Double Screening in the Labour Market

Luis Pinedo Caro*

University of Southampton

May 21st, 2014

Abstract

We set up a model of the labour market where firms do not observe the quality of the workers hired. Performance pay and signalling are common screening mechanisms used to improve the outcome of these situations. This paper analyzes the trade-off firms face when deciding the intensity of each screening method. In addition we propose two more applications of the model; one analyzes the optimal level of human capital accumulation. The other one provides insights on the effect of double screening in the income distribution.

Keywords: Signalling, Performance Pay, Adverse Selection.


*lpc1e09@soton.ac.uk
1 Introduction

Economists have long been aware of the importance of asymmetric information in the proper functioning of markets. The labour market is not an exception as employers might not have full information about workers’ productivity. The problems arising from this lack of information are split in two categories, adverse selection (unknown workers’ type) and moral hazard (unknown workers’ effort). This paper deals with how adverse selection is solved, while dismissing issues of moral hazard as they are out of the scope of this research.

Under adverse selection firms lack knowledge about the actual ability of workers. It is generally assumed actual ability is a worker’s private information and thus hard to observe quickly and reliably. Not knowing the workers’ ability means, in principle, that an employer has to believe a worker’s declaration of his ability. Since wages are normally higher for high ability workers, low ability ones have an incentive to pretend otherwise. This might induce, if the asymmetry is severe enough, a cessation of economic activity as shown by Akerlof (1970). The problem so far is that sometimes workers cannot be trusted. Some of their -observable- actions, though, may carry information correlated with their actual ability.

As noted in Riley’s (2001) survey two such common observable variables are the workers’ educational and performance choices of workers. The former gives space to the signaling literature pioneered by Spence (1973), where firms screen workers’ educational choices and allocate tasks more efficiently; the latter was initially approached by Rothschild & Stiglitz (1977) and uses the ability to screen the performance of workers to set wages that induce workers to self-select themselves onto the right task. In this paper the objective is to merge both approaches so that I obtain a model of double screening.

The importance of double screening can be observed in empirical studies like the ones of Altonji & Pierret (2001) or Lang & Siniver (2011). These two papers show that even though starting salaries do depend on easily available information (years of schooling, name of the school attended), experienced workers’ earnings tend to depend on actual ability. This finding suggests that firms are capable of learning a worker’s ability and thus both types of screening are used at the same time.

The use of signals arising from the education system as a screening mechanism by firms is pervasive but performance pay is also becoming a common feature in high skilled workers’ contracts as pointed out in Lemieux et al. (2009). These authors also point that income inequality increases in jobs affected by performance pay.

This paper offers three main contributions. The first one is to offer a rigorous theory of the labour market where employers’ learning is modelled. This theory allows us to learn from the optimal screening behaviour of firms and to provide some insights into the variables affecting it. I shows how firms’ screening behaviour is affected by the information carried by the ex-ante signals and ex-post signals and I link it to empirical studies on the same issue.

Then I link income distributions obtained from simulations of the model in steady state with the evolution of the US income distribution over the last 30 years. I then offer an explanation for
the changes in the shape of the US income distribution based on an increase in performance related payments.

The third contribution is in the field of human capital. I consider whether the level of human capital accumulation implied by the model under perfect competition is optimal or not. To assess that consideration we include a constrained planner who faces the same informational asymmetry as firms. I compare the investments in human capital generated by both the planner and firms in perfect competition. The results suggest perfect competition induces either under-investment or over-investment in human capital as well as lower welfare compared to the one achieved by the planner.

The last result opens the door to considering different market structures other than perfect competition. The last contribution relates market structure to workers’ welfare. I compare two opposite market structures, perfect competition and a zero-profit monopsonist. The welfare induced by the zero-profit monopsonist is always higher than the one attained under perfect competition. This situation is softened as I allow the monopsonist to obtain positive levels of profits. I find that how much profit the monopsonist makes, defines the situations under which a monopsony is preferred. For mild profit levels I find that a monopsony induces higher welfare in economies where the correlation between ex-ante signals and the skill of workers are either relatively high or low.

2 The model

I develop an overlapping generations model with asymmetric information in the labour market. The asymmetry is generated by an employer who does not know the productivity of his employees, which is the private information of the workers. This form of asymmetric information gives rise to a problem of adverse selection as the one shown in Akerlof (1970). Akerlof, in his paper, shows that asymmetric information could even cause the cessation of market activities. Typical remedies to the problem of asymmetric information include the use of signals from the education system and ex-post observations of workers’ on-the-job performance, hereinafter referred as ex-ante and ex-post signals, respectively.

The model developed in this paper has the spirit of Coate & Loury’s (1993) signaling model, although wages are endogenous following Moro & Norman’s (2004) general equilibrium model. A crucial innovation that is not present in the two papers just mentioned is the ability of the firms to condition wages on observed performance. When allowing the use of ex-ante and ex-post signals by firms a model of double screening arises. This feature has scarcely been utilised in the literature but is implemented by Gall et al. (2006).

This model also features endogenous human capital accumulation. This part is crucial to show the effects of firms’ current hiring policies on tomorrow’s educational choices. In this model workers have the typical motivations for a worker investing in human capital they expect stronger signals and to become more productive when handling demanding tasks.

The description of the model starts with the agents (workers and firms) and their strategic inter-

---

1 See Becker (1975) or Spence (1973)
actions. Then a definition of equilibrium is provided.

2.1 Workers

2.1.1 Demographic structure

Every period a unit continuum of workers is born. Each generation of workers lives for three periods that represent the workers’ youth, adulthood and retirement. Workers are endowed with one unit of time which is inelastically supplied to the labour market in their adulthood.

2.1.2 Types and skills

Workers’ type is given by the cost of investing in human capital. This disutility, denoted \( \phi \), is randomly drawn from the distribution \( F_{\phi} \).

The investment in human capital takes place when workers are young; if a worker willing to face the cost \( \phi \) he becomes high skilled (\( H \)), otherwise he becomes low skilled (\( L \)). I denote the investment in human capital with an indicator function, \( I = 1 \), if a worker invests and \( I = 0 \) if he does not. The cost of investing in human capital is assumed to enter the utility function linearly\(^2\). Finally I denote the share of high skilled individuals by \( \pi \).

2.1.3 Preferences

Workers derive utility from consuming during their adulthood and their retirement\(^3\). Their preferences over consumption are summarised by an isoelastic utility function that weights consumption during adulthood \( c^a \) and old age \( c^o \),

\[
    u(c^a, c^o) = \frac{c^a^{1-\epsilon} - 1}{1 - \epsilon} + \beta \frac{c^o^{1-\epsilon} - 1}{1 - \epsilon} - \phi,  
\]

where \( \epsilon \) is the coefficient of relative risk aversion and \( \beta \) the discount rate. I assume workers face a borrowing constraint. Given wages when they are adults and old, \( \{w^a, w^o\} \), the indirect utility function -net of investment costs- is defined as:

\[
    V(c^a, c^o) = \max_{c^a, c^o} u(w^a, w^o)  
\]

2.1.4 Outside options

High skilled and low skilled workers have outside options denoted by \( \{\bar{U}^H, \bar{U}^L\} \). Outside options reflect the utility that low \( L \) skilled and high \( H \) skilled workers could obtain if they put their labour back in the market.

\(^2\)This is assumed for simplicity. It does not affect qualitatively the results shown.

\(^3\)They do not explicitly consume while young. I can assume current adults feed the new generation and that is embedded in their utility.
2.2 Firms

There are two infinitely-lived firms competing á la Bertrand. These firms are myopic because they only take into account current profit when deciding their actions. At time 0, I assume that firms hold a belief on the share of high skilled workers in the population \( \{\pi_0\} \).

2.2.1 Production

Firms produce output, \( Y \), that represents a final consumption good. Production of the good requires the performance of both a simple and a complex task. The technology that firms use to produce the consumption good is represented by a CES function mapping complex and simple tasks’ efficiency units of labour \( \{Y : \mathbb{R}_+^2 \to \mathbb{R}\} \) into output,

\[
Y = A (\alpha C^\gamma + (1 - \alpha)S^\gamma)^{1/\gamma},
\]

where \( A \) stands for the total factor productivity, \( \alpha \) is the output share associated with the complex task and \( \gamma \) determines the degree of substitutability between the efficiency units of each task. The production function satisfies the Inada conditions,

\[
\lim_{C \to \infty} \frac{\partial Y}{\partial C} = \lim_{S \to \infty} \frac{\partial Y}{\partial S} = 0, \quad \text{and}
\]

\[
\lim_{C \to 0} \frac{\partial Y}{\partial C} = \lim_{S \to 0} \frac{\partial Y}{\partial S} = \infty.
\]

The complex task is done more efficiently by high skilled workers (with productivity \( A_h \)) than by low skilled workers (with productivity \( A_L \)), with \( A_h > A_L \geq 0 \). On the contrary, both types are equally efficient at doing the simple task, with productivities equal to 1.\(^5\) The efficiency units of labour for each task are made of the mass of high (\( H \)) and low (\( L \)) skilled workers performing those tasks times their respective productivities. Note the superscripts \( c \) and \( s \) reflect the task performed.

\[
C = A_h H^c + A_L L^c
\]

\[
S = H^s + L^s
\]

2.2.2 Information

Firms have a belief, \( \pi \), -that might be wrong- about the share of high skilled workers, yet they do not know, who is high skilled and who is not. This information asymmetry generates a problem of

\(^4\)The myopia is assumed to prevent the model being in steady state all the time. By doing this I obtain dynamics to the steady state.

\(^5\)This assumption does not affect the final result. The firms’ objective of matching high skilled workers with complex tasks would remain unaltered had I assumed low skilled workers were more productive than high skilled workers at doing simple tasks.
adverse selection. To counter this problem I assume that firms observe two noisy signals that are correlated with the skill level. Adverse selection is, as a result, partially mitigated because firms have the opportunity to use the signals to improve the skill composition of the workers on each task.

**Ex-ante signal:** The first of these signals is denoted \( \theta \). This signal is randomly drawn from one of two distributions \( F_h \) or \( F_L \). High skilled workers send a signal drawn from \( F_h \) while low skilled workers draws it from \( F_L \), with \( F_h \) first order stochastically dominating \( F_L \). Employers observe this signal at the moment of hiring workers. The ex-ante signal is best thought as a combination of factors that are observable by a firm when meeting an applicant. These factors might include the CV, the job interview or the neighbourhood where he lives, among others.

**Ex-post signal:** The second signal, denoted \( y \), is generated after the completion of a complex task by a worker. The ex-post signal takes on two values, ‘pass’ or ‘fail’ \( y \in (p, f) \). High skilled workers performing the complex task send out a ‘pass’ signal with probability \( \eta \) one. On the contrary, the low skilled workers who perform the complex task send out ‘pass’ signals with probability \( \eta \in (0, 1) \). I refer to \( \eta \) as the probability of fooling the firm. Low skilled workers are heterogeneous in the probability of obtaining a ‘pass’ signal as \( \eta \) is a random variable drawn from the distribution \( F_\eta \).

### 2.2.3 Hiring game

When workers become adults they randomly choose one of the two existing firms. At this point there is an exchange of information between the firm and the worker; the worker sends a signal, \( \theta \), to the employer and, at the same time, he gets to know the probability, \( \eta \), of obtaining a ‘pass’ signal if he ended up doing complex tasks.

Then, the firm offers a contract that involves either performing simple or complex tasks. Contracts in our setting are defined as follows:

---

6This is a simplification that does not affect the results. We could assume a lower probability insofar the signal remains informative.

7Firms are assumed to be identical. This means all firms receive the same signal \( \theta \) from each worker and it is equally easy -or difficult- for a worker to receive a ‘pass’ signal from any of them.
**Definition** A contract between a worker and a firm is defined as a triplet \( \{w^a, w^o, T\} \) that specifies the ex-ante wage \( w^a \), the ex-post wage \( w^o \) and the task \( T \in \{C, S\} \) that will be performed.

For the simple task the firm offers a fixed ex-ante wage \( w^a = w^s \) and no ex-post wage \( w^o = 0 \). None of the payments in the simple task contract depend on the signals because firms know with certainty the workers’ productivity when doing simple tasks. The contract for the complex task uses, though, both signals. The form of the ex-ante wage is given by \( w^a = p(\theta, \bar{\pi})w^c \), where \( p(\theta, \bar{\pi}) \) is the Bayesian posterior probability of being high skilled given the signal,

\[
p(\theta, \bar{\pi}) = \frac{\bar{\pi}f_h(\theta)}{\bar{\pi}f_h(\theta) + (1 - \bar{\pi})f_l(\theta)}.
\]

In turn, the ex-post wage is given by:

\[
w^o = \begin{cases} w^p & \text{if } y = p \\ 0 & \text{otherwise} \end{cases}
\]

where \( w^p \in \mathbb{R}_+ \) and firms pay \( w^p \) conditional on the ex-post signal being a ‘pass’. The structure of the ex-ante payment in the complex task contract is taken from Moro & Norman (2003) and I keep it to allow for comparability in the results.

Firms can potentially use the ex-ante and ex-post signals to improve the sorting of high skilled workers into complex tasks; this is done in two ways called Firms’ Selection (FS) and Self-Selection (SS). On the one hand Firms’ Selection refers to the use of a hiring threshold. A hiring threshold, \( \{\bar{\theta}\} \), represents the minimum signal a worker needs to be offered a complex task job. Since those who invest in human capital receive, on average, higher signals, setting hiring thresholds may allow the firms to attract more high skilled workers to the complex task.

On the other hand, Self-Selection refers to type revelation by low skilled workers initially assigned to a complex task. We denote as \( F_D \) the proportion of low skilled workers invited to perform the complex task and who accept the invitation, hereinafter called ‘deceivers’. Firms can decrease the size of \( F_D \) with the help of contracts by making deception less attractive.

Given hiring thresholds and contracts for simple and complex tasks, I obtain the mass of high and low skilled workers firms expect to hire for each task,

\[
H^c = [1 - F_h(\bar{\theta})] \bar{\pi}, \quad L^c = [1 - F_l(\bar{\theta})](1 - \bar{\pi})F_D
\]

\[
H^s = F_h(\bar{\theta}) \bar{\pi}, \quad L^s = F_l(\bar{\theta})(1 - \bar{\pi}) + [1 - F_l(\bar{\theta})](1 - \bar{\pi})[1 - F_D].
\]

In the above definitions \( H \) and \( L \) stand, respectively, for the mass of high and low skilled workers

---

8In terms of timing \( w^a \) is given when a worker starts his adulthood and \( w^o \) as soon as he becomes old and retires.

9The firm has no incentive to postpone the payment and so we make the calculations simpler by taking away this option.

10Think, for instance, of higher \( w^s \) or \( w^p \).
while the superscripts $c,s$, stand for the task to which they are allocated. As an example, to obtain the expected mass of high skilled workers performing the complex task I multiply the perceived share of high skilled in society $\bar{\pi}$ times the mass of high skilled with signals above the hiring threshold $[1 - F^q_h(\bar{\theta})]$. Calculations for the mass of low skilled on each task follow a similar logic. It should be noted that some low skilled workers initially assigned to a complex task may end up performing simple tasks after revealing their type (the fraction $[1 - F_D]$).

### 2.2.4 Profit

The profit function consists of the revenue earned by the firm minus the wages paid to the workers. Note the price of the consumption good is normalised to 1 and profit is

$$P(\bar{\theta}, w) = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w).$$

Output is defined in equation 1.5; even though the expression in 1.5 does not explicitly depend on either wages or hiring thresholds it does so implicitly through the efficiency units. The expected payroll is calculated assuming that the law of large numbers applies. In particular, the ex-ante payments for workers performing the complex task are calculated by multiplying the expected posterior probability of those who accept the job times the base payment $w^c$. The ex-post, or performance payments are calculated by multiplying the expected mass of workers with a ‘pass’ signal times the individual payment $w^p$. Finally, workers doing the simple task are paid $w^s$. The formula to calculate the payroll is

$$\omega = w^c \left( E[p(\theta)|I=1]H^c + E[p(\theta)|I=0]L^c \right) + w^p \left( H^c + E[\eta]|D=1]L^c \right) + w^s N^s,$$

where $E[p(\theta)|I=1]$ is the expected Bayesian posterior probability of investors doing complex tasks and $E[\eta]|D=1]$ is the expected probability of obtaining a pass in the ex-post signal by low skilled workers doing the complex task.

**Updating of beliefs:** I assume the actual share of high skilled workers can be recovered by the firm after the tasks are done. Firms use this information to update their current beliefs as follows:

$$\bar{\pi}_{t+1} = q\bar{\pi}_t + (1 - q)\pi_t \quad \text{with} \quad q \in (0, 1)$$

where $q$ is an exogenous smoothing parameter and $\pi$ is the current share of high skilled workers in the population. This updating rule generalises the one used in Coate & Loury (1993) and I keep it for its simplicity.\textsuperscript{11} Note that the firm possesses private information about former workers but, since the information is learnt after they have left the firms it does not matter for the analysis.

\textsuperscript{11}See Kim & Loury (2009) for a version of Coate & Loury with forward looking agents.
2.3 Equilibrium

The concept of equilibrium of the economy is Bayes-Nash. Since the properties of the equilibrium differ depending on the time span I define the equilibrium in the short-run and in the steady state. The difference lies in the correctness of the firms’ beliefs with respect to the workers’ investment rate. These beliefs can be wrong in short-run but not in the steady state. Since the short-run equilibrium contains the steady state as a sub-case I define it first.

Equilibrium in the short-run: The equilibrium in the short-run is characterised by the agents maximising their objective functions in addition to the firms updating their beliefs.

Definition Given an initial belief, $\{\tilde{\pi}_t\}$, allocations, $\{\hat{c}_{n,i}, \hat{c}_{o,i}\}$ $\forall i \in (0, 1)$, along with workers’ strategies, $\{\hat{I}_i, \hat{D}_i\}$ $\forall i \in (0, 1)$, firms’ strategies, $\{\hat{\theta}\}$, outside options, $\{\hat{U}^H, \hat{U}^L\}$, and wages $\{\hat{w}^c, \hat{w}^p, \hat{w}^s\}$ constitute a Bayes-Nash equilibrium if they are such that:

i Given market outside options $\{\hat{U}^H, \hat{U}^L\}$ and firms’ beliefs, firms’ strategies $\{\hat{\theta}\}$ and wages $\{\hat{w}^c, \hat{w}^p, \hat{w}^s\}$ solve the firms’ problem.

ii $\forall i \in (0, 1)$, given firms’ strategies, $\{\hat{\theta}\}$, and wages, $\{\hat{w}^c, \hat{w}^p, \hat{w}^s\}$, workers’ strategies, $\{\hat{I}_i\}$, solve the workers’ investment problem.

iii $\forall i \in (0, 1)$, given firms’ strategies, $\{\hat{\theta}\}$, and wages, $\{\hat{w}^c, \hat{w}^p, \hat{w}^s\}$, workers’ strategies, $\{\hat{D}_i\}$, solve the workers deceiving problem.

iv Firms’ expected profit given their beliefs are 0.

v Beliefs are updated according to equation 1.13.

Steady state: A steady state of this economy is characterised by a sequence of short-run equilibria such that the state variables (firms’ beliefs) remain constant. A formal definition follows.

Definition The steady state of the economy is a Bayes-Nash equilibrium constituted by outside options, $\{\hat{U}^H, \hat{U}^L\}$, and firms’ belief, $\{\pi_t\}$, such that the wages $\{\hat{w}^c, \hat{w}^p, \hat{w}^s\}$, and the hiring threshold, $\{\hat{\theta}\}$, set by the firms cause the workers’ investment rate to be $\pi = \tilde{\pi}$.

2.4 Solving for the equilibrium

This section solves for the equilibrium of the model with double screening. I show how the agents solve their problems in the same order as the decisions are taken for a given generation of workers. First firms, given their beliefs, decide the wages and the hiring thresholds. Workers then use the firms’ actions to decide whether to invest in their own human capital or not. Lastly, low skilled workers who are invited to perform the complex task decide whether or not to deceive the firm.
Firms set contracts and thresholds
Young choose human capital
Workers choose to deceive or not
Firms learn $\pi$ and update beliefs

### 2.4.1 The firm’s problem

Since the two firms in the economy are identical I show the behaviour of a representative firm behaving as in perfect competition. Given beliefs $\{\pi_t\}$ and outside options $\{\bar{U}^H, \bar{U}^L\}$ the firm solves,

$$\max_{\bar{\theta}, w^c, w^p, w^s} P = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w)$$

subject to four individual rationality constraints,

$$\bar{U}^H \leq (1 - F_h(\bar{\theta}))V_{E}^{h,c} + F_h(\bar{\theta})V^s$$ \hspace{1cm} IR.H (16)

$$\bar{U}^L \leq (1 - F_l(\bar{\theta}))V_{E}^{l,c}F_D + (1 - F_l(\bar{\theta}))V^s[1 - F_D] + F_l(\bar{\theta})V^s$$ \hspace{1cm} IR.L (17)

and four incentive compatibility constraints,

$$V^{h,c} \geq V^s \quad \forall \theta > \bar{\theta}$$ \hspace{1cm} IC.H (18)

$$V^s \geq V^{l,c}(\theta, \eta), \quad \text{for some } (\eta, \theta)$$ \hspace{1cm} IC.L (19)

where output is given by equation 1.5 and the expected payroll of the firm is defined in equation 1.13. In addition, four Individual Rationality and four Incentive Compatibility constraints must be satisfied. On the one hand, the four IR constraints reflect that the firm must offer at least the outside option for a worker to choose that firm as his workplace, i.e. it is an ex-ante constraint. The IR constraints are skill dependant.

The right hand sides of equations 1.16 and 1.17 reflect the expected utility of, respectively, a high and a low skilled worker, before taking the decision of whether to invest in human capital or not. In terms of notation, value functions are assigned superscripts depending on the task $\{c, s\}$ and the skill of the worker $\{h, l\}$. More specifically, $V_{E}^{h,c}$ is the expected utility that a worker would obtain by investing if he were chosen to perform the complex task. Likewise, $V_{E}^{l,c}$ is the expected utility a worker would obtain if he did not invest but was given the opportunity to perform the complex task and accepted a contract to do so -a deceiver-. Lastly $V^s$ is the utility drawn from performing simple

---

12 Expected because the ex-ante wage depends on the signal and the signal itself is a random variable.
13 There is no expectation as this value is known with certainty by the workers.
tasks.

On the other hand, I add four IC constraints to the problem of the firm, one per skill. First, the IC.H constraints make sure high skilled workers assigned to the complex task do not have an incentive to pretend to be low skilled and thus to switch to the simple task. Then, the IC.L constraints have the opposite purpose; they are intended for low skilled workers assigned to the complex task to leave and perform the simple one instead.

It is important to notice that IC.H is, in equilibrium, satisfied and thus no high skilled worker assigned to a complex task would pretend to be low skilled. This is not necessarily true in the case of the IC.L, and a semi-separating equilibrium is normally\textsuperscript{14} achieved. Only in special cases\textsuperscript{15} would this constraint be satisfied for all low skilled workers and a pure separating equilibrium would exist.

2.4.2 Workers’ problems.

Workers’ payoffs depend on whether they become high skilled or not, and -only for the low skilled workers who are invited to perform the complex task- on whether they deceive their employer or not. Next we show how these decisions are taken and aggregated.

**Human capital decision:** Young workers make their human capital acquisition decisions based on their investment costs and expected future earnings. I define the benefit of investing as

$$\xi = E[U|I = 1] - E[U|I = 0],$$

where $E[U|I = 1]$ and $E[U|I = 0]$ are the expected utility of a worker who, respectively, has and has not invested in human capital. Given information on contracts and hiring thresholds workers compare the payoffs of investing and not investing. This is done by weighting the value at each state by its respective probability.

$$E[U|I = 1] = [1 - F_h(\bar{\theta})]V_{E}^{h,c} + F_h(\bar{\theta})V^s$$

$$E[U|I = 0] = [1 - F_l(\bar{\theta})]V_{E}^{l,c} F_D + [1 - F_l(\bar{\theta})]V^s[1 - F_D] + F_l(\bar{\theta})V^s$$

Every worker whose benefit to investing is higher than his utility cost acquires human capital. The share of workers that choose to acquire human capital, $\pi$, is computed as the mass of workers whose investment cost is below the benefit to invest $\pi = F_\phi(\xi)$.

**Deceiving decision:** Workers invited by the firm to perform complex tasks have to decide whether to accept it or to reject the offer. The incentive compatibility constraint IC.H ensures all high skilled workers assigned to the complex task will accept the invitation. On the contrary, low skilled workers invited to perform complex tasks must decide whether they want to deceive the firm or not. Given

\textsuperscript{14}Suppose $V^s = V^{l,c}$, $\forall (\eta, \theta)$, then $V^s = V^{h,c}$, $\forall \theta$ and no worker wants to perform complex tasks.

\textsuperscript{15}For instance if $E[\eta] = 0$ and $w^a = w^c$, setting $w^c \leq w^a$ guarantees full separation.
workers’ knowledge of their ex-ante signals, \(\theta\), and their probabilities, \(\eta\), of fooling -obtaining a ‘pass’ signal \(y = p\) - the firm, I define a worker’s expected utility of deceiving (D=1) and revealing the type (D=0) as,

\[
E[U|D = 1] = V_E^{l,c}(p(\theta)w^e, w^p, \eta)
\]  

(23)

\[
E[U|D = 0] = V^s(w^s, 0).
\]  

(24)

To calculate the share of deceivers I first define the benefit of deceiving the firm, \(\delta = E[U|D = 1] - E[U|D = 0]\).

(25)

Workers with combinations of \((\theta, \eta)\) such that \(V^l_{E} > V^s\) will attempt to deceive the firm (D=1) and perform the complex task. The share of these low skilled workers invited to perform the complex task who deceive the firm is given by the joint CDF of the signals and the fooling probabilities denoted \(F_D\).

\[
F_D = \int_{\theta_{\min}}^{\theta_{\max}} \int_{\eta_{\min}(\theta)}^{\eta_{\max}} f_\eta(\eta, \eta^l, \eta^h)d\eta d\theta
\]  

(26)

In addition, I need to calculate the expected probability of obtaining a ‘pass’ signal \(y = p\) given deception \(E[\eta|D = 1]\). This is given by:

\[
E[\eta|D = 1] = \int_{\theta_{\min}}^{\theta_{\max}} \int_{\eta_{\min}(\theta)}^{\eta_{\max}} \eta f_\eta(\theta, \eta, \infty)f_\eta(\eta, \eta_{\min}, \eta) d\eta d\theta
\]  

(27)

Note that in equation 1.27 the lower truncation of \(f_\eta\) depends on \(\eta_{\min}\). Also the lower bound of the inner integral \(\eta_{\min}(\theta)\) depends on the outer variable. The function \(\eta_{\min}(\theta)\) is obtained by first making \(\delta = 0\). This function combines pairs of signals and fooling probabilities such that a worker is indifferent between performing complex or simple tasks. If I solve in terms of the fooling probability\(^{16}\) the result is

\[
\eta_{\min} = \frac{V^s}{V_E^{l,c}(p(\theta)w^e, w^p)^{-1}}.
\]  

(28)

This equation provides the lower bound fooling probabilities that are needed for a worker to deceive given his signal. See appendix A for extra details on these calculations.

### 2.4.3 Outside options

The outside options, \(\{\bar{U}^H, \bar{U}^L\}\), one per skill, are determined in equilibrium using the zero-profit condition. To pin them down we choose \(\bar{U}^H\) such that the profit of the representative firm is equal to 0. Then, \(\bar{U}^L\) is given by the expected utility of not acquiring human capital (right hand side of equation 1.20). Note I do not impose a value for \(\bar{U}^L\).

The way to pin down \(\bar{U}^L\) as a by-product of \(\bar{U}^H\) follows an insurance motive. Given the imperfection

---

\(^{16}\)Provided the inverse of \(V\) exists.
of the educational signals, in the unlucky event that a high skilled worker is sent to perform the simple task, a higher salary, \( w^s \), helps the firm to reduce the overall\(^{17}\) payroll. This procedure has implications for the functioning of the labour market. Under perfect information, and assuming workers cannot opt for unemployment benefits and have no other form of wealth, the procedure would lead to low skilled workers being paid the minimum necessary to survive (zero in my model).

### 2.4.4 Market clearing

Two markets must clear in each period, a goods market and two labour markets.

**Goods market:** Equilibrium in the goods market means that the amount of good supplied equals the amount demanded. In the short-run equilibrium the demand of goods might be higher that the production (if the firms’ beliefs are wrong), in that case I assume firms hold a finite stock of the consumption good that is used to keep up their payment promises. The equation for the goods market equilibrium is

\[
C_t = Y_t + S_t, \tag{29}
\]

where \( C \) is total consumption, \( Y \) is output and \( S \) is stocked consumption good by the firm.

**Labour markets:** Firms’ demand for high and low skilled workers shape the market. Since workers supply their labour inelastically labour markets always clear.

\[
N_{dH} = N_{sH}; \quad N_{dL} = N_{sL} \tag{30}
\]

Firms demand for workers is set when they announce hiring thresholds and wages. If their beliefs are wrong the mass of workers with signals above and below the hiring threshold may not be the expected one. In those cases I allow the firm to change the hiring threshold. If the mass of workers with signals above the threshold is higher than the expected one, then market clearing occurs by the firm tightening the hiring threshold for the complex task, thus only accepting as many workers as were expected (those with the highest signals). The reverse is partially true and the firm would lower the threshold to accept more workers but only up to \( \bar{\theta}^{min} \); where \( \bar{\theta}^{min} \) is the lowest signal such that the \( IC.H^g \) constraint is still satisfied, i.e. \( V^{h,c} = V^s \).

### 3 Applications of the model

This section offers three applications of the model in the fields of human capital accumulation, firms’ optimal screening and income distribution. These applications use three specifications of the model.

The first specification is called Firm’s Selection\(^{18}\) (FS). In this specification the ex-post signals are excluded and, thus, performance pay does not exist. This specification resembles a signalling

\(^{17}\)By Jensen’s inequality as the utility function is concave.

\(^{18}\)Because the firm ‘selects’ the workers’ task via the hiring threshold.
model, partially mimicking the one developed by Moro & Norman (2004), even though I add concave utility and allow wages to be different from marginal products. The second specification, called Self-Selection\(^{19}\) (SS), assumes that the ex-ante signals are not informative. The SS specification is close in spirit to Rothschild & Stiglitz (1977), as firms choose from a menu of contracts to enforce separation. The third specification is called Double Screening (DS); it combines both types of signals.

The (DS) specification is explained in depth in section 2. Appendices B and C show how the specifications SS and FS are solved. A summary of the main features is offered in Table 1.

Table 1: Different model specifications

<table>
<thead>
<tr>
<th>Model</th>
<th>Ex-ante Signals</th>
<th>Ex-post signals</th>
</tr>
</thead>
<tbody>
<tr>
<td>(FS) Firm-Selection</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>(SS) Self-Selection</td>
<td>✗</td>
<td>✔</td>
</tr>
<tr>
<td>(DS) Double Screening</td>
<td>✔</td>
<td>✔</td>
</tr>
</tbody>
</table>

3.1 Welfare and income distributions

The objective of this section is to stress the role that performance pay may play on the US income distribution. My strategy for understanding the effect of performance pay is to isolate it. Since performance pay arises in the model from the existence of ex-post signals I use all three specifications of the model; one with only ex-ante signals, another one where only ex-post signals are informative and the last one where I use both screening mechanisms. Apart from the just described differences the economic environment is identical.

Simulations: To better understand the importance of allowing for both, ex-ante and ex-post signals, I show the implied income and welfare distribution of each specification. In particular I simulate 5000 worker histories for all three specifications, SS, FS and DS with the parameters shown in table 2 and 3. Parameters in table 2 are common while the ones in table 3 belong to each of the specifications.

Table 2: Common parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta)</td>
<td>0.6</td>
<td>(\alpha)</td>
<td>0.9</td>
</tr>
<tr>
<td>(\epsilon)</td>
<td>0.99</td>
<td>(\mu_\phi)</td>
<td>0.8</td>
</tr>
<tr>
<td>(A)</td>
<td>10</td>
<td>(\sigma_{\phi})</td>
<td>0.8</td>
</tr>
<tr>
<td>(A_h)</td>
<td>1</td>
<td>(\phi)</td>
<td>0</td>
</tr>
<tr>
<td>(A_l)</td>
<td>0.2</td>
<td>(\phi^l)</td>
<td>1.6</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{19}\)The name arises as the screening mechanism is enabled by workers self-selecting themselves out of the complex task.
Double Screening in the Labour Market

In the FS specification there are no ex-post signals and, therefore, we do not define the distribution that they come from. Also note that I assume the ex-ante signals not to be informative in the SS case. This is done by assuming that the signals of both, high and low skilled workers come from the same distribution.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>FS</th>
<th>SS</th>
<th>DS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_\eta$</td>
<td>-</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>$\sigma_\eta$</td>
<td>-</td>
<td>0.4</td>
<td>0.4</td>
</tr>
<tr>
<td>$\eta^h$</td>
<td>-</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\eta^l$</td>
<td>-</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\mu^h$</td>
<td>2.098</td>
<td>0</td>
<td>2.098</td>
</tr>
<tr>
<td>$\mu^l$</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3: Model specific parameters.

Income distributions: Income distributions arising from the 3 specifications are shown in figure 1. Technical details on how income is computed can be found in Appendix D.

Figure 1: Income distribution

First let us compare the FS and SS specifications as it is easier to understand separately the implications of each screening method on the income distribution. On the one hand, in the SS specification, performance pay creates an economy where some individuals earn much more than the rest. On the other hand, the FS specification creates a much smoother income distribution. The rationale is as follows, in the FS specification payments to complex task workers exclusively depend on the ex-ante wage, since this wage depends on a continuum of ex-ante signals wages are smoothed out. Meanwhile, in the SS specification we find more income dispersion since every complex task worker earns the same ex-ante wage and, it is performance pay that makes the difference; those who obtain it are located in the right tail of the distribution while the unlucky deceivers plus the simple task workers are found in
Double Screening in the Labour Market

The double screening specification DS combines features from both FS and SS. In the income distribution that arises from the DS specification, we see, first, a smoother distribution than in the SS specification, and, second, the creation of a hump on the left tail of the distribution. We can say that a more accurate ex-post signal (i.e. monitoring device) tends to increase income dispersion. This is not surprising since the more information is revealed about the workers’ skill level, the more accurately the contracts offered would reflect the true value of a worker.

Welfare distributions: The effects of double screening can also be seen in terms of welfare. High skilled workers would be better off if there were perfect information in the economy. This is shown in figure 2.2 where I show the welfare distribution of two economies, with and without perfect information, sorted by the workers’ cost of investment. I see that under perfect information high skilled workers benefit while low skilled lose out, increasing welfare dispersion.

Figure 2: Assymmetric vs perfect information: Utility distribution

![Utility distribution graph](image)

A more general comparison of welfare among the models is shown in figure 1.4, with the welfare distributions from all three specifications sorted in descending order of utility. It is remarkable that the welfare distribution from the DS specification is not skewed like the income one. The reason can be found in the contracts’ shape. Since workers are budget constrained, performance pay puts more weight on the right tail of the income distribution than on the welfare distribution one.

Empirical validity: Sala I Martín (2006) shows the US income distribution for the last 30 years. The evolution is characterised by the appearance of a hump in the left tail of the distribution as shown in Figure 2.3. The hump, in the model, is the result of an increase in the informativeness of the ex-post signals. I argue that in the US an increase in performance-related pay might have occured, perhaps
because of an improvement in the quality of the ex-post signals (better monitoring technologies, easier implementation due to the nature of jobs) or due to firms now using this reward system.

This story is, first of all, consistent with Béna"{b}ou & Tirole’s (2013) research, where it is argued that performance-related payments have increased since the 1970’s. It also goes along the same lines as Altonji & Pierret’s (2001) theory of employers’ learning which suggests the existence of double screening.

3.2 Firms’ screening usage

One of the main features of the DS model is the firm’s ability to choose between two screening mechanisms. The double screening system makes use of the hiring thresholds and the contracts. The thresholds screen the workers thanks to the stochastic dominance of the signals obtained by the high skilled workers and through the posterior probability of investing. Since the posterior $p(\theta)$ increases monotonically in $\theta$ it harms the low skilled more than the high skilled therefore allowing the firm to set a relatively a lower hiring threshold than it would otherwise.$^{20}$

On the other hand the rationale on how to use the performance pay is simple, the firm will set a relatively high performance pay with a simple task wage that is also high so as to disincentivise deception by low skilled workers.

Figure 2.4 shows the reaction function of the workers to a range of firms’ initial beliefs for two of the specifications discussed, double screening DS and Firms’ Selection FS. The steady state is found where the reaction function meets the 45 degree line. We observe that double screening is associated with higher investment rates both, in the steady state and out of the steady state (short run equilibrium). This is because the economy rewards more accurately the skill when there are two screening mechanisms and the incentive to invest to grows.

$^{20}$For instance if the ex-ante wage were flat $w^a(\theta) = w^c$. 

Figure 3: Welfare distribution.
To show the firm’s screening choice I simulate an economy for a vector of signals’ quality. Signals quality is defined as $S_q = 1 - \text{OVC}$, where OVC is the overlapping coefficient between the distributions $F_h$ and $F_l$. When $S_q$ is 0 firms cannot gain any valuable information when they meet the workers. When it is 1 the correlation between the investment decision and the signal is perfect. With regards to the parameters of the signal distribution an OVC equal to 1 is associated with $\mu_h = 0$ and an OVC of 0 corresponds to $\mu_h = \infty$.

Figure 2.5 shows shares of workers sent to perform simple tasks for each screening method. To build the graph I fix the probabilities of obtaining a ‘pass’ signal using the ones shown in table 2.3. As expected, the more informational the ex-ante signals are, the more the firm relies on using hiring thresholds rather than on provoking low skilled workers to reveal their type through contracts (self-selection). At some point, when the ex-ante signals’ quality is relatively high, the firm no longer cares about potential foolers as their number is low and the cost of disincentivize them is higher than the marginal profit obtained by ensuring a better pool of workers doing complex tasks.

In Figure 2.6 I also show the wages offered by the firm because they offer valuable information on how firms set the contracts. Because the quality of the signals alters the pattern of human capital accumulation the interpretation of this figure is more subtle.

On the left hand side of the figure I expect the firm to set contracts that put off deceivers. Such contracts imply a relatively high simple task wage and rising performance pay. I do not see the latter though, because the number of investors is also increasing and so the firm’s need for a good pool of complex task workers is alleviated.

At some point in the middle of the graph the firm switches off screening through self-selection and we see $w^s$ decreasing for two reasons: First the firm does not care about deceivers self-selecting.

---

21This is further expanded in section 2.4.
themselves to do simple tasks. For the second reason remember that firms pay a non-zero simple task wage as insurance for investors who could be wrongly placed in the simple task. As the information carried by the ex-ante signals improve the insurance motive vanishes and so the need for a higher $w^s$.

Finally the complex task base payment, $w^c$, (the maximum a worker could earn provided $p(\theta, \tilde{\pi}) = 1$) decreases sharply. This occurs because as the informational quality of the signals improves the firm does not need to pay such a high $w^c$ to satisfy the IC.H constraint because $p(\theta)$ is doing a better job of identifying high skilled workers. An implication of this result is that economies -or sectors within an economy- with bad quality ex-ante signals should show smoother income distributions as it is difficult to reward ability and the ex-ante payment is going to be almost equally distributed among all the workers assigned to the complex task.

### 3.3 Human capital accumulation

I develop a benevolent planner\(^\text{22}\) to test whether the market outcome under perfect competition is efficient or not. The planner’s objective function is the social welfare function and the planner has the power to set wages and hiring thresholds. Welfare is defined as aggregate welfare, assuming every individual in society has the same weight (utilitarian criterion).

Notice the planner knows his decisions affect the investment rate in human capital. This planner shall be regarded as a second best though, as he cannot observe the workers’ investment choices and therefore is constrained. I prefer to omit the first best result as the informational requirements seem unreasonable and far from reality. The constrained planner, thus, also faces the problem of asymmetric information.

The differences in screening choice are obvious with respect to the perfect competition case. The

\(^{22}\)Details of the problem can be found in appendix E.
The planner keeps using self-selection irrespective of the signals’ quality. This can be seen in figure 2.7.

The main difference between the planner and the versions of the model under perfect competition is in the benefit of investing to workers. The analysis follows Figure 2.8, where I see the steady state investment rate, \( \pi_{ss} \), under the constrained planner and under perfect competition for a vector of ex-ante signal quality. On the one hand high ex-ante signal quality (to the right of the figure) makes firms under perfect competition set contracts that create a broader gap between the expected utilities of investing and non-investing, thus, creating a bigger benefit to investing than under the planner’s rule. On the other hand when the ex-ante signal quality is low (to the left of Figure 2.8) the planner makes a
more intense usage of contracts to induce self-selection than do firms under perfect competition. This creates a higher benefit to investing and therefore investment rates rise in the case of the constrained planner.

As a conclusion I can say that perfect competition entails under-investment when the ex-ante signal quality is low and over-investment when the ex-ante signal quality is high. Only by chance could the level of human capital under perfect competition be optimal.

Figure 8: Investment rate. Perfect competition vs. constrained planner.

The role of market structure: The lack of efficiency found in the level of human capital under perfect competition opens the door to considering the possibility that different market structures are better in terms of welfare. In this regard I analyse the market structure that most closely fits the constrained planner; a monopsony.

A monopsony is made of a sole employer with market power to set wages and hiring thresholds without any threats from any other firm. More importantly it also internalises the effect of wages on the incentives to invest (just like the planner discussed). The existence of two differences between the constrained planner and the monopsonist are likely, though, to influence the analysis. First the monopsonist is a profit, not a welfare, maximiser. Second, positive levels of profit monopsonies are known to exist in monopsonies that are not given to workers.

To make the analysis I build a zero profit monopsonist as an upper bound for welfare. The differences in human capital accumulation between these two resemble the one shown in Figure 2.8 and it is omitted. It is more interesting to show the level of welfare achieved under the monopsony and in an economy with firms under perfect competition. I see in Figure 2.9 that under a monopsony, the economy always achieves higher welfare than under perfect competition. Obviously if I shift up

---

23 Note the planner’s feasibility condition is a zero profit constraint.

24 Please refer to appendix G for detailed definition of welfare.
Figure 9: Welfare, perfect competition vs. monopsony.

the level of profit that the monopsony can obtain, the set of signal quality values where a monopsony is more efficient becomes smaller.

As a rule of thumb, and provided a monopsonist always makes some level of profit, I could say a monopsony is preferable in terms of welfare to perfect competition in economies where signal quality is relatively extreme (either high or low).

**Taxation:** Modifying an economy’s market structure might be complex and not have such desirable consequences. Given that we know what the optimal policy for investment in human capital is, if I knew how informative the ex-ante and ex-post signals are (i.e. the informational structure of the labour market) I could tax the firms’ payroll to obtain an outcome similar to the constrained planner.

Taxes should depend on the case. If there is under-investment in human capital the incentive to invest should grow. Subsidies towards performance-related payments or a lower minimum wage for the simple tasks would work in the right direction.

The contrary holds if it is over-investment in human capital that exists. Heavier taxation against performance pay or higher minimum wages would shrink the incentive to invest so that the level of human capital moves closer to the optimal.

4 Conclusion

I build a model of adverse selection where firms have two screening mechanisms. In terms of labour market development the work undertaken in this paper goes further than the treatment offered by Coate & Loury (1993) or Moro & Norman (2004), where firms can only use the workers’ signals to allocate them either to the complex or the simple task. In addition, I propose a different way to solve for
the equilibrium, where wages are not necessarily the marginal products. This approach suggests firms only reward workers insofar as this is needed to incentivise workers to undertake a costly investment in human capital.

In this paper I use the model to focus on the effects of double screening on labour market and economy-wide outcomes. I show that the skill composition of the workers employed in the complex task improves when firms take advantage of the two screening mechanism. I also offer insights on how the screening usage depends on the informativeness of each signal method.

Finally I use the model for two more purposes, as an explanation for the changes in the shape of the US income distribution, and for testing whether or not perfect competition offers an optimal accumulation of human capital. I find for the former that double screening might play an important role in explaining changes to the income distribution due to an increase in performance-related payments. As for the latter I show not only that perfect competition implies a non-optimal level of human capital, but also that competition-restrictive market structures such a monopsony might offer better results in terms of welfare. Finally I hint at how taxes might approximate a monopsony’s human capital level without changing the market structure.
References


Appendix A: Joint CDF and expected value

In section 4, as part of the model solution, I need to calculate the joint CDF of two disjoint truncated normal random variables and the expected value of one of the variables. With regards to the former the joint CDF is defined as

$$F_D = \int_{\theta_{\min}}^{\theta_{\max}} \int_{\eta_{\min}(\theta)}^{\eta_{\max}} f_{\theta}(\theta, \tilde{\theta}, \infty) f_{\eta}(\eta, \eta^l, \eta^h) d\eta d\theta,$$

but a word must be said about the integration limits. The upper bounds of both integrals are given by the upper limits of the variables’ truncated normal distributions i.e. \( \theta_{\max} = \infty \), \( \eta_{\max} = \eta^h \). On the contrary the calculations for the lower bounds are more subtle. On the one hand the lower bound of the outer integral \( \theta_{\min} \) might depend on the upper bound of the inner integral \( \eta_{\max} \) in the following way:

$$\theta_{\min} = \max \left\{ \frac{V^s}{(V^l,c(\eta_{\max}))^{-1} \tilde{\theta}} \right\} \quad (32)$$

On the other hand the lower bound of the inner integral depends on the outer variable. The idea is that I must ensure the pairs \( \theta, \eta \) satisfy the incentive compatibility constraint for a low skilled to accept a complex task.

$$\eta_{\min} = \frac{V^s}{(V^l,c(\theta))^{-1}} \quad (33)$$

In addition I calculate the expected probability of fooling given that a worker deceives the firm \( E[\eta|D = 1] \). This is done as follows,

$$E[\eta|D = 1] = \int_{\theta_{\min}}^{\theta_{\max}} \int_{\eta_{\min}(\theta)}^{\eta_{\max}} \eta f_{\theta}(\theta, \theta_{\min}, \infty) f_{\eta}(\eta, \eta_{\min}, \eta^h) d\eta d\theta,$$

where it has to be noted, in addition to what was explained before, that the lower truncation of the distributions follows the lower bounds of the respective integrals. This is because I must ensure the volume adds up to one. In addition the order of integration (first \( \theta \), then \( \eta \)) is not trivial; since the signal occurs first in time, I need to know what probabilities of fooling are IC and not vice-versa.

Appendix B: Model with only Firm-Selection

This is, in essence, a signaling model with human capital accumulation where workers pay to become high skilled and obtain an, on average, higher signal. Firms then select those with the highest signals to perform complex task. In the context of the double screening model this case is constructed by the deletion of the detection technology. This has two side effects, first, no worker invited to perform complex tasks would turn the offer down making \( F_D = 1 \). In addition the firm has no incentive to pay money ex-post, thus for simplicity I assume all money is paid ex-ante and \( w^p \) is erased from the
problem.

To solve for the equilibrium I maximize the representative firm’s profit function. Given outside options \( \{ \bar{U}^H, \bar{U}^L \} \) and a belief with respect the workers’ investment rate \( \bar{\pi} \), the firm solves,

\[
\max_{\bar{\theta}, w^c, w^s} P = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w) \tag{35}
\]

where the output is given by \( Y = A (\alpha C^{\gamma} + (1 - \alpha)S^{\gamma})^{\frac{1}{\gamma}} \) and the efficiency units are defined as \( C = A_h \bar{\pi}(1 - F_h) + A_l (1 - \bar{\pi})(1 - F_l) \) and \( S = \bar{\pi} F_h + (1 - \bar{\pi}) F_l \). The payroll function also suffer some modifications,

\[
\omega(\bar{\theta}, w) = w^c (E[p(\theta)|I=1]H^c + E[p(\theta)|I=0]L^c) - w^s N^s \tag{36}
\]

Now I turn to the constraints of the problem. The IR constrains take the following form:

\[
\bar{U}^H \leq [1 - F_h(\bar{\theta})]V_{E}^{h,c} + F_h(\bar{\theta})V^s \quad \text{IR.H} \tag{37}
\]

\[
\bar{U}^L \leq [1 - F_l(\bar{\theta})]V_{E}^{l,c} + F_l(\bar{\theta})V^s \quad \text{IR.L} \tag{38}
\]

As for the IC constrains, the IC.L is never satisfied. This is because \( V^{h,c} \geq V^s \) implies \( V^{l,c} \geq V^s \).

\[
V^{h,c} \geq V^s \quad \text{IC.H} \tag{39}
\]

\[
V^s \geq V^{l,c}(\theta, \eta), \quad \text{broken} \forall(\eta, \theta) \quad \text{IC.L} \tag{40}
\]

The workers’ investment rate it follows the same logic as in the main document. As for the deceiving decisions every low skilled invited to do complex tasks accepts them thus making \( F_D = 1 \) and \( E[\eta|D=1] = \mu_e ta \).

**Appendix C: Model with only Self-Selection**

This appendix describes a model where firms set a menu of contracts so that some low skilled workers reveal their type and opt to take on simple tasks. To achieve its aim firms condition part of the wage on actual productivity that is paid ex-post. Firms, therefore, are able to find out, with a certain degree of accuracy, workers’ productivity. How they find out is explained in section 2.2.2 of this paper.

There are two ways to build this specification that yield the same result. One is assuming signals still exist but are meaningless (i.e. both signals cdf’s have the same distribution), in which case every worker has a posterior probability of having invested of 0.5. The second one, and the one I follow, simplifies even further the problem and assumes signals do not exist at all. This leaves the representative firm to invite a random sample of the applicants for the complex tasks.
**Profit:** Given outside options \( \{\bar{U}^{H}, \bar{U}^{L}\} \) and a belief with respect the workers’ investment rate \( \bar{\pi} \), the firm solves,

\[
\text{Max}_{\bar{\theta}, w^{c}, w^{s}} \quad P = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w) \tag{41}
\]

where the output is given by \( Y = A (\alpha C^{\gamma} + (1 - \alpha)S^{\gamma})^{\frac{1}{\gamma}} \) and the efficiency units are defined as \( C = A_0 h\bar{\pi} + A_l \bar{\theta}(1 - \bar{\pi})F_D \) and \( S = \bar{\pi}(1 - \bar{\theta}) + (1 - \bar{\pi})(1 - \bar{\theta}) + \bar{\theta}(1 - \bar{\pi})[1 - F_D] \). The payroll function also suffer some modifications,

\[
\omega(\bar{\theta}, w) = w^{c}N^{c} - w^{p}(H^{c} + E[\eta|D = 1]L^{c}) - w^{s}N^{s} \tag{42}
\]

**Constraints:** Finally I need to define 2 individual rationality and 2 incentive compatibility constraints. With regards to the Individual Rationality constraints I have,

\[
\bar{U}^{H} \leq \bar{\theta}V^{h,c} + (1 - \bar{\theta})V^{s} \quad \text{IR.H} \tag{43}
\]

\[
\bar{U}^{L} \leq \bar{\theta}V^{l,c}[1 - F_\eta(\bar{\eta})] + \bar{\theta}V^{s}F_\eta(\bar{\eta}) + (1 - \bar{\theta})V^{s} \quad \text{IR.L} \tag{44}
\]

A key difference arises on the IC.L constraint compared to the FA model. The IC.L constraint is not broken for a certain subset of the low skilled workers invited to perform complex tasks and, therefore, some will refuse to carry them out and will opt for simple tasks.

\[
V^{h,c} \geq V^{s} \quad \text{IC.H} \tag{45}
\]

\[
V^{s} = V^{l,c}(\bar{\eta}) \quad \text{IC.L} \tag{46}
\]

The IC.L holds true for all \( \eta < \bar{\eta} \) meaning there will be a semi-separating equilibrium. Workers investment decisions are not modified. As for the deceiving decision all workers with a fooling probability higher than the deceiving threshold \( \bar{\eta} \) will deceive. The threshold is calculated as follows,

\[
\bar{\eta} = \left( \frac{V^{s}}{V^{l,c}} \right)^{-1} \tag{47}
\]

provided the inverse of \( V^{l,c} \) exists.

**Appendix D: Model with perfect information**

This appendix shows how to solve the model presented in section 2 if firms were able to observe with certainty and ex-ante whether a worker has invested in his human capital or not. That clears out the informational asymmetry from the model and also a source of inefficiency.

The maximization problem that the representative firm faces varies slightly. Now the firm does
not worry about low skilled workers potentially doing complex tasks as there are no deceivers. In addition I simplify the model further by eliminating the performance payment \( w^p \), because the firm knows ex-ante the performance of every worker and there is no reason to withhold payments.

**Profit:** Given outside options \( \{ \bar{U}^H, \bar{U}^L \} \) and a belief with respect the workers’ investment rate \( \bar{\pi} \), the firm solves,

\[
\max_{\theta, w^c, w^s} P = Y(\theta, w) - \omega(\bar{\theta}, w) \tag{48}
\]

where the output is given by \( Y = A(\alpha C^\gamma + (1 - \alpha)S^\gamma)^{1/\gamma} \) and the efficiency units are defined as \( C = A_h \bar{\pi} \bar{\theta} \) and \( S = \bar{\pi}(1 - \bar{\theta}) + (1 - \bar{\pi}) \). The payroll simplifies to just

\[
\omega(\bar{\theta}, w) = w^c N^c - w^s N^s \tag{49}
\]

**Constraints** Both Incentive Compatibility constraint are now missing because workers cannot pretend to be a different type. The maximization problem of the firm is left with the 2 Individual rationality constraints.

\[
\bar{U}^H \leq V^{h,c} \tag{50}
\]

\[
\bar{U}^L \leq V^s \tag{51}
\]

**Appendix E: Constrained planner**

This appendix show the calculations undertaken to obtain the solution of a benevolent planner constrained to take into account the lack of knowledge about the workers’ type. This planner cares about the society’s welfare and maximizes total utility by choosing wages a hiring threshold and the investment rate of the society. The latter can be done because the planner controls the incentive to invest and knows the shape of the investment cost cdf.

\[
\max_{\bar{\theta}, w^c, w^p, w^s, \pi} \quad W = W^{h,c} H^c + W^{l,c} L^c + W^s N^s - \int_{\phi} F(\phi) d\phi \tag{52}
\]

where total welfare \( W \) is constructed as the addition of the high skilled and low skilled individual welfare\(^{27}\) weighted by their size in the society minus the total cost of investment. The total cost of investment can be added altogether because it enters the individual utility functions linearly. The integral of the costs goes from the lower bound of the cost distribution to the incentive to invest, which is the highest cost a worker would pay.

First the planner faces a feasibility constraint. The planner must ensure that the earnings\(^{28}\) of the

---

\(^{27}\)Calculations for the welfare measure is shown in appendix G.

\(^{28}\)Calculations for the earnings measure is shown in appendix G.
workers do not exceed what it is produced.

$$H^c E^{h,c} + L^c E^{l,c} + N^s E^s \leq Y$$ (53)

where output is given by $Y = A (\alpha C^\gamma + (1 - \alpha)S^\gamma)^{\frac{1}{\gamma}}$ and the efficiency units are defined exactly as in the DS model. Then I just add a set of IC constraints because the IR’s are substituted by equation E.5,

$$V^{h,c} \geq V^s \quad \text{IC.H}$$ (54)

$$V^s \geq V^{l,c}(\eta, \theta) \quad \text{for some } \theta, \eta \quad \text{IC.L}$$ (55)

The last constraint I add to the model states that the amount of investors is equal to,

$$\pi = F_\phi(\xi)$$ (56)

Where the incentive to invest, $\xi$, is defined in section 2.4.2. Finally I also refer the reader to section 2.4.2 to check how the deceiving decisions are made.

**Appendix F: Workers’ utility maximization problem**

In the main body of the thesis I state the workers’ utility function and its liquidity constraint. This section specify how the workers maximize their utility for a given contract that pays some money ex-ante and some money ex-post, perhaps with uncertainty.

**Utility under certain payments**

The utility obtained by high skilled workers in the complex task, with contract $\{w^a = p(\theta)w^c, w^o = w^p\}$, and any worker in the simple task with contract $\{w^a = w^s, w^o = 0\}$, is known with certainty. To solve for the optimal allocation, workers maximize the following function:

$$\text{Max } U(c^a, c^o) = \frac{c^a^{1-\epsilon} - 1}{1 - \epsilon} + \beta \frac{c^o^{1-\epsilon} - 1}{1 - \epsilon}$$ (57)

subject to

$$c^a \leq w^a$$ (58)

$$c^a + c^o = w^a + w^o$$ (59)
The utility function evaluated at the optimum takes two values:

\[
V = \begin{cases} 
  w^a_1 - \epsilon - 1 \frac{w^a}{1+\beta} - 1 & \text{if } w^a \leq w^a_1 + w^o_1 \\
  \left(\frac{w^a_1 + w^o_1}{1+\beta}\right)^{1-\epsilon} - 1 + \beta \left(\frac{w^a_1 + w^o_1}{1+\beta}\right)^{1-\epsilon} - 1 & \text{if } w^a > w^a_1 + w^o_1 
\end{cases}
\]  

(60)

Utility under uncertain payments

If the performance pay is uncertain, as it occurs to the low skilled workers when joining the complex task, some precautionary savings will be taken by the workers in order to insure themselves against the risk of not having any consumption good when they are old. The maximization program for a given contract \(\{w^a = p(\theta)w^c, w^o = w^p, C\}\) is,

\[
\text{Max } U(c^a, c^{o,h}, c^{o,l}) = c^{a^{1-\epsilon}} - 1 + \beta \left[ \eta c^{o,h^{1-\epsilon}} - 1 + (1 - \eta) c^{o,l^{1-\epsilon}} - 1 \right]
\]  

(61)

subject to

\[
c^{o,h} = w^a + w^o - c^a
\]  

(62)

\[
c^{o,l} = w^a - c^a
\]  

(63)

The value function analytical expression is not easy to find so I provide, on the one hand, the first order condition that gives the optimum:

\[
c^{a^{1-\epsilon}} = \beta \left( \zeta (w^a + w^o - c^a)^{-\epsilon} + (1 - \zeta)(w^a - c^a)^{-\epsilon} \right)
\]  

(64)

Once \(c^a\) is known I can pin down future consumption using the constraints of the problem. The value function is the utility function evaluated at the optimum \(V = U(c^{a*}, c^{o,h*}, c^{o,l*})\).

Appendix G: Welfare and income measures

This appendix is written to define the measures of welfare and earnings used in section 3 and appendix D. First I define welfare by workers’ occupation and type (i.e. high skilled in complex task). Then I turn to do the same for the earnings.

Welfare

Using some previously defined objects I define average welfare for high skilled and low skilled workers, at both, the simple and the complex task. This is done for the double screening specification. Welfare, net of investment costs, for high skilled workers doing the complex task is:

\[
W^{h,c} = \int_\theta^\infty V^{h,c}(p(\theta)w^c, w^p) f_h(\theta, \bar{\theta}, \infty) d\theta
\]  

(65)
For the non-investors who deceive the firm I need to take into account the bias produced by the self-selection\textsuperscript{29} of workers:

\[
W^{l,c} = \int_{\theta_{\text{min}}}^{\infty} \int_{\eta_{\text{min}}}^{1} V^{l,c}(p(\theta)w^c, w^p, \eta) f_1(\theta, \theta_{\text{min}}, \infty) f_\eta(\eta, \eta_{\text{min}}, \eta^h) d\eta d\theta
\]  
(66)

The average welfare obtained by high skilled doing simple tasks and low skilled doing simple task is given by:

\[
W^{h,s} = V^s
\]  
(67)

And total welfare in the economy is defined by the addition of all the partial welfare measures minus the investment costs incurred by the high skilled.

\[
W = H^c W^{h,c} + L^c W^{l,c} + N^s W^s - \int_{\phi}^{\xi} F_\phi d\phi
\]  
(68)

**Earnings**

I do a similar exercise to compute workers’ aggregate earnings. For high skilled workers doing complex task, earnings are given by:

\[
E^{h,c} = \int_{\theta}^{\infty} p(\theta)w^c f_h(\theta, \theta_{\text{min}}, \infty) d\theta + w^p
\]  
(69)

For the non-investors who deceive the firm I need to take into account the bias produced by the self-selection\textsuperscript{30} of workers:

\[
E^{l,c} = \int_{\theta_{\text{min}}}^{\infty} \int_{\eta_{\text{min}}}^{\eta^h} \eta w^p f_1 f_\eta d\eta d\theta + \int_{\eta_{\text{min}}}^{\infty} \int_{\theta_{\text{min}}}^{\infty} w^c p(\theta) f_1 f_\eta d\eta d\theta
\]  
(70)

The earnings obtained by high skilled and low skilled doing simple tasks is given by:

\[
E^{h,s} = w^s
\]  
(71)

And the total earnings in the economy are defined by the addition of all the partial earnings measures weighted by the share of each group in the population,

\[
E = H^c E^{h,c} + L^c E^{l,c} + N^s E^s
\]  
(72)

\textsuperscript{29}Refer to Appendix A for information about the bounds of this integral.

\textsuperscript{30}Refer to Appendix A for information about the bounds of this integral.