsuperscript{14} Following this argument, in our model, one should interpret individual wealth as long term parents’ wealth, and borrowing constraints as long term ones. As a consequence, since early skills are shaped by long term family factors, we should then assume positive correlation between wealth and skills.

However, Brown, Scholz, and Seshadri, 2012 – show that identifying borrowing constraints based on the relationship between family income and children’s schooling might be misleading in a model in which parents choose whether to invest in children’s education, and whether to make cash transfers to them. Accounting for this identification problem, they show that – based on data from the Health and Retirement Studies – the set of children who face quantitatively important short-term borrowing constraints for higher education is substantial.\textsuperscript{15} Accordingly, in our model, we do not interpret individual wealth as long term parents’ wealth, and therefore, since the skills we focus on are shaped by long term family traits, assuming independence between wealth and skills, seems natural.

Aside from borrowing constraints, one could think of many other individual characteristics – including

\textsuperscript{12}See Brown, Scholz, and Seshadri, 2012, for a discussion.
\textsuperscript{13}See Carneiro and Heckman, 2002, page 706, figure 1.
\textsuperscript{14}As they put it ”[...] short run income constraints play a role, albeit a quantitatively minor one. […]” (page, 707). According to their estimates, up to 8% of the US population is short run credit constrained.
\textsuperscript{15}In a similar vein, using a quantitative financial cycle model with college enrollment whereby parental transfers are endogenously determined calibrated to the US data, Winter, 2013 finds that 24% of US households were constrained in their college decision in the 80s.
some related to contingent aspects of the individual’s family – that could shape individuals’ willingness (or ability/capacity) to engage in education, independently of early skills. One could focus on any of these factors, and construct a model based on the idea that workers are – conditional on skills – heterogeneous with respect to that characteristic, which could yield predictions similar to the ones we derive here. To this extent, we believe that the economic intuition underlying our main results is more general than the specific setup we adopt.

The paper is organized as follows. In sections 2 and 3 we present the baseline model. Section 4 extends the results to the case of exogenous skill biased technological change. Section 5 concludes the paper.

2 The baseline model

We consider a competitive economy populated by a continuum of size one of workers and a continuum of size one of firms. All agents are price-taker and risk-neutral. Each worker is endowed with an amount of wealth $\omega \in \Omega$, where $\Omega = \{\omega, \overline{\omega}\}$, with $0 \leq \omega < \overline{\omega}$, an amount $T$ of time, and a level of skills $\theta \in \Theta$, where $\Theta = \{\theta, \overline{\theta}\}$, with $\overline{\theta} > \theta$. Workers’ individual type is defined by a pair $\{\theta, \omega\}$, with $\Theta \times \Omega$. Type is assigned to each worker by nature, where, $\pi, (1 - \pi)$ is the probability of a worker being assigned skills $\theta, (\theta)$, and $\delta, (1 - \delta)$, is the probability of a worker being assigned wealth, $\omega, (\omega)$.

Workers spend their time studying and working. Studying for an education level (or degree) of length $n$ costs $C(n) = cn$, which is to be paid upfront, where $c$ is the fee rate per unit of time. By studying for a degree of length $n$, a worker also suffers a disutility $n/\theta$, measured in pecuniary terms.

Firms are homogeneous. A firm employing $l$ units of labor from a worker of skills $\theta$ produces output $y$ according to $y = \phi(\theta)l$, where $\phi(\theta)$ is the marginal product of a unit of labor supplied by a worker of skills, $\theta$. We assume that $\phi(\theta)$ is strictly increasing in $\theta$.

Both firms and workers take wages as given. Workers’ individual skills are unobservable. The distribution of skills and wealth across workers is public information, and workers’ individual levels of education are observable.

Finally, we assume there exists no market for financial resources, so that workers cannot borrow in order to finance the cost of education.

**Timing.** The sequence of events is the following:

**Stage 0.** Nature decides workers’ type \( \{\theta, \omega\} \);

**Stage 1.** Having observed their individual type, workers simultaneously allocate their time to studying and working;

**Stage 2.** Having observed workers’ educational choices, firms make their employment decisions;

**Stage 3.** Wages are set so to clear the labor market. Exchange and production, if any, take place, and payoffs are realized.

**Firms’ beliefs, and expected marginal productivity.** At stage zero, firms’ prior beliefs assign probability \( \pi, (1 - \pi) \), that the level of skills of an individual worker is \( \overline{\theta}, (\theta) \). Upon observing a worker with a degree of education, \( n \), firms’ posterior beliefs assign a probability, \( \mu(n|\theta) \in [0, 1] \) that the worker has skills, \( \theta \). Accordingly, the expected level of productivity of a worker with education, \( n \), is

\[
E(\phi(\theta)|n) = \mu(\theta|n)\phi(\theta) + [1 - \mu(\theta|n)] \phi(\overline{\theta}), \tag{1}
\]

**Workers’ payoff.** A worker of type \( \{\theta, \omega\} \), who studies for a degree of length \( n \leq T \), is able to offer an amount \( T - n \) of labor, incurring a cost of education \( cn \) and a skill-dependent disutility \( n/\theta \). Let \( w \) the wage paid to workers with a level \( n \) of education. Then, worker’s individual payoff is

\[
v(\theta, \omega, w, n) = (T - n)w + \omega - n \left( c + \frac{1}{\theta} \right). \tag{2}
\]
Since workers cannot borrow to finance education, a worker can engage in a degree of length $n$ if and only if her wealth, $\omega$, weakly exceeds $C(n)$. Accordingly,

$$n^{\text{max}}(\omega) : C(n^{\text{max}}(\omega)) = \omega \Rightarrow n^{\text{max}}(\omega) = \frac{\omega}{c},$$

(3)

defines the maximum amount education that can be self-financed by a worker endowed with an amount $\omega$ of wealth. Clearly, rich workers can afford higher levels of education than poor ones, $n^{\text{max}}(\omega_R) > n^{\text{max}}(\omega_P)$. We assume that,

$$\frac{\omega_R}{c} > T > \frac{\omega_P}{c},$$

(4)

so that while rich workers can self-finance any feasible degree of education, the same is not true for poor workers.

**Single crossing.** Let $\{n_1, w_1\}$ and $\{n_2, w_2\}$ two education-wage pairs, with $n_2 > n_1$, $w_2 > w_1$. Then, the net payoff from choosing $\{n_2, w_2\}$ instead of $\{n_1, w_1\}$ is the same across workers who differ in wealth and are homogeneous in skills, and it is greater for skilled workers than for unskilled ones, irrespectively of whether they are rich or poor. Given these properties it is immediate to verify that the following result holds.

**Lemma 1** (Sorting condition). Let $\{n_1, w_1\}$ and $\{n_2, w_2\}$ two education-wage pairs, with $n_2 > n_1$, and $w_2 > w_1 \geq 0$. If $v(\theta, \omega, w_2, n_2) \geq v(\theta, \omega, w_1, n_1)$, then $v(\theta, \omega, w_2, n_2) > v(\theta, \omega, w_1, n_1)$ follows.

That is, skilled workers benefit from education more than unskilled ones.

**3 Equilibrium**

An equilibrium is a set of strategies for workers and firms; a wage schedule $w(n)$ that maps possible levels of education, $n$, into wages, i.e. $w(n) : n \rightarrow w$, and; a belief function $\mu(\theta | n)$, such that:

i. Firms and workers’ strategies are optimal based upon the available information;

ii. Beliefs are derived from the strategy profiles using Bayes’ rule whenever possible;
iii. The wage schedule, \( w(n) \), is consistent with agents’ optimal strategies and beliefs, and clears the market for labor.

This standard signaling game yields many possible equilibria, broadly classifiable into two categories:

1. Skill separating equilibria (SSE), in which all skilled workers, independently of their wealth \( \omega \), separate from unskilled ones;

2. Skill pooling equilibria (SPE), in which (some) skilled workers pool with (some) of the unskilled ones.

### 3.1 Robust equilibria: Existence and characterization

We restrict our attention to symmetric equilibria robust to the intuitive criterion.\(^{17}\) First, we analyze robustness and existence of SSE, and then focus on SPE.

**Proposition 1** (Robust SSE: Existence and characterization). A robust SSE exists if and only if the marginal productivity of skilled workers, \( \phi(\theta) \), is not too large, i.e.,

\[
\phi(\theta) \leq \frac{c(T\phi(\theta) + \omega) + \omega / \theta}{Tc - \omega} \equiv \phi^{\text{max}}. \tag{5}
\]

The robust SSE is unique. Skilled workers choose a level of education,

\[
n(\theta) = \frac{T[\phi(\theta) - \phi(\theta)]}{\phi(\theta) + c + \omega}, \tag{6}
\]

and earn a wage, \( w(n(\theta)) = \phi(\theta) \), while unskilled ones choose \( n(\theta) = 0 \), and earn a wage \( w(n(\theta)) = \phi(\theta) \).

Relative supply of workers by level of education is as follows: \( l(n(\theta)) = \pi \), and \( l(n(\theta)) = (1 - \pi) \).

**Proof.** See appendix.

If all skilled workers are able to finance the cost of the degree of length \( n(\theta) \), which allows them to signal their high skills and get the wage, \( w(n(\theta)) \), then it must be the case that \( n(\theta) \) does not exceed the maximum amount of education, \( \omega / c \), that poor workers – irrespectively of their skills – can afford. As it can be seen from equation (6), \( n(\theta) \) is – other things equal – strictly increasing in \( \phi(\theta) \), in such a way that there exist a (unique) critical value for the marginal productivity of skilled workers, \( \phi(\theta) \), which we

\(^{17}\)A model-specific definition of the intuitive criterion can be found in the appendix.
call $\phi^{\max}$, such if and only if $\phi(\theta) > \phi^{\max}$, $n(\theta)$, exceeds the maximum amount of education that poor workers can self-finance; in which case the SSE cannot exist.

From a related perspective, according to Proposition 1 a SSE exists only if there is not too much difference between the marginal productivity of skilled and unskilled workers, respectively, so that the wage paid to educated (and skilled) workers is not too large than that paid to uneducated (and unskilled) ones.\(^{18}\)

Proposition 1 states that a robust SSE does not exists if and only if the marginal productivity of skilled workers, $\phi(\theta)$, is not too high. What happens if that condition is not satisfied? Define,

$$
\phi(\theta_p) = \frac{\pi(1-\delta)}{(1-\pi)(1-p)+\pi(1-\delta)}\phi(\theta) + \frac{(1-\pi)(1-p)}{(1-\pi)(1-p)+\pi(1-\delta)}\phi(\theta),
$$

the expected marginal productivity of a worker coming from a pool of workers including all the poor-skilled ones and a fraction $1-p$ of the unskilled ones, with $p \in [0, 1-\pi]$, where

$$
\theta_p = \frac{\pi(1-\delta)}{(1-\pi)(1-p)+\pi(1-\delta)}\theta + \frac{(1-\pi)(1-p)}{(1-\pi)(1-p)+\pi(1-\delta)}\theta,
$$

is the expected level of skills of workers in the pool.

**Proposition 2** (Robust SPE: Existence and characterization). A robust SPE exists if and only if the marginal productivity of skilled workers, $\phi(\theta)$, is large enough, i.e.,

$$
\phi(\theta) > \phi^{\max}.
$$

Any robust SPE is characterized as follows:

i. Rich-skilled workers always separate by choosing a level of education

$$
n(\theta, \omega) = \frac{T \left[ \phi(\theta) - \phi(\theta_p) \right] + n(p) \left[ \phi(\theta_p) + c + \theta^{-1} \right]}{\phi(\theta) + c + \theta^{-1}},
$$

and get a wage, $w(n(\theta, \omega)) = \phi(\theta)$;

ii. If the expected marginal productivity of the pool of workers containing all workers but the rich-skilled ones, $\phi(\theta_0)$, is strictly lower than $\phi^{max}$, a continuum of equilibria exists each associated with a different value of $p \in [0, p^{\text{max}}]$, of the fraction of unskilled workers who separate, where

$$
p^{\text{max}} : \phi(\theta_{p^{\text{max}}}) = \phi^{\max},
$$

\(^{18}\)As usual, the robust SSE is the Riley outcome, that is the equilibrium associated with the lowest investment in education that allows skilled workers to separate.
such that: i. a fraction $p$ of unskilled workers separate by playing $n(\theta) = 0$, and receive a wage $w(n(\theta)) = \phi(\theta)$; A fraction $1-p$ of the unskilled ones pool together with poor-skilled workers by playing $n(\theta_p) = \overline{n}(\theta_p)$, where

$$\overline{n}(\theta_p) = \frac{\phi(\theta_p) - \phi(\theta)}{\phi(\theta_p) + c + \theta^{-1}},$$

and receive a wage $w(n(\theta_p)) = \phi(\theta_p)$.

iii. If $\phi(\theta_0) \geq \phi_{\text{max}}$, all workers but rich-skilled ones pool together, i.e. $p = 0$, by playing $n(\theta_0) \in [0, \overline{n}(\theta_0)]$, and receive a wage $w(n(\theta_0)) = \phi(\theta_0)$.

iv. Relative supply of workers by education level is as follows: $l(n(\overline{\theta})) = \delta \pi$, $l(n(\theta_p)) = \pi (1 - \delta) + (1 - \pi)(1 - p)$; $l(n(\theta)) = (1 - \pi)p$.

Proof. See appendix.

The intuition is the same one discussed above for SSE. Consider a candidate SPE, in which poor-skilled workers are pooled with some of the unskilled ones, while rich-skilled workers, who are rich enough to afford the cost of any feasible level of education, separate by choosing $n(\overline{\theta}, \omega)$. If and only if the wage paid to rich-skilled workers, $\phi(\overline{\theta})$ is below $\phi_{\text{max}}$, could poor-skilled workers afford the cost associated with the level of education, $n(\overline{\theta}, \omega)$, which is necessary to separate from the unskilled they are pooled with. Viceversa, if $\phi(\overline{\theta}) > \phi_{\text{max}}$, they cannot afford it, and they are forced to pool with some of the unskilled workers. This also implies that, according to Proposition 2, SPEs are robust if only if there is enough dispersion between the marginal productivity of high skill workers, $\phi(\overline{\theta})$, and that of low skill workers,$\phi(\theta)$.

3.2 Levels of education, wage premia for education, and relative supply of workers by education level

We now analyze how the levels of education, the wage premia for education, and the relative supply of workers by level of education change with the marginal productivity of skilled workers, other things equal.

If $\phi(\overline{\theta}) \leq \phi_{\text{max}}$, the robust equilibrium outcome is unique and characterized by two levels of education. All skilled workers are able to perfectly signal their skills by studying for a degree of length, $n(\overline{\theta}) > 0$. These workers are paid a wage equal to $\phi(\overline{\theta})$. Unskilled workers get no education, $n(\theta) = 0$, and they
a wage, $\phi(\theta)$. The relative supply of educated workers equals $\pi$, while that of uneducated workers equals $1 - \pi$, and the wage premium for education is $\phi(\bar{\theta}) - \phi(\theta)$.

If $\phi(\bar{\theta}) > \phi^{\text{max}}$, in any robust equilibrium outcome, rich-skilled workers still perfectly signal their skills by studying for a degree of length, $n(\bar{\theta}, \bar{\omega}) > 0$, and receive a wage $\phi(\bar{\theta})$. However, they have now to study more: $n(\bar{\theta}, \bar{\omega}) > n(\bar{\theta})$. The relative supply of workers at the highest level of education, $n(\bar{\theta}, \bar{\omega})$, is $\delta \pi$.

Differently, poor-skilled workers fall behind as they cannot afford to finance the cost of a degree of length, $n(\bar{\theta}, \bar{\omega})$. They pool with some unskilled workers. As stated in proposition 2 – depending on parameter values – either (case 1) all workers but the rich-skilled ones are pooled together at some positive level of education, $n(\theta_0)$, or (case 2) poor-skilled workers pool with a fraction $1 - p < 1$ of unskilled ones at some strictly positive level of education, $n(\theta_p)$, while a fraction $p$ of the unskilled separates by staying uneducated, $n(\theta) = 0$. In both cases, the level of education achieved by the workers pooling together is strictly lower than that played by rich-skilled ones.

In case 1, any robust equilibrium outcome, is characterized by two levels of education, $n(\bar{\theta}, \bar{\omega})$, and $n(\theta_0) \geq 0$. The relative supply of workers playing $n(\theta_0)$ will be $1 - \delta \pi$. The wage paid to workers who choose the level of education $n(\theta_0)$ would be, $\phi(n(\theta_0))$, and the wage premium paid to workers who choose the highest level of education, $n(\bar{\theta}, \bar{\omega})$, relative to those who choose the lowest level of education, $n(\theta_0)$, is $\phi(\bar{\theta}) - \phi(n(\theta_0)) > 0$.

Viceversa, in case 2, three levels of education emerge. Rich-skilled play $n(\bar{\theta}, \bar{\omega})$, poor-skilled workers pool with a fraction $1 - p$ of the unskilled ones and play an intermediate level of education $n(\theta_p)$, while a fraction $p$ of the unskilled separate by play $n(\theta) = 0$. The relative supply of workers at the intermediary level of education, $n(\theta_p)$, will be equal to $[1 - p](1 - \pi) + \pi(1 - \delta)$, while the relative supply of workers at the lowest level of education will be equal to, $(1 - \pi)p$. The wage premium for the highest level of education measured relatively to the wage paid to uneducated workers is $\phi(\bar{\theta}) - \phi(\theta) > 0$. Similarly, the wage premium for workers with an intermediate level of education relative to those the uneducated
ones is $\phi(\theta_p) - \phi(\theta) > 0$. Note that the wage premium associated with the highest level of education is strictly greater than that for the intermediate level of education, given $p < 1$. In other words there is a wage premium, $\phi(\theta) - \phi(\theta_p) > 0$, for the highest level of education relative to the intermediate level of education. The possible cases are summarized in table 1.

Table 1: Relative supply of workers and wages by education level

<table>
<thead>
<tr>
<th>Relative labor supply (RLS) and wages (w) by education level</th>
</tr>
</thead>
<tbody>
<tr>
<td>High $(n(\theta), n(\bar{\theta}, \overline{\theta}))$</td>
</tr>
<tr>
<td>RLS</td>
</tr>
<tr>
<td>SSE $(\phi(\overline{\theta}) \leq \phi_{\text{max}})$</td>
</tr>
<tr>
<td>SPE</td>
</tr>
<tr>
<td>$(\phi(\theta) &gt; \phi_{\text{max}})$</td>
</tr>
</tbody>
</table>

3.3 Trends in wage education premia and education of the workforce in the US

Can these predictions help explaining the joint trends of wage premia and relative labor supply for PGs vs CGs in the US?

Consider an the economy that is initially in a SSE equilibrium in which only skilled workers and all of them, get a college degree (CG). In this equilibrium, the relative supply of PGs equals zero; the relative supply of CGs equals $\pi$; and, the relative supply of workers with a level of education lower than CG is $1 - \pi$. The wage premium for CGs equals $\phi(\theta) - \phi(\theta)$. The wage premium for a worker who were to get a postgraduate degree (never mind the fact that it is never optimal to do so) would also be $\phi(\theta) - \phi(\theta)$. In other words, PG and CG degrees buy the same wage premium over lower levels of education.

Suppose now that $\phi(\overline{\theta})$ increases above $\phi_{\text{max}}$, and the economy switches to a SPE. In the new equilibrium, rich-skilled workers are getting more education than in the previous one: They get some sort of postgraduate degree (PG). Relative supply of PGs is now positive, and equal to $\delta \pi$. Poor-skilled workers cannot afford to finance the cost associated with that level of education, and fall behind. Suppose they
We measure the difference between the relative supply of PGs and CGs, $\Delta \text{RLS}_{PGs-CGs}$, as a function of $\phi(\theta)$ on the vertical axis to the left, and the difference between the wage of PGs, $w_{PGs}$, and that of CGs, $w_{CGs}$, on the vertical axis to the right. The economy is in a SSE for values of $\phi(\theta) \leq \phi^{max}$, and in a SPE for values of $\phi(\theta)$ such that the reverse inequality holds.

Then, we know they must be pooled with some of the unskilled workers. As we know, there are different possible situations. Suppose that all the unskilled workers also study for a college degree. In this case, the relative supply of CGs is $1 - \delta \pi$. Then, so long as

$$\delta < \frac{1 - \pi}{2\pi},$$

(13)

the relative supply of PGs increases less than the relative supply of CGs. This, in spite of the fact that PGs receive a wage premium relative to CGs. In fact, (13), is the necessary condition for the relative supply of PGs to increase less than that of CGs as we move from a SSE to a SPE.

Note that, according to the above condition, as the fraction $\pi$ of skilled workers decreases, the lower is the fraction, $1 - \delta$, of poor workers in the workforce population, required for the result to hold. As skills become a more scarce resource, wealth distribution becomes more important in determining workers’ ability so signal their skills through education.

\footnote{There is no guarantee of this being the case, as in the new equilibrium, workers other than rich-skilled ones can be pooled at any level of education so long as such level is below the level played by the rich-skilled.}
Figure (4) summarizes the relevant trends. As the difference between the wage of PGs and that of CGs increases from $\phi_1 - \phi(\theta)$ to $\phi_2 - \phi(\theta)$, the relative labor supply of PGs drops relative that of CGs. This is coherent with the observed trends for the US economy, see figure 3.

4 Exogenous skill bias technical progress (ESBTP)

In this section, we explore a variant of the above model, in which firms’ production is based upon a technology that combines low skill and high skill labor, according to the following CES production function:

$$Y = \left[ \sum_{\theta \in \Theta} [\psi(\theta)L(\theta)]^{\frac{1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}},$$

where $L(\theta)$ is the amount of labor assigned to tasks that require a minim level of skills, $\theta$, and $\psi(\theta)$ is the related factor-augmenting technology term, which is assumed to be strictly increasing in $\theta$. We assume that while only skilled workers can perform high skill tasks, skilled workers are as productive as unskilled workers when undertaking unskilled tasks. Furthermore, in the following discussion, we focus on the case in which the elasticity of substitution between skilled and unskilled workers is greater than one, i.e. $\sigma > 1$. Aside from these new assumptions, the model is unchanged.

4.1 Firms’ beliefs, and equilibrium wages

The same equilibrium concept applies as in the baseline model. Workers are paid according to the following wage schedule,

$$w(n) = \mu(\theta|n) \bar{w} + (1 - \mu(\theta|n)) \underline{w},$$

where

$$\bar{w} \equiv L(\theta)^{-\frac{1}{\sigma}} \psi(\theta) \frac{\sigma}{\sigma-1} \left[ \sum_{\theta \in \Theta} [\psi(\theta)L(\theta)]^{\frac{1}{\sigma}} \right]^{\frac{\sigma-1}{\sigma}},$$

$$\underline{w} \equiv L(\theta)^{-\frac{1}{\sigma}} \psi(\theta) \frac{\sigma}{\sigma-1} \left[ \sum_{\theta \in \Theta} [\psi(\theta)L(\theta)]^{\frac{1}{\sigma}} \right]^{\frac{\sigma-1}{\sigma}},$$
are the equilibrium levels of wages that would be paid to workers assigned to high skill and low skill
tasks respectively, under full information for given values, \( L(\theta), L(\theta) \), of the supply of labor employed in
skilled and unskilled tasks, respectively.

It is important to note that, under perfect information, unskilled workers would be always allocated
to low skill tasks, in any equilibrium. Differently, all skilled workers will be all allocated to high skill
tasks, if and only if \( \psi(\theta) \geq \psi_{\text{min}} \), where

\[
\psi_{\text{min}} \equiv \left( \frac{\pi}{1 - \pi} \right)^{\frac{1}{\sigma - 1}} \psi(\theta).
\]  

(18)

If the above inequality is not satisfied, then, in any candidate equilibrium in which all skilled workers were
to be assigned to high skill tasks, these workers would be paid a lower wage than that paid to low skill
workers assigned to low skill tasks; which implies that such equilibrium would not exist as skilled workers
would prefer to deviate and undertake low skill tasks. In fact, if \( \psi(\theta) < \psi_{\text{min}} \), the unique equilibrium
under perfect information is such that a fraction,

\[
f = \frac{\left( \frac{\psi(\theta)}{\psi(\theta)} \right)^{\frac{\sigma}{\pi - 1}}}{\pi \left( 1 + \left( \frac{\psi(\theta)}{\psi(\theta)} \right)^{\frac{\sigma}{\pi - 1}} \right)},
\]  

(19)
of skilled workers, where \( f < 1 \) holds, is assigned to high skill tasks so that, \( l(\theta) = \pi f \) is the relative
supply of workers assigned to high skill tasks and \( l(\theta) = (1 - \pi) + (1 - f)\pi \) is the supply of workers
assigned to low risk tasks. In such equilibrium, \( \bar{w} = \underline{w} \), holds, where equilibrium wages are given by
equations (16), and (17), once we substitute in for the equilibrium values of labor supply for low and
high skill tasks, which equal \( L(\theta) = N\pi f \), and \( L(\theta) = [(1 - p) + (1 - f)\pi]N \).\(^{20}\)

4.2 Robust equilibria

As in the baseline model, skilled workers benefit more from education than unskilled ones. Yet, poor-
skilled workers are at a disadvantage compared to rich ones, since they cannot borrow to finance education.

As a consequence, SSE or SPE emerge as the robust equilibria depending on parameter values.

\(^{20}\)Note that under perfect information, workers do not engage in education.
Let us define the critical level \( \psi_{\text{max}} \) of the productivity of skilled workers, \( \phi(\theta) \), such that:

\[
\left( \frac{T - n_{\text{max}}(\omega)}{T} \right) \left( \frac{L(\theta)}{L(\theta)} \right)^{\frac{1}{2}} \left( \frac{\psi_{\text{max}}}{\psi(\theta)} \right)^{\frac{1}{2}} = \frac{n_{\text{max}}(\omega)(c + \theta^{-1})}{TL(\theta) - \frac{1}{2} \psi(\theta) \sigma^{2}} \left[ \frac{\psi(\theta)L(\theta)}{\sigma^{2}} + \frac{\psi_{\text{max}}L(\theta)}{\sigma^{2}} \right]^{\frac{1}{2}} + 1,
\]

(20)

where, \( L(\theta) = L_{\text{min}} \), with \( L_{\text{min}} = \pi[T - n_{\text{max}}(\omega)] \), and \( L(\theta) = (1 - \pi)T \).

As explained in the proof of the proposition below, the above value of \( \psi(\theta) \) is associated with a SSE in which in order to signal their skills, skilled workers have to play a value of education equal to the maximum that poor workers can afford, \( n_{\text{max}}(\omega) \). Then,

**Proposition 3** (SSE with ESBTP: Existence and Characterization). A robust SSE exists if and only if the level of the factor-augmenting technology term for skilled labor, \( \psi(\theta) \) is neither too small or too large, i.e. \( \psi(\theta) \in [\psi_{\text{min}}, \psi_{\text{max}}] \). The robust SSE is unique it is characterized as follows:

i. Unskilled and skilled workers choose \( n(\theta) = 0 \), and \( n(\theta) > 0 \) respectively, where

\[
n(\theta) : \left[ \frac{T - n(\theta)}{T_{\text{w}}} \right] \frac{w}{w} = \frac{n(\theta)}{T_{w}}(c + \theta^{-1}) + 1;
\]

(21)

ii. Supply of labor by level of skills is \( L(\theta) = \pi[T - n(\theta)] \), \( L(\theta) = (1 - \pi)T \);

iii. Wages are \( w(n(\theta)) = w \), \( w(n(\theta)) = w \);

iv. Relative supply of workers by education level is: \( l(n(\theta)) = \pi \), and \( l(n(\theta)) = 1 - \pi \).

**Proof.** See appendix.

The intuition is the same as in the baseline model. Suppose that the factor-augmenting term for high-skill tasks, \( \psi(\theta) \) increases due to exogenous skill-biased technical progress. Then, the wage earned in a SSE by high skill workers who signal their level of skills by investing in education increases relatively to that of those who do not invest in education, thereby signaling a low level of skills. Therefore, mimicking high skill workers by investing in education becomes more attractive. Accordingly, the equilibrium level of investment in education that sustains a SSE, \( n(\theta) \), increases. Eventually, for \( \psi(\theta) \) greater than \( \psi_{\text{max}} \), such level of investment becomes greater than the maximum level of education that poor-skilled workers can afford, \( n_{\text{max}}(\omega) \), so that no SEE exist.

\[\text{21} \text{The values of } w \text{ and } w \text{ are found by substituting for the equilibrium values of } L(\theta), \text{ and } L(\theta), \text{ in equations (16) and (17).}\]
When no SSE exists, any equilibrium will be characterized by the fact that some skilled workers are pooled together with some of the unskilled ones. The following result holds,

**Proposition 4** (SPE with ESBTP: Characterization and existence). A robust SPE exists if and only if the level of the factor-augmenting technology term for skilled labor, $\psi(\bar{\theta})$ satisfies, $\psi(\bar{\theta}) > \psi_{\text{max}}$. SPE are characterized as follows:

1. If the factor-augmenting technology term for skilled labor, $\psi(\bar{\theta})$ exceeds the critical value $\psi_{\text{min}}$,
   i. All rich-skilled workers always separate by playing $n(\bar{\theta}, \bar{\omega})$, where
      $n(\bar{\theta}, \bar{\omega}) : \frac{(T - n(\bar{\theta}))\bar{w}}{T \bar{w}} = \frac{n(\bar{\theta})}{T \bar{w}}(c + \bar{\theta}^{-1}) + 1$; \hspace{1cm} (22)
   
   ii. A fraction $p \geq 0$ of unskilled workers separate, playing $n(\bar{\theta}) = 0$.
   iii. All poor-skilled workers pool with a fraction $1 - p$ of the unskilled ones, by playing $n(\theta_p)$ where,
      $n(\theta_p) : [T - n(\theta_p)]w(n(p)) - (c + \theta_p^{-1}) n(\theta_p) \geq \bar{w}T$; \hspace{1cm} (23)
      with strict equality if $p > 0$.
   iv. Wages are, $w(n(\bar{\theta}, \bar{w})) = \bar{w}$, $w(n(\bar{\theta})) = \bar{w}$, and $w(n(\theta_p)) = \mu(\bar{\theta}|n(\theta_p))\bar{w} + (1 - \mu(\bar{\theta}|n(\theta_p)))\bar{w}$, where
      \[ \mu(\bar{\theta}|n(\theta_p)) = \frac{(1 - \delta)\pi}{(1 - \delta)\pi + (1 - \pi)p}; \] \hspace{1cm} (24)

2. If $\psi(\bar{\theta}) > \psi_{\text{min}}$, then no worker acquires education, all unskilled workers plus a fraction
   \[ f = \frac{\left(\frac{\psi(\bar{\theta})}{\psi(\bar{\theta})}\right)^{\frac{\pi}{\bar{\pi}}} \pi}{\pi + 1 + \left(\frac{\psi(\bar{\theta})}{\psi(\bar{\theta})}\right)^{\frac{\pi}{\bar{\pi}}} \pi} \] \hspace{1cm} (25)
   perform low skill tasks, and a fraction $1 - f$ of skilled workers perform high skill tasks.

**Proof.** See appendix.

Equivalently to the baseline model, as the factor-augmenting term associated with skilled labor increases due to skill-biased technological progress, the amount of education that skilled workers have to invest in to perfectly signal their skills becomes unaffordable to poor-skilled workers who fall behind in the education race and are pooled together with some unskilled at some lower, intermediate level of education.
4.3 Wage premium for education, relative supply of workers by education level, and the education race

The equilibrium is not unique. Yet, all robust equilibrium outcomes share some qualitative properties, which enable us to draw some predictions about the effect of skill biased technological progress on the supply of labor by education level, and the wage premia for education. As it emerges from the following discussion, the results of the baseline model still hold, at qualitative level.

For intermediate values of $\psi(\theta)$, the equilibrium is characterized by two levels of education, $n(\theta) > 0$, and $n(\theta) = 0$. All skilled workers invest in educated, while unskilled stay uneducated, and there is a strictly positive wage premium for education, $w(n(\theta)) > w(n(\theta))$. So long as $\psi(\theta)$ does not exceed the critical value $\psi^{\text{max}}$, further skill-biased technological progress results in a higher wage premium for education, while supply of educated and uneducated workers stay constant, and equal to $\pi$ and $1 - \pi$, respectively.

For values of $\psi(\theta)$ above $\psi^{\text{max}}$, a further (intermediate) level of education, $n(\theta_p)$, emerges. Rich-skilled workers still separate from the by educating themselves more than other workers. They choose an amount of education, $n(\theta, \omega)$, that exceeds any of the possible levels they would choose for values of $\psi(\theta)$ below $\psi^{\text{max}}$. Poor-skilled workers pool together with some fraction $p$ of the unskilled ones at the intermediate level of education, $n(\theta_p)$, while the rest of the unskilled stay uneducated.

The supply of workers at the top level of education, is now $\delta \pi$, while the supply at the intermediate or lower levels of education is $1 - \delta \pi$. There is a strictly positive wage premium for education at all levels of education: $w(n(\theta, \omega)) > w(n(\theta))$, and $w(n(\theta_p)) > w(n(\theta))$, hold. Given any equilibrium, the wage premium is increasing in the level of education: rich-skilled workers, who are the most educated ones, are paid more than workers who get an intermediary level of education. An increase in $\psi(\theta)$ due to further skill-biased technological progress results in a higher wage premium at all levels of education, and in a higher level of education for the rich-skilled workers.
Wage premia for education and relative supply of educated workers. The result is confirmed that, in spite of the fact that skill biased technological progress results in higher education premia for higher levels of education, as poor-skilled workers fall behind in the education race, the relative supply of workers at the highest level of education might well increase less than the relative supply of workers at intermediate levels of education. In particular, condition (13) is still necessary and sufficient for the relative supply of workers at the highest level of education (PG, for example) to go down relative to that of workers with intermediate levels of education (CG, for instance) as the level of the factor-augmenting term for skilled labor grows large enough.

SBETP and the education race. If the condition $\psi(\theta) < \psi^{\min}$ there is no positive level of education such that the resulting SSE equilibrium guarantees skilled workers a wage high enough to compensate for the cost of education. The only equilibrium is one in which both skilled and unskilled workers get the same wage, and make no investment in education.

Hence, in a context in which education has no intrinsic content but only a signaling value, and marginal returns to labor are decreasing, education emerges as a by product of skill biased technological progress. Moreover, as skill biased technological progress results in higher wage premia for education, it also induces an education race. To the extent that – independently of their skills – not all workers have the same access to education (which in this model is due to capital market imperfections and heterogeneity in initial wealth), such race is associated with a certain degree of sluggishness of the supply of highly educated workers, compared to what would happen if workers where homogeneous with respect to their opportunities to access to education.

5 Conclusion

There is a vast literature documenting the rise of wage inequality across educational groups of American workers and the increase in the supply of skills in the US labor market over the last forty years. Many
economists have proposed a demand driven explanation for such phenomenon based upon the idea of skill bias technical progress.

We propose a complementary explanation based on a model of the labor market where workers are heterogeneous with respect to wealth and skills – which are both unobservable – and costly investment in education might have a role in signaling the level of skills. We show that as the wage premium for education increases, the minimum investment in education needed to sustain a perfectly separating equilibrium (SSE) in which skilled workers are able to perfectly signal their skills increases. Hence, an increase in the wage premium for education generates an education race with skilled workers investing more and more to signal themselves. However, if capital markets are imperfect so that the borrowing capacity of poor workers is lower than that of those who are rich, this race will finally lead to a situation in which – for a sufficiently large increase in the endowment of the accumulable factor – poor-skilled workers are no longer able to invest enough to signal themselves and fall behind, ending up pooled together with some (possibly all) of the unskilled. Hence, at the highest level of education (PGs, for instance) the supply of skills is sluggish with respect to an increase in the correspondent wage premium. The model offers a supply side explanation for the observed trends of relative supply and wage skill differentials of PGs and CGs, as well as for the falling quality of CGs, in the US. In the present model, such education race only affect workers’ welfare and has no consequences on aggregate productivity of workers. We leave the analysis of these potential consequences, which could be analyzed in the context of a model with endogenous skill biased technological progress, for future research.
References


A Appendix

Definition 1 (Intuitive criterion). Consider a candidate equilibrium $E$. Let $\Theta_1$ and $\Theta_2$ be two subsets of the set of individual skills, $\Theta$, such that $\Theta_1 \cup \Theta_2 = \Theta$ and $\Theta_1 \cap \Theta_2 = \emptyset$. Let $\Theta_1$ be the subset of skills such that workers with skills $\theta \in \Theta_1$ are worse off from a deviation, $n$, no matter what the beliefs of the firms observing $n$ are. Let $\Theta_2$ be the subset of skills such that workers with skills $\theta \in \Theta_2$ always strictly benefit from the same deviation, $n$, provided that firms assign probability zero to the event that an individual of skills $\theta \in \Theta_1$ has deviated. In other words, any individual with skills $\theta \in \Theta_2$ strictly benefits from the deviation, $n$, for any system of firms’ beliefs that assign probability zero to the event that an individual of skills $\theta \in \Theta_1$ has deviated. Then, if the subset $\Theta_2$ is non-empty, $E$ is not robust.

A.1 Proof of proposition 1

Let $N_\theta$ denote the set of values of $n$ played with positive probability by a worker of skills $\theta$, given equilibrium, so that a SSE can be defined as an equilibrium in which $N_\theta \cap N_\theta' = \emptyset$, holds. Then,

Lemma 2. In any SSE, $N_\theta$ and $N_\theta'$ must be singletons.

Proof. Consider a candidate SSE equilibrium in which, $n', n'' \in N_\theta$, with $n'' > n'$, for a given level of skills, $\theta$. Since, $N_\theta \cap N_\theta' = \emptyset$ for $\theta \neq \theta'$ holds (given that the candidate equilibrium is SSE), upon observing $n'$, or $n''$ firms should conclude that the worker playing $n'$, or $n''$ is of type $\theta$. Hence, $\mu(\theta|n') = \mu(\theta|n'')$ must hold, which implies $E(\phi(\theta)|n') = E(\phi(\theta)|n'')$. Market clearing requires $w(n) = E(\phi(\theta)|n)$, so that $w(n') = w(n'')$ must hold. But then, since the disutility of education is strictly increasing in $n$, workers of skills $\theta$ strictly prefer playing $n'$ to $n''$, which destroys the candidate equilibrium. Therefore, in a SSE, the sets $N_\theta$ and $N_\theta'$ must be singletons. □

Characterization of robust equilibria. According to the above lemma, in any SSE, workers of skills $\theta$ play the same value of $n$, $n(\theta)$. Skilled workers play $n(\overline{\theta})$, and unskilled workers play $n(\overline{\theta})$. $\mu(\theta|n(\theta)) = 1$ holds, so that $w(n(\overline{\theta})) = \phi(\overline{\theta})$, and $w(n(\overline{\theta})) = \phi(\overline{\theta})$. As for the equilibrium values of education, consider first candidate SSE equilibrium where unskilled workers choose $n(\overline{\theta}) > 0$. It is immediate to verify that these workers will be strictly better off – for any possible off equilibrium beliefs – by playing $n' < n(\overline{\theta})$, which destroys the equilibrium. Hence, in any SSE, $n(\overline{\theta}) = 0$ must hold. As for the level of $n(\overline{\theta})$, the following applies. In general, any SSE must satisfy the following set of participation, and incentive
compatibility constraints (PC and ICC, respectively):

\[
PC(\theta, \omega): [T - n(\theta)]w(n(\theta)) + \omega - n(\theta)(c + \frac{1}{\theta}) \geq 0, \quad (A.1)
\]

\[
PC(\theta, \omega): [T - n(\theta)]w(n(\theta)) + \omega - n(\theta)(c + \frac{1}{\theta}) \geq 0, \quad (A.2)
\]

\[
ICC(\theta, \omega): [T - n(\theta)]w(n(\theta)) - n(\theta)(c + \frac{1}{\theta}) \geq 0, \quad (A.3)
\]

\[
ICC(\theta, \omega): [T - n(\theta)]w(n(\theta)) - n(\theta)(c + \frac{1}{\theta}) \leq 0, \quad (A.4)
\]

It is immediate to verify that PCs are always satisfied when ICCs are. Imposing strict equality on ICC(\theta, \omega) and n(\theta) = 0, and substituting for the equilibrium values of wages, we find

\[n = \frac{[\phi(\theta) - \phi(\theta)]T}{\phi(\theta) + c + \theta^{-1}}, \quad (A.5)\]

as the maximum value of n(\theta) such that skilled workers have the incentive to separate. Similarly, imposing strict equality on ICC(\theta, \omega), we find

\[n = \frac{[\phi(\theta) - \phi(\theta)]T}{\phi(\theta) + c + \theta^{-1}}, \quad (A.6)\]

as the minimum value of n(\theta) such that unskilled workers have no incentive to mimic skilled ones.

Consider now a candidate SSE in which, n(\theta) > n. There exist sufficiently pessimistic off equilibrium beliefs to sustain such equilibrium. However, the equilibrium is not robust to the intuitive criterion. Consider a deviation n' \in [n, n(\theta)).\textsuperscript{22} Unskilled workers will be worse off from such deviation independently of the value of off equilibrium beliefs, \mu(\theta|n'). Differently, skilled workers will be strictly better off provided that firms assign a probability zero that the deviation comes from an unskilled worker, that is provided that \mu(\theta|n') = 1. Therefore, SSE with n(\theta) > n are not robust. Consider now a candidate SSE, such that n(\theta) = n. Consider a deviation n' such that n' < n. Both unskilled and skilled workers strictly benefit from such deviation if \mu(\theta|n'). Hence, such candidate SSE is robust to the intuitive criterion. Therefore, we conclude that any robust SSE is characterized by n(\theta) = n.

Finally, since all skilled workers play n(\theta) = n, while uneducated workers play n(\theta) = 0, relative supply of educated and uneducated workers equal \pi and 1 − \pi, respectively.

**Existence.** In order for the SSE characterized above to exist, given n(\theta) = 0, n(\theta) = n, (i) ICCs and PCs must be satisfied; (ii) \textsuperscript{22} Only downward deviations are to be considered, as upward deviation make everyone worse off for any possible beliefs.
\[ \phi(\bar{\theta}) \leq \frac{c(T\phi(\theta) + \omega) + \omega/\theta}{Tc - \omega} \equiv \phi_{\text{max}}. \] (A.7)

\[ \square \]

A.2 Proof of proposition 2

Characterization of robust equilibria. Let \( \mathcal{N}_{\theta_p} \) is the set of values of education played by workers heterogeneous in skills in a given SPE. The following result applies

Lemma 3. A SPE is non-robust if rich-skilled workers are pooling with other types.

Proof. By contradiction, assume that rich-skilled workers are playing some \( n(\theta_p) \in \mathcal{N}_{\theta_p} \) receiving a wage \( w(n(\theta_p)) < \phi(\bar{\theta}) \). Let us define,

\[ \pi(\theta) : (T - n(\theta_p)) w(n(\theta_p)) - (c + 1/\theta)n(\theta_p) + \omega = (T - \pi(\theta)) \phi(\bar{\theta}) - (c + 1/\theta)\pi(\theta) + \omega, \] (A.8)

the maximum value of education that a worker of skills \( \theta \) playing \( n(\theta_p) \) would deviate to if perceived as a skilled worker by doing so. Solving the above equality for \( \pi(\theta) \), we have

\[ \pi(\theta) = \frac{T \left[ \phi(\bar{\theta}) - w(n(\theta_p)) \right] + n(\theta_p) [w(n(\theta_p)) + c + \theta^{-1}]}{\phi(\bar{\theta}) + c + \theta^{-1}}. \] (A.9)

It is immediate to verify that – other things equal – \( \pi(\theta) \) increases with \( \theta \). Hence, there always exist deviations \( n \in (\pi(\theta), \pi(\bar{\theta})) \) that, for off equilibrium beliefs such that \( \mu(\theta|n) = 1 \), are strictly beneficial to rich-skilled workers, only. Accordingly, any SPE where rich-skilled are pooled with other types is never robust. \( \square \)

Therefore, in any candidate robust SPE, rich-skilled workers always separate. Furthermore, the same argument of lemma 2 applies so that these workers all play the same level of \( n \), which we call \( n(\theta_p, \bar{\omega}) \). As for the other types, the following preliminary result holds.

Lemma 4 (Levels of education in a robust SPE). Any robust SPE is characterized by at most three levels of education, \( n(\bar{\theta}), n(\theta), \) and \( n(\theta_p) \), where \( n(\bar{\theta}) > n_{\text{max}}(\omega) \) is played by rich-skilled workers, \( n(\theta) = 0 \) is played by a fraction \( p \) unskilled ones, and \( n(\theta_p) \in (0, n_{\text{max}}(\omega)) \) is played by a pool including a fraction \( 1 - \delta \) of skilled workers and a fraction \( 1 - p \) of unskilled ones.

Proof. Consider a candidate equilibrium in which a fraction of unskilled workers separate playing \( n(\theta) > 0 \). The standard argument holds that these unskilled workers are strictly better off by deviating to \( n(\theta) \) for any possible value of firms’ off equilibrium beliefs, \( \mu(\theta|n) \), which destroys the equilibrium. Hence, in any equilibrium, where a fraction of unskilled agents separate, \( n(\theta) = 0 \) must hold.

30
Since rich-skilled workers always separate, in any SPE where a fraction of unskilled workers also separate, there must be at least one intermediate level of education, which is strictly positive and lower than the level of education played by rich-skilled workers.

Consider now a candidate equilibrium such that the set of average skills associated with the various levels of education played by workers in equilibrium, includes more than one intermediate level of skills. In particular, let, \( n', \) and \( n'' \) be two intermediate levels of education played with positive probability by pools of workers heterogeneous in individual skill levels, with \( n(\bar{\theta}) > n' > n'' > n(\bar{\theta}) = 0. \) Clearly, both unskilled and -skilled workers must be indifferent between playing \( n'' \) and \( n' \). This implies,

\[
[T - n']w(n') - n'(c + \bar{\theta}^{-1}) = [T - n'']w(n'') - n''(c + \bar{\theta}^{-1}),
\]

must hold for unskilled workers. But then, it follows directly from lemma 1 that,

\[
[T - n']w(n') - n'(c + \bar{\theta}^{-1}) > [T - n'']w(n'') - n''(c + \bar{\theta}^{-1}),
\]

which contradicts the hypothesis that in equilibrium, poor-skilled workers are playing both \( n' \) and \( n'' \) with positive probability. Clearly enough, the above argument applies –in general– to any candidate equilibrium where the number of intermediate skill levels associated with the set of levels of education played with positive probability by workers is greater than one.

The above results imply that we have two possible kinds of SPE. Either, (i) All workers but rich-skilled, who separate playing some \( n(\bar{\theta}, \omega) \), pool together \((p = 0)\) by playing \( n(\theta_0) \), or; (ii) A fraction \( p > 0 \) of unskilled workers separate, by playing \( n(\bar{\theta}) \), the rest of the unskilled workers pool with poor-skilled ones, playing \( n(\theta_p) \), while rich-skilled workers separate playing some \( n(\bar{\theta}, \omega) \). We first characterize robust equilibria and then we discuss existence.

**Robustness. Case i, \( p = 0 \).** Workers playing \( n(\theta_0) \geq 0 \) are paid their expected marginal productivity, \( w(n(\theta_0)) = \phi(\theta_0) \). These workers have always the possibility to deviate and play \( n' = 0 \) which yields at least a wage equal to \( \phi(\bar{\theta}) \), Therefore, the minimum value of the equilibrium payoff for unskilled workers in a SPE is associated with a critical value of \( n(\theta_0) \), which we call \( \pi(\theta_0) \), such that

\[
[T - \pi(\theta_0)] \phi(\theta_0) - (c + \bar{\theta}^{-1})\pi(\theta_0) = T\phi(\bar{\theta}) \Rightarrow \pi(\theta_0) = \frac{\phi(\theta_0) - \phi(\bar{\theta})}{\phi(\theta_0) + c + \bar{\theta}^{-1}}.
\]

Correspondingly, the maximum value of \( n \) that unskilled workers are willing to deviate to if perceived as skilled workers by doing so – in the SPE considered – is,

\[
\pi^{\max}(\bar{\theta}) = \frac{T [\phi(\bar{\theta}) - \phi(\bar{\theta})]}{\phi(\bar{\theta}) + \bar{\theta}^{-1} + c}.
\]
We note that the correspondent value for skilled workers, $\pi^{\text{max}}(\theta)$, satisfies, $\pi^{\text{max}}(\theta) > \pi^{\text{max}}(\theta)$. Then, only if $\pi^{\text{max}}(\theta)$ exceeds the maximum value of $n$ that poor workers can afford, $\pi(\omega)$, there exist robust SPE. Otherwise, poor-skilled workers always separate and the SPE is not robust. The relevant condition is

$$\pi^{\text{max}}(\theta) > \pi(\omega) \Rightarrow \frac{T [\phi(\theta) - \phi(\theta)]}{\phi(\theta) + \theta^{-1} + c} > \frac{\omega}{c}. \quad (A.14)$$

which reduces, to $\phi(\theta) > \phi^{\text{max}}$, which is necessary and sufficient for a SPE to be robust.

**Robustness. Case ii, $p > 0$.** If a fraction $p > 0$ of unskilled workers separate by playing $n(\theta) = 0$, and gets a wage $w(n(\theta)) = \phi(\theta)$, while a fraction $1 - p$ pools with poor-skilled workers playing $n(\theta_p) > 0$. Unskilled workers must be indifferent between playing $n(\theta) = 0$ and $n(\theta_p) > 0$. But then, $n(\theta_p) = \pi(\theta_p)$ must hold, where,

$$[T - \pi(\theta_p)] \phi(\theta_p) - (c + \theta^{-1}) \pi(\theta_p) = T \phi(\theta) \Rightarrow \pi(\theta_p) = \frac{\phi(\theta_p) - \phi(\theta)}{\phi(\theta)} + c + \theta^{-1}. \quad (A.15)$$

Hence, the same analysis of case i applies, so that we conclude that $\phi(\theta) > \phi^{\text{max}}$ is the necessary and sufficient condition for a robust SPE, also when $p > 0$.

We are now ready to characterize robust SPE. We look first at the equilibrium values of $n$ and then at labor supply and wages.

**Case i, $p = 0$.** All types of worker but rich-skilled ones are playing $n(\theta_0) \in [0, \pi(\theta_0)]$, and get a wage $w(n(\theta_0)) = \phi(\theta_0)$. Given $n(\theta_0)$, the value of $n(\theta, \omega)$ played by rich-skilled is

$$n(\theta, \omega) = \frac{T [\phi(\theta) - \phi(\theta_0)] + n(\theta_0) [\phi(\theta_0) + c + \theta^{-1}]}{\phi(\theta) + c + \theta^{-1}}, \quad (A.16)$$

which is the minimum value of $n$ such that the ICC of rich-unskilled workers is satisfied, which completes the equilibrium characterization of education levels.

**Case ii, $p > 0$.** In this case, a fraction $p$ of unskilled workers is playing $n(\theta)$, and gets a wage $w(n(\theta)) = \phi(\theta)$, so that as we already know, the level of education played by the pool of poor-skilled and the fraction $1 - p$ of unskilled must be equal to $\pi(\theta_p)$. Correspondingly, $\pi(\theta, \omega)$ is given by equation (A.16) as before.

Finally, relative supply of workers by level of education is as follows.

$$L(n(\theta)) = p(1 - \pi). \quad (A.19)$$
A.2.1 Existence

It is immediate to show that a (non robust) SPE where all workers but rich-skilled pool together always exists. Hence, $\phi \leq \phi^{\text{max}}$ is both necessary and sufficient for the existence of a robust SPE. Let us now analyze the necessary and sufficient conditions for (i) A continuum of SPE in which all workers but the rich-skilled ones pool together playing $n(0) \in [0, \pi(\theta_0)];$ (ii) A continuum of SPE in which a fraction $p \in [0, p^{\text{max}}]$ of unskilled workers separate, while the rest pools with the poor-skilled ones.

We know that in a SPE where a fraction $p > 0$ unskilled workers separate, $n(\theta_p) = \pi(\theta_p)$ must hold. Hence, $\pi(\theta_p) \leq n^{\text{max}}(\omega)$ is a necessary and sufficient condition for the existence of such equilibrium, that is,

$$\frac{T[\phi(\theta_p) - \phi(\theta)]}{\phi(\theta_p) + \theta^{-1} + c} \leq \frac{\omega}{c}. \tag{A.20}$$

Solving for $\phi(\theta_p)$ we find,

$$\phi(\theta_p) \leq \frac{c(T\phi(\theta) + \omega) + \frac{\omega}{2}}{Tc - \omega} \equiv \phi^{\text{max}}. \tag{A.21}$$

$\phi(\theta_p)$ is strictly increasing in $p$. Hence,

$$\phi(\theta_0) \leq \phi^{\text{max}}, \tag{A.22}$$

is the necessary condition for the existence of a pooling equilibrium where a fraction $p > 0$ of unskilled workers separate. Whenever the above condition is satisfied, there exist a maximum value of $p^{\text{max}} \in [0, 1)$ such that, $\phi(\theta_p) \leq \phi^{\text{max}}$ holds for $p \leq p^{\text{max}}$. It is then immediate to verify for any $p \in [0, p^{\text{max}}]$ a SPE as characterized above exists. Viceversa, if the above condition does not hold, no SPE with $p \in (0, p^{\text{max}}]$ exists, and the only SPE is one in which no unskilled worker separate, i.e. $p = 0$. □

A.3 Proof of proposition 3

Characterization of robust equilibria. Lemma 2 applies. In any robust SSE, (i) workers of skills $\theta$ play all the same value of $n$, $n(\theta)$, with $n(\theta) \neq n(\bar{\theta})$; (ii) $n(\theta) = 0$, holds. Accordingly, relative supply of workers conditional on education is given by

$$l(n(\bar{\theta})) = \pi, \tag{A.23}$$

$$l(n(\theta)) = (1 - \pi). \tag{A.24}$$

Supply of labor by level of skills is given by $L(\bar{\theta}) = \pi[T - n(\bar{\theta})]$, and $L(\theta) = (1 - \pi)T$, respectively. Equilibrium salaries are $w(n(\bar{\theta})) = \bar{w}$, and $w(n(\theta)) = w$, where the values of $\bar{w}$, and $w$ are found substituting for the equilibrium values of labor supply in equations, (16) and (17).

Constraints A.1-A.4 apply. Given (A.4) unskilled workers have an incentive mimic skilled ones so long as $n(\theta)$ satisfies,
$\frac{|T - n(\bar{\theta})|w(n(\bar{\theta}))}{Tw(n(\bar{\theta}))} \leq \frac{n(\bar{\theta})}{Tw(n(\bar{\theta}))}(c + \theta^{-1}) + 1$, \hspace{1cm} (A.25)

where

$\frac{w(n(\bar{\theta}))}{w(n(\theta))} = \left(\frac{(T - n(\bar{\theta}))\pi}{T(1 - \pi)}\right)^{-\frac{1}{\sigma}} \left(\frac{\psi(\bar{\theta})}{\psi(\theta)}\right)^{\frac{\sigma - 1}{\sigma}}$. \hspace{1cm} (A.26)

Substituting in for the equilibrium values of wages, define

$n : \frac{|T - n|w}{Tw} = \frac{n}{Tw}(c + \theta^{-1}) + 1$, \hspace{1cm} (A.27)

such that unskilled workers are willing to mimic skilled ones for any $n(\bar{\theta}) \leq n$.

Similarly, skilled workers have incentive to signal their skills by choosing $n(\bar{\theta})$ so long as,

$\frac{|T - n(\bar{\theta})|\bar{w}}{Tw(n(\bar{\theta}))} \geq \frac{n(\bar{\theta})}{Tw(n(\bar{\theta}))}(c + \theta^{-1}) + 1$, \hspace{1cm} (A.28)

so that we can define

$\bar{n} : \frac{|T - \bar{n}|\bar{w}}{Tw} = \frac{\bar{n}}{Tw}(c + \theta^{-1}) + 1$, \hspace{1cm} (A.29)

such that skilled workers are willing to play $n(\bar{\theta})$ and separate for any $n(\bar{\theta}) \leq \bar{n}$.

It is crucial to note that,

$\frac{\bar{n}}{Tw}(c + \theta^{-1}) + 1 > \frac{n}{Tw}(c + \theta^{-1}) + 1$,

for any $w$, which directly implies $\bar{n} > n$, which guarantees that there always exist a non empty set of values for $n(\bar{\theta})$ such that workers’ ICCs are satisfied. Clearly, in any robust SSE, $n(\bar{\theta}) = n$, must hold (Riley’s outcome).

**Existence.** It can be easily verified that:

1. the LHS of (A.28) is strictly decreasing in $n(\bar{\theta})$, tends to 0 for $n(\bar{\theta}) \to T$, and to

   $\left(\frac{\pi}{1 - \pi}\right)^{-\frac{1}{\sigma}} \left(\frac{\psi(\bar{\theta})}{\psi(\theta)}\right)^{\frac{\sigma - 1}{\sigma}}$, for $n(\bar{\theta}) \to 0$.

2. The RHS of (A.28) is strictly increasing in $n(\bar{\theta})$, tends to 1 for $n(\bar{\theta}) \to 0$, and to

   $\frac{1}{\psi(\bar{\theta})}(c + \theta^{-1}) + 1$, for $n(\bar{\theta}) \to T$. 

34
Accordingly, if and only if,
\[ \psi(\bar{\theta}) > \psi(\theta) \left( \frac{\pi}{1 - \pi} \right)^{\frac{1}{\sigma - 1}} \equiv \psi_{\text{min}}, \quad (A.30) \]
there exist \( n > 0 \). If condition (A.30) does not hold, given a candidate SSE, skilled workers will never prefer to separate, which destroys the candidate equilibrium.

Existence requires not only (A.30) to be satisfied, but also, \( n \leq n_{\text{max}}(\omega) \), so that poor-skilled workers can afford the investment in education necessary in order to separate. Provided that (A.30) holds, imposing, \( n = n_{\text{max}}(\omega) \), substituting in for the values of \( w, \) and \( w \) in condition (A.27), we define the critical level \( \psi_{\text{max}} \) of the productivity of skilled workers, such that:
\[ \left( \frac{T - n_{\text{max}}(\omega)}{T} \right) \left( \frac{L(\bar{\theta})}{L(\theta)} \right)^{\frac{1}{\sigma}} \left( \frac{\psi_{\text{max}}}{\psi(\bar{\theta})} \right)^{\frac{1}{\sigma - 1}} = \frac{n_{\text{max}}(\omega)(c + \theta^{-1})}{TL(\bar{\theta})} \left[ \frac{\psi(\bar{\theta})L(\bar{\theta})}{\psi(\theta)L(\theta)} \right]^{\frac{1}{\sigma}} + 1, \quad (A.31) \]
where, \( L(\bar{\theta}) = L_{\text{min}} \), with \( L_{\text{min}} = \pi[T - n_{\text{max}}(\omega)] \), and \( L(\theta) = (1 - \pi)T \). The RHS of (A.27) is decreasing in \( \psi(\bar{\theta}) \) and increasing in the level of education played by skilled workers, \( n \). Differently, the LHS is increasing in \( \psi(\bar{\theta}) \) and decreasing in \( n \). Hence, other things equal, for \( \psi(\bar{\theta}) > \psi_{\text{max}}, n > n_{\text{max}}(\omega) \), so that no SSE exist, while the opposite is true if the reverse inequality holds. \( \square \)

### A.4 Proof of proposition 4

We characterize SPE characterized by multiple levels of education and then we discuss the related existence conditions. Then, we characterize and discuss existence of SPE in which all workers but rich-skilled pool by staying uneducated.

**Characterization of SPE with multiple .** Lemmata 3-4 hold. Therefore, in any SPE with multiple levels of education: 1. Rich and skilled workers always separate; 2. There are at most three levels of education. Moreover, an equivalent argument of lemma 2 holds, in that subset of workers homogeneous in skills who separate, all play the same level of education.

We analyze two separate cases: (i) All workers but rich-skilled pool together \( (p = 0) \); (ii) A fraction \( p > 0 \) of unskilled workers separate, while the rest unskilled workers are pooled with poor-skilled ones.

**Case i, \( p = 0 \).** Rich and skilled workers always separate, by playing \( n(\bar{\theta}, \bar{\omega}) > n_{\text{max}}(\omega) \); while all other workers pool together playing \( n(\theta_0) \geq 0 \). Relative supply of workers by education level is
\[ l(n(\bar{\theta}, \bar{\omega})) = \delta \pi, \quad (A.32) \]
\[ l(n(\theta_0)) = (1 - \delta)\pi + (1 - \delta)(1 - \pi). \quad (A.33) \]
Supply of labor by level of skills is given by
\[ L(\theta) = \pi \delta [T - n(\theta, \omega)] + (1 - \delta) \pi [T - n(\theta_0)], \]
and
\[ L(\theta) = (1 - \pi) [T - n(\theta_0)], \]
respectively. Salaries are,
\[ w(n(\theta)) = \bar{w}, \quad w(n(\theta_0)) = \mu(\theta_0) + [1 - \mu(\theta_0)]w \]
where
\[ \mu(\theta_0) = \frac{(1 - \delta)\pi}{(1 - \delta)\pi + 1 - \pi}, \quad (A.34) \]
and – as usual – the values of \( \bar{w} \) and \( w \) are computed substituting for the equilibrium values of labor supply in equations (16) and (17).

\[ n(\theta_0) \quad \text{and} \quad n(\theta, \omega), \quad \text{must satisfy unskilled ICC, which imply} \]
\[ n(\theta_0) : \frac{[T - n(\theta_0)](\mu(\theta_0) + [1 - \mu(\theta_0)]w)}{T\bar{w}} \geq n(\theta_0)(c + \bar{\theta}^{-1}) + 1, \quad (A.35) \]
and
\[ n(\theta, \omega) : \frac{[T - n(\theta, \omega)]w}{T\bar{w}} = n(\theta, \omega)(c + \theta^{-1}) + 1. \quad (A.36) \]

**Case ii,** \( p > 0 \). Rich and skilled workers always separate, by playing \( n(\theta, \omega) > n^\max(\omega) \). A fraction \( p \) of unskilled workers separate by playing \( n(\theta) = 0 \). A fraction \( 1 - p \) of unskilled workers pool together with poor-skilled workers playing \( n(\theta_p) \). Supply of workers by education is
\[ l(n(\theta, \omega)) = \delta \pi, \quad (A.37) \]
\[ l(n(\theta_p)) = (1 - \delta)\pi + (1 - \delta)(1 - \pi), \quad (A.38) \]
\[ l(n(\theta)) = \pi(1 - \pi). \quad (A.39) \]

Supply of labor by level of skills is given by
\[ L(\theta) = \pi \delta [T - n(\theta, \omega)] + (1 - \delta)\pi [T - n(\theta_p)], \]
and
\[ L(\theta) = (1 - \pi)(1 - p)[T - n(\theta_p)] + (1 - \pi)pT, \]
respectively. Salaries are,
\[ w(n(\theta, \omega)) = \bar{w}, \quad w(n(\theta)) = \bar{w}, \quad \text{and} \quad w(n(\theta_p)) = \mu(\theta_p) + [1 - \mu(\theta_p)]w \]
where
\[ \mu(\theta_p) = \frac{(1 - \delta)\pi}{(1 - \delta)\pi + (1 - \pi)(1 - p)}, \quad (A.40) \]
\[ n(\theta_p) \quad \text{and} \quad n(\theta, \omega), \quad \text{must satisfy unskilled ICC, which imply} \]
\[ n(\theta_p) : \frac{[T - n(\theta_p)](\mu(\theta_p) + [1 - \mu(\theta_p)]w)}{T\bar{w}} = n(\theta_p)(c + \bar{\theta}^{-1}) + 1, \quad (A.41) \]
and
\[ n(\theta, \omega) : \frac{[T - n(\theta, \omega)]w}{T\bar{w}} = n(\theta, \omega)(c + \theta^{-1}) + 1. \quad (A.42) \]

36
Robustness. The minimum value of the equilibrium payoff for unskilled workers in a SPE is associated with a critical value of \( n(\theta_p) \), which we call \( \pi(\theta_p) \), such that

\[
[T - \pi(\theta_p)] \{ \mu(\theta_\square(\theta_p)) \overline{w} + (1 - \mu(\theta_\square(\theta_p))) \} \overline{w} = \left[ c + \theta^{-1} \right] \pi(\theta_p) + T \overline{w}.
\]

(A.43)

Note that in case ii above, \( n(\theta_p) = \pi(\theta_p) \) holds in any SPE. Rearranging terms, the above condition can be rewritten as:

\[
[T - \pi(\theta_p)] \mu(\theta_\square(\theta_p)) \overline{w} - \pi(\theta_p) \overline{w} = \frac{\pi(\theta_p)}{w + c + \theta^{-1}}.
\]

(A.44)

Given that the lowest expected payoff of unskilled workers in a SPE is \( \overline{w} T \), the maximum value of \( n \) that unskilled workers are willing to deviate to in a SPE, if perceived of skill \( \theta \) by doing so, is \( \pi(\theta) \) such that

\[
[T - \pi(\theta)] \overline{w} = \left[ c + \theta^{-1} \right] \pi(\theta) + \overline{w} T.
\]

(A.45)

Other things equal, given \( n(\theta_p) > 0 \), as \( p \) increases, \( L(\theta) \) increases. In turns, this implies that \( \overline{w} \) will decrease, while \( \overline{w} \) will increase. Accordingly, both \( \pi(\theta_p) \) and \( \pi(\theta) \) are both strictly increasing functions of \( p \). Therefore, if

\[
\lim_{p \to 1} \pi(\theta) \leq n_{\text{max}}(\omega),
\]

(A.46)

no SPE is robust. Otherwise, there exist a critical value of \( p \), strictly lower than one, call it \( \check{p} < 1 \), such that SPE with \( p \geq \check{p} \) are robust.

Consider an SPE in which the intermediate level of education is \( \pi(\theta_p) \) and \( p \to 1 \). Then, \( \mu(\theta_\square(\theta_p)) \to 1 \), so that \( w(\pi(\theta_p)) = \overline{w} \). In turns, this implies \( \pi(\theta_p) = n(\theta, \overline{w}) \) must hold. By imposing \( n(\theta, \overline{w}) = n_{\text{max}}(\omega) \), we define a critical value \( \psi^\text{max} \) for \( \psi(\theta) \), such that:

\[
\left( \frac{T - n_{\text{max}}(\omega)}{T} \right) \left( \frac{L(\theta)}{L(\theta)} \right)^{\frac{1}{\hat{p}}} \left( \frac{\psi^\text{max}}{\psi(\theta)} \right)^{\frac{\sigma}{\sigma - 1}} = \frac{n_{\text{max}}(\omega)(c + \theta^{-1})}{TL(\theta)^{-\frac{1}{\hat{p}}} \psi(\theta)^{\frac{\sigma}{\sigma - 1}} \left[ \psi(\theta)L(\theta)^{\frac{\sigma}{\sigma - 1}} + \left( \psi^\text{max} L(\theta)^{\frac{\sigma}{\sigma - 1}} \right)^{\frac{\sigma}{\sigma - 1}} \right]^{\frac{\sigma}{\sigma - 1}}} + 1.
\]

(A.47)

It is then immediate to verify that condition A.46 is equivalent to \( \psi(\theta) \leq \psi_{\text{max}}(\omega) \). Furthermore, given \( n(\theta_p) = n(\theta, \overline{w}) = n_{\text{max}}(\omega) \), the values of supply of labor by level of education are \( L(\theta) = L_{\text{min}} \), with \( L_{\text{min}} = \pi[T - n_{\text{max}}(\omega)] \), and \( L(\theta) = (1 - \pi)T \). It then follows that the values of wages are the same as in the SSE equilibrium when \( n(\theta) = n_{\text{max}}(\omega) \) so that \( \psi^\text{max} = \psi_{\text{max}} \), holds.

Hence, we conclude that a necessary and sufficient condition for a robust SPE with multiple levels of education is \( \psi(\theta) > \psi_{\text{max}} \).

Existence. It is immediate to verify that, by definition, \( \psi_{\text{min}} < \psi_{\text{max}} \) so that if \( \psi(\theta) < \psi_{\text{min}} \), no robust with multiple levels of education SPE exists.
Given that \( \psi(\theta) \geq \psi_{\text{min}} \) is also necessary for the existence of a SSE, if \( \psi(\theta) < \psi_{\text{min}} \), it is immediate to verify that the only equilibrium is a perfect pooling equilibrium with no education, in which a fraction \( f < 1 \) of skilled workers performs high skills tasks, and a fraction \( 1 - f \) performs low skills tasks, so that \( w(\theta) = w(\theta) \). This fraction \( f \) satisfies,

\[
\left( \frac{\pi f}{\psi(\theta)} \right)^{\frac{\sigma}{\sigma-1}} = \left( 1 - \pi f \right)^{-\frac{1}{\sigma}} \left( \psi(\theta) \right)^{\frac{\sigma}{\sigma-1}} \Rightarrow f = \frac{\left( \frac{\psi(\theta)}{\psi(\theta)} \right)^{\frac{\sigma}{\sigma-1}}}{\pi \left( 1 + \left( \frac{\psi(\theta)}{\psi(\theta)} \right)^{\frac{\sigma}{\sigma-1}} \right)}.
\]

(A.48)

Clearly, \( f \) is a strictly increasing function of \( \psi(\theta) \), with \( f = 0 \) for \( \psi(\theta) = 0 \), \( f = \infty \) for \( \psi(\theta) = \infty \), and \( f = 1 \) for \( \psi(\theta) = \psi_{\text{min}} \). \( \square \)

Figure 5: PGs vs CGs: Differences in (a) Real, log weekly wages for full time workers, and (b) Nominal, log average annual tuition and fees.

Source: Data on fees and tuition are taken from NCES Digest of Education Statistics, 2011, tables 349 and 352. Regarding data on wages, see note to figure 1. We used nominal data for fees, since we could not find official real data for undergraduate fees and tuition net of room, and board.

\[ \text{If education were positive, workers would find it optimal to deviate by choosing less education.} \]