

University of Southampton Research Repository ePrints Soton

Copyright © and Moral Rights for this thesis are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given e.g.

AUTHOR (year of submission) "Full thesis title", University of Southampton, name of the University School or Department, PhD Thesis, pagination

ADVERSE SELECTION AND RACE
IN THE LABOUR MARKET

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ECONOMICS
AND THE COMMITTEE ON GRADUATE STUDIES
OF SOUTHAMPTON UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Luis Pinedo Caro

June 2014

© Copyright by Luis Pinedo Caro 2014
All Rights Reserved

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(John Knowles) Principal Advisor

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Árpád Ábrahám, External Examiner
European University Institute)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(Thomas Gall, Internal Examiner
University of Southampton)

Approved for the University Committee on Graduate Studies

Preface

This thesis is the work of 4 years at the University of Southampton. As a person, I am both, curious and surprised, about issues like inequality, discrimination and education opportunities as a form of social lift. These, and other related issues are the ones you will find in this dissertation.

Acknowledgments

I will never be able to say thanks enough times to John Knowles, my supervisor. He did not just supervised my research, he taught me to be a researcher myself, which is priceless.

In addition I was lucky enough to meet my internal examiner, Thomas Gall. I want to thank him for his effort and generosity to make this document better.

I also want to thank the continuous effort (both, emotional and financial) my parents have put on giving me the best possible education.

Fianlly I have to acknowledge financial support from a 3 year grant awarded by the University of Southampton.

Contents

Preface	iv
Acknowledgments	v
Introduction	1
1 A model of double screening	4
1.1 Introduction	4
1.2 The model	5
1.2.1 Workers	6
1.2.2 Firms	7
1.3 Equilibrium	12
1.4 Solving for the equilibrium	13
1.4.1 The firm's problem	13
1.4.2 Workers' problems.	15
1.4.3 Outside options	17
1.4.4 Market clearing	18
1.5 Conclusion	18
2 Applications of the model with 1 group	20
2.1 Introduction	20
2.2 Welfare and income distributions	21
2.3 Firms' screening usage	25
2.4 Human capital accumulation	27
2.5 Conclusion	31
3 Applications of the model with 2 groups	32
3.1 Introduction	32
3.2 Discriminatory beliefs	35

3.2.1	Without contract freedom	36
3.2.2	With contract freedom	38
3.3	The case of the US	41
3.3.1	The racial gap in the US	42
3.3.2	The effect of schooling differences	44
3.3.3	The effect of banning race from hiring	46
3.4	Conclusion	48
Conclusion		51
A Joint CDF and expected value		52
B Model with only Firm-Selection		54
C Model with only Self-Selection		56
D Model with perfect information		58
E Constrained planner		60
F Workers' utility maximization problem.		62
F.1	Utility under certain payments	62
F.2	Utility under uncertain payments	63
G Welfare and income measures		64
G.1	Welfare	64
G.2	Earnings	65
H Identification of outside options for competitive equilibria		66
H.1	Model with one group	66
H.2	Model with 2 groups	66
Bibliography		67

List of Tables

2.1	Different model specifications	22
2.2	<i>Common parameters.</i>	22
2.3	<i>Model specific parameters.</i>	22
3.1	Steady state. Discriminatory steady-state.	46
3.2	<i>Preferences and technology</i>	47
3.3	<i>Detection tech., ex-ante signals dist. and investment costs.</i>	48

List of Figures

2.1	Income distribution	23
2.2	Assymmetric vs perfect information: Utility distribution	24
2.3	Welfare distribution.	24
2.4	Fixed points	26
2.5	Firm's usage of double screening.	27
2.6	Firm's usage of double screening.	28
2.7	Planner's usage of double screening.	28
2.8	Investment rate. Perfect competition vs. constrained planner.	29
2.9	Welfare, perfect competition vs. monopsony.	30
3.1	<i>Convergence. Double screening vs. Firm's Selection (Signaling)</i>	37
3.2	Investment reaction functions.	38
3.3	<i>Screening policy with different initial belief.</i>	39
3.4	Firms' wage policy with varying beliefs.	40
3.5	Black/White income gap by birth cohort	43
3.6	Beliefs converging to a non-colour-blind steady state.	45
3.7	Black and white's welfare.	46
3.8	Black/White simulated income ratio	49

Introduction

The original title of this thesis was ‘Amending the Racial Wage Gap, What Helps and What Does not?’. However the path to amending the racial gap suffered several diversions that extended the scope of this thesis and therefore its title.

The path starts where wages are paid, the labour market. The labour market is not a place where everybody knows everything about every other person; perfect information, as it is called in economics, does not exist. Employers, when choosing how to allocate responsibilities and how much to pay for each responsibility, are aware that they could be hiring the wrong person.

The consequence of imperfect information to firms is called *adverse selection* and leads to the skill composition of the hired sample being lower than desired by the firm. In fact, adverse selection might provoke, in extreme cases, a cessation of economic activity in a market. Fortunately economists have found ways to ameliorate this problem. In this thesis I examine two well-known solutions. Both solutions are related to how the firm can change its actions to minimise the impact of asymmetric information.

One of the solutions requires the existence of an education system. This solution is based on the expectation by firms that workers will send signals about their skill set whilst still in school. The best performers will each have a stronger signal and employers will be keener to hire them for positions of a more demanding nature. The solution just described is called *signaling* and it allows employers to allocate more demanding jobs to workers with better signals. I refer to this solution as firms screening the applicants’ education choices or signalling.

The second solution developed requires the output a worker produces to be ex-post observable -with some degree of accuracy- by the employer. When this is possible, firms change the wage structure by conditioning ex-post payments on actual productivity. The latter means underskilled workers self-select themselves out of the demanding jobs. Observable output, thus, allows the firm to pay piece-rate. In this thesis we refer to this solution as firms screening the workers’ occupational choices or in short self-selection.

Notice in the literature on this topic self-selection is named *screening*. We believe

this is unfortunate because both signalling and self-selection involve some sort of screening by the firm.

Chapter 1 of this thesis tackles the adverse selection problem. It presents a model of adverse selection where the two most common solutions to this problem, screening of educational and occupational choices, are put together creating a model of double screening.

Models incorporating both solutions are appealing as they more realistically reflect what is observed in data. Empirically, firms seem to condition initial wages on signals and later wages on variables correlated with actual productivity. This is known as employers' learning. I also assume employers in this model have a belief (which might be wrong) with regards to how many workers are able to perform demanding tasks properly. This assumption is added to observe the evolution of the belief over time and how long it takes to be equal to the actual share of skilled workers. In summary, chapter 1 lays down a model of the labour market which is close to the hypothesis of employers' learning.

Chapter 2, in turn, provides three applications of the double screening model with one group. First it explains how the possibility of double screening impacts the contracts that firms offer and the assignment of jobs. In addition I change the quality of the ex-ante signals and the accuracy of the firms' observations of workers' productivity. These changes allow us to better understand how the firm chooses the profit maximising contract.

Second, I use the double screening model to simulate income and welfare distributions. The income distribution is not immune to the screening mechanism chosen by the firms. Furthermore I link the availability of screening methods with the US actual income distribution and its evolution. In particular I think that the observed increase in piece-rate payments might be behind the changes to the US income distribution.

The last point of chapter is an analysis of the influence of market structure on human capital accumulation. Market structure refers to how many firms exist in a given market producing the same good. When applied to the labour market the word *producing* is replaced by *hiring*. The extremes of this classification are perfect competition, a situation where many employers co-exist, and a monopsony, where there exists only one employer.

The last chapter uses the the double screening model with two groups workers that can be recognized because of some physical characteristic like skin colour. The first part of this chapter allows the firms to use an applicant's physical characteristics when setting contracts and allocating tasks.

I show results from a simulation where one of the groups is believed by the firm to be inferior (in productivity terms). Two lessons emerge. Firstly, convergence in earnings and beliefs between the two groups is almost immediate, not taking more than 2 generations to occur. Secondly, the more powerful is the ability of the firm to observe performance the faster the above-mentioned convergence is. The reason for the faster convergence is the firms' ability to set group-specific hiring strategies. Firms condition wage more on the performance of the applicants who belong to the disfavoured group; that increases the incentives of workers in the disfavoured group to invest in human capital and thus closes the earnings and beliefs' gap faster.

The last part of chapter 3 contains applied research on the black/white income gap in the US. It starts by asking why the exercise carried out in the first part of the same chapter suggests quick earnings convergence between groups yet we do not see it in reality. We follow the literature on US racial history to note that black and whites do not enjoy the same family backgrounds and schooling opportunities. We adapt the model to include these differences by assuming, first, different investment cost distributions. Cost that is meant to represent family background. Second, we assume that investment in human capital rewards more whites than blacks. More rewarding in terms of productivity to perform demanding jobs and in terms of the schooling signal observed by employers. I show that this differences in opportunities can actually explain the persistence of the US racial gap.

Finally we scrutinise the role played by one important policy, the ban on using race when hiring, as stems from the Title VII of the US civil rights act of 1964. I show that the easier it is for firms to observe the actual productivity of workers the more damaging the ban for the disfavoured group is. This result, even though shocking at first, has a very simple rationale; banning the use of race when hiring reduces the incentives to invest for the disfavoured group. The latter implies a reduction in the next generation's investment rate and as a direct consequence on performance in the labour market. In particular I show that, over the course of three generations, the economy where the use of race is partially excluded from hiring decisions from being used in contracts closes the income gap by 15 percentage points while on economy where the use of race is allowed closes the gap by 28 percentage points.

Chapter 1

A model of double screening

1.1 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 chapter 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 chapter 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.

The main contribution of this chapter 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.

1.2 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 work developed in this chapter 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).

The model developed in this chapter 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 capital¹; 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 interactions. Then a definition of equilibrium is provided and a method for how to solve the equilibrium with one and two groups, respectively.

1.2.1 Workers

Demographic structure

Every period a unit continuum of workers is born. Workers belong to one of the two existent groups, $g \in \{W, B\}$. The size of each group is given by λ^W and λ^B with respect to groups W and B. $\lambda^W + \lambda^B = 1$. I assume that group identity is easily observable and public knowledge.

Each generation of workers lives for three periods. They represent the workers' youth, adulthood and retirement. Workers are endowed with one unit of time which is inelastically supplied in their adulthood.

Types and skills

The type of the workers is given by the cost of investing in human capital. This disutility, denoted ϕ , is randomly drawn from the distribution F_ϕ^g whose shape might depend on the group.

When young, a worker willing to face the cost ϕ and invests in his human capital 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 each group's high skilled individuals by π^g .

Preferences

Workers derive utility from consuming during their adulthood and their retirement.³ Their preferences over consumption are summarised by an isoelastic utility function

¹See Becker (1975) or Spence (1973)

²This is assumed for simplicity. It does not affect qualitatively the results shown.

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

that weights consumption during adulthood c^a and old age c^o ,

$$u(c^a, c^o) = \frac{c^{a^{1-\epsilon}} - 1}{1 - \epsilon} + \beta \frac{c^{o^{1-\epsilon}} - 1}{1 - \epsilon} - \phi, \quad (1.1)$$

where ϵ is the coefficient of relative risk aversion and β 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:

$$V(c^{a,*}, c^{o,*}) = \max_{c^a, c^o} u(w^a, w^o) \quad (1.2)$$

Outside options

High skilled and low skilled workers have outside options denoted by $\{\bar{U}^{H,g}, \bar{U}^{L,g}\}$ for $g \in \{w, b\}$. Outside options reflect the utility that these workers could obtain if they put their labour back in the market.

1.2.2 Firms

There are two infinitely-lived firms competing á la Bertrand. These firms are myopic⁴ because they only take into account current profit when deciding their actions. At time 0, I assume that firms hold beliefs on the share of high skilled workers in each group $\{\tilde{\pi}_0^w, \tilde{\pi}_0^b\}$.

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 $\{Y : \mathbb{R}_+^2 \rightarrow \mathbb{R}\}$ into output,

$$Y = A (\alpha C^\gamma + (1 - \alpha) S^\gamma)^{1/\gamma}, \quad (1.3)$$

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

$$\lim_{C \rightarrow \infty} \frac{\partial Y}{\partial C} = \lim_{S \rightarrow \infty} \frac{\partial Y}{\partial S} = 0, \quad \text{and} \quad (1.4)$$

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

$$\lim_{C \rightarrow 0} \frac{\partial Y}{\partial C} = \lim_{S \rightarrow 0} \frac{\partial Y}{\partial S} = \infty. \quad (1.5)$$

The complex task is done more efficiently by high skilled workers (with productivity A_h^g) than by low skilled workers (with productivity A_ℓ), with $A_h^g > A_\ell > 0$ for $g \in \{W, B\}$. 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.

$$C = \sum_{g \in \{W, B\}} A_h^g H^{c,g} + A_\ell L^c \quad (1.6)$$

$$S = H^s + L^s \quad (1.7)$$

Information

Firms have a belief, $\tilde{\pi}^g$, -that might be wrong- about the proportion of high skilled workers in each group, yet they do not know, within a group, who is high skilled and who is not. This information asymmetry generates a problem of 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 θ . This signal is randomly drawn from one of two distributions F_h^g or F_ℓ . High skilled workers send a signal drawn from F_h^g while low skilled workers draw it from F_ℓ , with F_h^g 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.

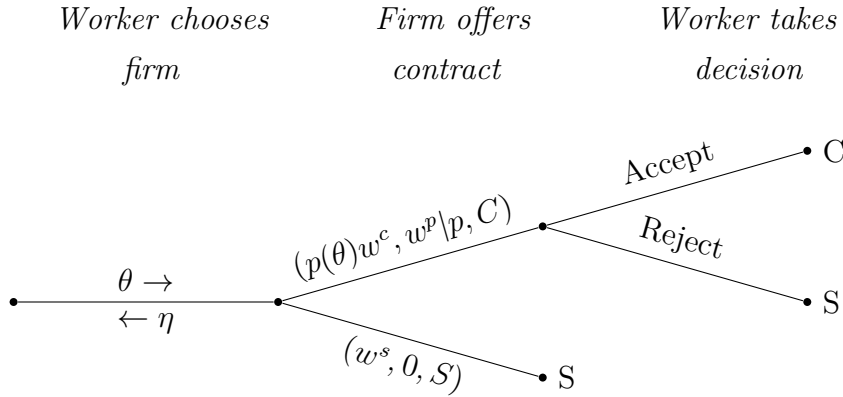
Ex-post signal: The second signal, denoted y , is generated after the completion of a complex task by a worker. The ex-post signal takes on two values, ‘pass’ or ‘fail’

⁵This 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.

$y \in (p, f)$. High skilled workers performing the complex task send out a ‘pass’ signal with probability⁶ 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 η as the probability of fooling the firm. Low skilled workers are heterogeneous in the probability of obtaining a ‘pass’ signal as η is a random variable drawn from the distribution F_η .

Hiring game

When workers become adults they randomly choose one of the two existing firms.⁷ At this point there is an exchange of information between the firm and the worker; the worker sends a signal, θ , to the employer and, at the same time, he gets to know the probability, η , 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:

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 performed⁸.

For the simple task the firm offers a fixed ex-ante wage $w^a = w^s$ and no⁹ 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.

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 θ from each worker and it is equally easy -or difficult- for a worker to receive a ‘pass’ signal from any of them.

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.

The contract for the complex task uses, though, both signals. The form of the ex-ante wage is given by $w^a = p(\theta, \tilde{\pi}^g)w^c$, where $p(\theta, \tilde{\pi}^g)$ is the Bayesian posterior probability of being high skilled given the group and the signal,

$$p(\theta, \tilde{\pi}^g) = \frac{\tilde{\pi}^g f_h(\theta)}{\tilde{\pi}^g f_h(\theta) + (1 - \tilde{\pi}^g) f_l(\theta)}. \quad (1.8)$$

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

$$w^o = \begin{cases} w^p & \text{if } y = p \\ 0 & \text{otherwise} \end{cases}, \quad (1.9)$$

where $w^p \in \mathfrak{R}_+$ 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 hiring thresholds for each group. Hiring thresholds, $\{\bar{\theta}^w, \bar{\theta}^b\}$, represent 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 attractive¹⁰.

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,

$$H^c = \sum_{g \in \{W, B\}} [1 - F_h^g(\bar{\theta}^g)] \tilde{\pi}^g, \quad L^c = \sum_{g \in \{W, B\}} [1 - F_l(\bar{\theta}^g)] (1 - \tilde{\pi}^g) F_D^g \quad (1.10)$$

$$H^s = \sum_{g \in \{W, B\}} F_h^g(\bar{\theta}^g) \tilde{\pi}^g, \quad L^s = \sum_{g \in \{W, B\}} F_l(\bar{\theta}^g) (1 - \tilde{\pi}^g) + [1 - F_l(\bar{\theta}^g)] (1 - \tilde{\pi}^g) [1 - F_D^g]. \quad (1.11)$$

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

¹⁰Think, for instance, of higher w^s or w^p .

workers 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 $\tilde{\pi}^g$ times the mass of high skilled with signals above the hiring threshold $[1 - F_h^g(\bar{\theta}^g)]$. 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 $[1 - F_D^g]$).

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

$$P(\bar{\theta}, w) = Y(\bar{\theta}, w) - \sum_g \omega^g(\bar{\theta}, w). \quad (1.12)$$

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

$$\omega^g = w^{c,g} (E[p(\theta)]_{I=1}^g H^{c,g} + E[p(\theta)]_{I=0}^g L^{c,g}) + w^{p,g} (H^{c,g} + E[\eta]_{D=1}^g L^{c,g}) + w^{s,g} N^{s,g}, \quad (1.13)$$

where $E[p(\theta)]_{I=1}^g$ is the expected Bayesian posterior probability of investors doing complex tasks from group g and $E[\eta]_{D=1}^g$ 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:

$$\tilde{\pi}_{t+1}^g = q\tilde{\pi}_t^g + (1 - q)\pi_t^g \quad \text{with } q \in (0, 1) \quad (1.14)$$

where q is an exogenous smoothing parameter and π^g is the current share of high skilled workers of the g^{th} group. This updating rule generalises the one used in Coate & Loury (1993) and I keep it for its simplicity.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.

1.3 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 a set of initial beliefs, $\{\tilde{\pi}_t^w, \tilde{\pi}_t^b\}$, allocations, $\{\hat{c}^{a,i}, \hat{c}^{o,i}\} \forall i \in (0, 1)$, along with workers' strategies, $\{\hat{I}^i, \hat{D}^i\} \forall i \in (0, 1)$, firms' strategies, $\{\hat{\theta}^W, \hat{\theta}^B\}$, outside options, $\{\hat{U}^{H,g}, \hat{U}^{L,g}\}_{g \in \{W,B\}}$, and wages $\{\hat{w}^{c,g}, \hat{w}^{p,g}, \hat{w}^{s,g}\}_{g \in \{W,B\}}$ constitute a Bayes-Nash equilibrium if they are such that:

- i Given market outside options $\{\hat{U}^{H,g}, \hat{U}^{L,g}\}_{g \in \{W,B\}}$ and firms' beliefs, firms' strategies $\{\hat{\theta}^W, \hat{\theta}^B\}$ and wages $\{\hat{w}^{c,g}, \hat{w}^{p,g}, \hat{w}^{s,g}\}_{g \in \{W,B\}}$ solve the firms' problem.
- ii $\forall i \in (0, 1)$, given firms' strategies, $\{\hat{\theta}^W, \hat{\theta}^B\}$, and wages, $\{\hat{w}^{c,g}, \hat{w}^{p,g}, \hat{w}^{s,g}\}_{g \in \{W,B\}}$, workers' strategies, $\{\hat{I}^i\}$, solve the workers' investment problem.
- iii $\forall i \in (0, 1)$, given firms' strategies, $\{\hat{\theta}^W, \hat{\theta}^B\}$, and wages, $\{\hat{w}^{c,g}, \hat{w}^{p,g}, \hat{w}^{s,g}\}_{g \in \{W,B\}}$, 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.

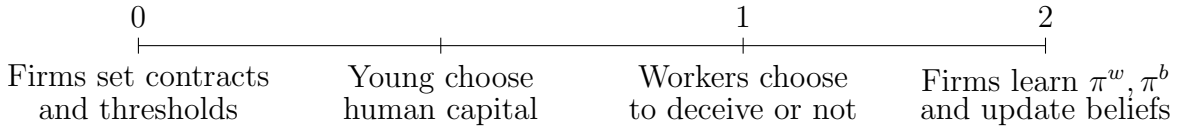
¹¹See Kim & Loury (2009) for a version of Coate & Loury with forward looking agents.

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, $\{\hat{U}^{H,g}, \hat{U}^{L,g}\}_{g \in \{w,b\}}$, and firms' beliefs, $\{\pi_t^w, \pi_t^b\}$, such that the wages $\{\hat{w}^{c,g}, \hat{w}^{p,g}, \hat{w}^{s,g}\}_{g \in \{W,B\}}$, and the hiring thresholds, $\{\hat{\theta}^W, \hat{\theta}^B\}$, set by the firms cause the workers' investment rates to be $\pi^W = \tilde{\pi}^W, \pi^B = \tilde{\pi}^B$.

1.4 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.



1.4.1 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 $\{\tilde{\pi}_t^w, \tilde{\pi}_t^b\}$ and outside options $\{\bar{U}^{H,g}, \bar{U}^{L,g}\}_{g \in \{W,B\}}$ the firm solves,

$$\text{Max}_{\bar{\theta}^g, w^{c,g}, w^{p,g}, w^{s,g}} P = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w) \quad (1.15)$$

subject to four individual rationality constraints,

$$\bar{U}^{H,g} \leq [1 - F_h(\bar{\theta}^g)]V_E^{h,c} + F_h(\bar{\theta}^g)V^s \quad IR.H^g \quad (1.16)$$

$$\bar{U}^{L,g} \leq [1 - F_l(\bar{\theta}^g)]V_E^{l,c}F_D^g + [1 - F_l(\bar{\theta}^g)]V^s[1 - F_D^g] + F_l(\bar{\theta}^g)V^s \quad IR.L^g \quad (1.17)$$

and four incentive compatibility constraints,

$$V^{h,c} \geq V^s \quad \forall \theta > \bar{\theta}^g \quad IC.H^g \quad (1.18)$$

$$V^s \geq V^{l,c}(\theta, \eta), \quad \text{for some } (\eta, \theta) \quad IC.L^g \quad (1.19)$$

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 group and skill dependant. Also note that the value functions should have a group superscript since wages may differ between groups; these are omitted for clarity.

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 expected¹² 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 utility¹³ drawn from performing simple tasks.

On the other hand, I add four IC constraints to the problem of the firm, one per group and 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 $IC.H^g$ 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 $IC.L^g$, and a semi-separating equilibrium is

¹²Expected because the ex-ante wage depends on the signal and the signal itself is a random variable.

¹³There is no expectation as this value is known with certainty by the workers.

normally¹⁴ achieved. Only in special cases¹⁵ would this constraint be satisfied for all low skilled workers and a pure separating equilibrium would exist.

1.4.2 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

$$\xi^g = E[U^g|I = 1] - E[U^g|I = 0] \quad g \in (W, B), \quad (1.20)$$

where $E[U^g|I = 1]$ and $E[U^g|I = 0]$ are the expected utility of a member of group g 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.

$$E[U^g|I = 1] = [1 - F_h(\bar{\theta}^g)]V_E^{h,c} + F_h(\bar{\theta}^g)V^s \quad g \in (W, B) \quad (1.21)$$

$$E[U^g|I = 0] = [1 - F_l(\bar{\theta}^g)]V_E^{l,c}F_D^g + [1 - F_l(\bar{\theta}^g)]V^s[1 - F_D^g] + F_l(\bar{\theta}^g)V^s \quad g \in (W, B) \quad (1.22)$$

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

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 workers' knowledge of their

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

¹⁵For instance if $E[\eta] = 0$ and $w^a = w^c$, setting $w^c \leq w^s$ guarantees full separation.

ex-ante signals, θ , and their probabilities, η , 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,

$$E[U^g|D = 1] = V_E^{l,c}(p(\theta)w^{c,g}, w^{p,g}, \eta) \quad g \in (W, B) \quad (1.23)$$

$$E[U^g|D = 0] = V^s(w^{s,g}, 0) \quad g \in (W, B). \quad (1.24)$$

To calculate the share of deceivers I first define the benefit of deceiving the firm,

$$\delta^g = E[U^g|D = 1] - E[U^g|D = 0] \quad g \in (W, B). \quad (1.25)$$

Workers with combinations of (θ, η) such that $V_E^{l,c} > V^s$ 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^g . F_D^g is calculated as:

$$F_D^g = \int_{\theta^{min}}^{\theta^{max}} \int_{\eta^{min,g}(\theta)}^{\eta^{max}} f_l(\theta, \bar{\theta}, \infty) f_\eta(\eta, \eta^l, \eta^h) d\eta d\theta \quad (1.26)$$

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:

$$E[\eta]_{D=1}^g = \int_{\theta^{min}}^{\theta^{max}} \int_{\eta^{min,g}(\theta)}^{\eta^{max}} \eta f_l(\theta, \theta^{min}, \infty) f_\eta(\eta, \eta^{min,g}, \eta^h) d\eta d\theta \quad (1.27)$$

Note that in equation 1.27 the lower truncation of f_η depends on η^{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 probability¹⁶ the result is

$$\eta^{min} = \frac{V^s}{V_E^{l,c}(p(\theta)w^c, w^p)^{-1}}. \quad (1.28)$$

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.

¹⁶Provided the inverse of V exists.

1.4.3 Outside options

The outside options, $\{\bar{U}^{H,g}, \bar{U}^{L,g}\}$, of the workers are determined in equilibrium using the zero-profit condition. Throughout the thesis I use the double screening model presented in this chapter with either one or two groups. Since the strategy to pin down the outside options is slightly different in each of the cases I start with the simplest one; the case where only one group exists¹⁷ in the economy. Refer to appendix H for extra details on how the algorithm is set.

One group model: There are two outside options, $\{\bar{U}^H, \bar{U}^L\}$, in this version of the model, one per skill. 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 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 overall¹⁸ 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).

Two groups model: In the model with two groups there exist 4 outside options $\{\bar{U}^{H,g}, \bar{U}^{L,g}\}_{g \in (W,B)}$, one per group and skill. It is obvious that, for the 2 groups model to be of any interest, there must exist a difference between the firms' beliefs about the groups W and B. Under this condition the strategy to pin down the outside options takes as a starting point a colour-blind steady state¹⁹ where firms have the same belief about each group. At this stage I shock the economy by shifting down the belief of group, say b , and I calibrate the outside options of the high skilled, $\{\bar{U}^{H,w}, \bar{U}^{H,b}\}$, until zero profit holds. The choice of changing $\bar{U}^{H,w}$ or $\bar{U}^{H,b}$ depends on the initial profit after the shock. If the profit is positive I shift up $\bar{U}^{H,w}$, if it is negative I shift $\bar{U}^{H,b}$ down.

¹⁷For instance if all were W or B.

¹⁸By Jensen's inequality as the utility function is concave.

¹⁹The strategy to pin down the outside options for the colour-blind steady state is the same as that sketched for the one group model.

1.4.4 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 than 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

$$C_t = Y_t + S_t, \quad (1.29)$$

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.

$$N_H^d = N_H^s \quad ; \quad N_L^d = N_L^s \quad (1.30)$$

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 $\bar{\theta}^{min,g}$; where $\bar{\theta}^{min,g}$ is the lowest signal such that the $IC.H^g$ constraint is still satisfied, i.e. $V^{h,c} = V^s$.

1.5 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 chapter 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.

Chapter 2

Applications of the model with 1 group

2.1 Introduction

This chapter provides three applications of the double screening model. To keep it as simple and informative as possible I work under the assumption that only one group exists in the economy.

First 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 second application 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.

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.

2.2 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 build 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 the full model presented in chapter 1 with both screening mechanisms. Apart from the just described differences the economic environment is identical.

Model specifications: The first specification is called Firm's Selection¹ (FS). In this specification the ex-post signals are excluded and, thus, performance pay does not exist. This specification resembles a signalling 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-Selection² (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 chapter 1. For details on how the (SS) and (FS) specifications are solved I refer the reader to appendixes B and C. A summary of the main features is offered in table 2.1.

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

¹Because the firm 'selects' the workers' task via the hiring threshold.

²The name arises as the screening mechanism is enabled by workers self-selecting themselves out of the complex task.

Table 2.1: Different model specifications

		Features	
		Ex-ante Signals	Ex-post signals
Model	(FS) Firm-Selection	✓	✗
	(SS) Self-Selection	✗	✓
	(DS) Double Screening	✓	✓

with the parameters shown in table 2.2 and 2.3. Parameters in table 2.2 are common while the ones in table 2.3 belong to each of the specifications.

Table 2.2: *Common parameters.*

<i>Parameter</i>	<i>Value</i>	<i>Parameter</i>	<i>Value</i>
β	0.6	α	0.9
ϵ	0.99	μ_ϕ	0.8
A	10	σ_ϕ	0.8
A_h	1	ϕ^l	0
A_l	0.2	ϕ^h	1.6
γ	0.01		

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.

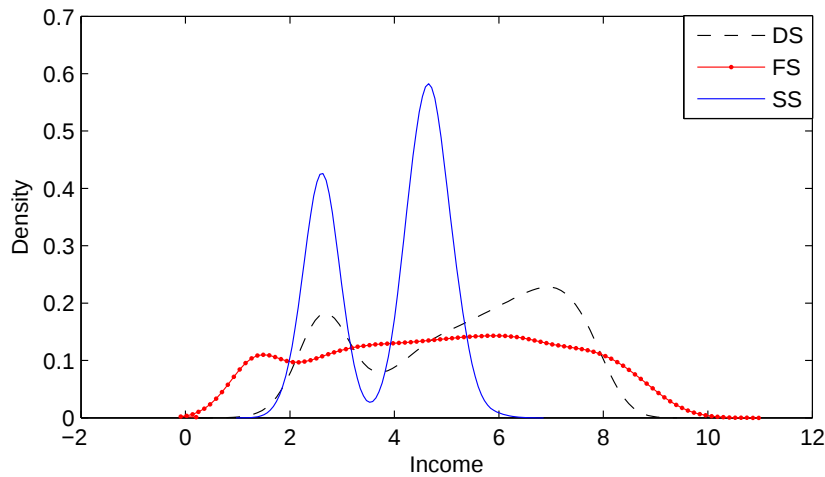
Table 2.3: *Model specific parameters.*

<i>Parameter</i>	<i>FS</i>	<i>SS</i>	<i>DS</i>
μ_η	-	0.5	0.5
σ_η	-	0.4	0.4
η^h	-	1	1
η^l	-	0	0
μ_h	2.098	0	2.098
μ_l	0	0	0

Income distributions: The income distributions that arise from these 3 specifications are shown in figure 2.1. Technical details on how income is computed can be

found in Appendix D.

Figure 2.1: Income distribution



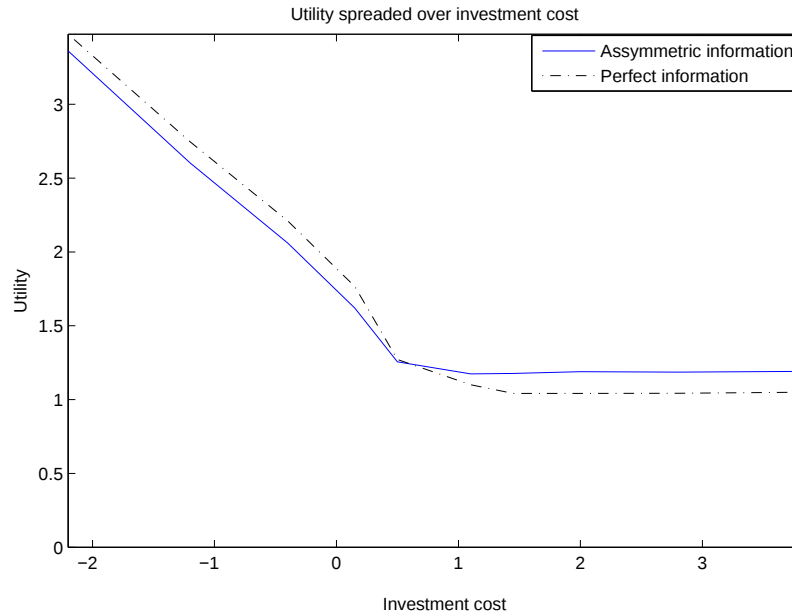
First let us compare the FS and SS specifications as it is easier to understand separately the implications of each screening method on the income distribution. On the one hand, in the SS specification, performance pay creates an economy where some individuals earn much more than the rest. On the other 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 the left tail.

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

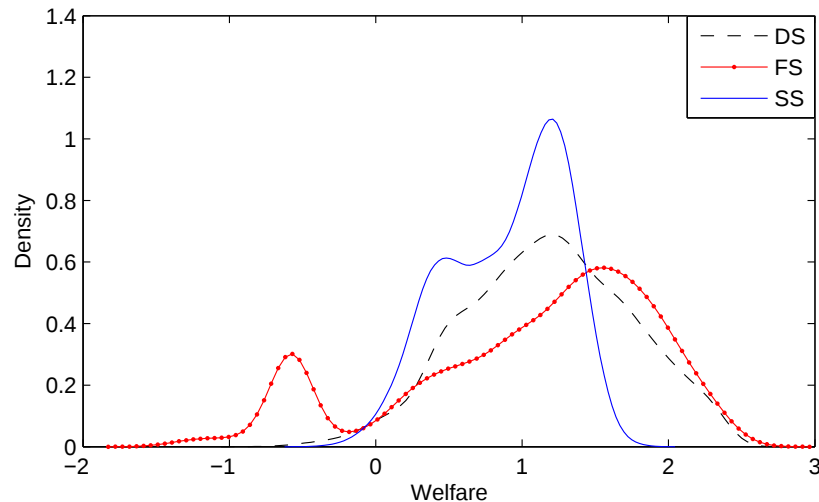
investment. I see that under perfect information high skilled workers benefit while low skilled lose out, increasing welfare dispersion.

Figure 2.2: Assymmetric vs perfect information: Utility distribution



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.

Figure 2.3: Welfare distribution.



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 occurred, 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 Bénabou & 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.

2.3 Firms' screening usage

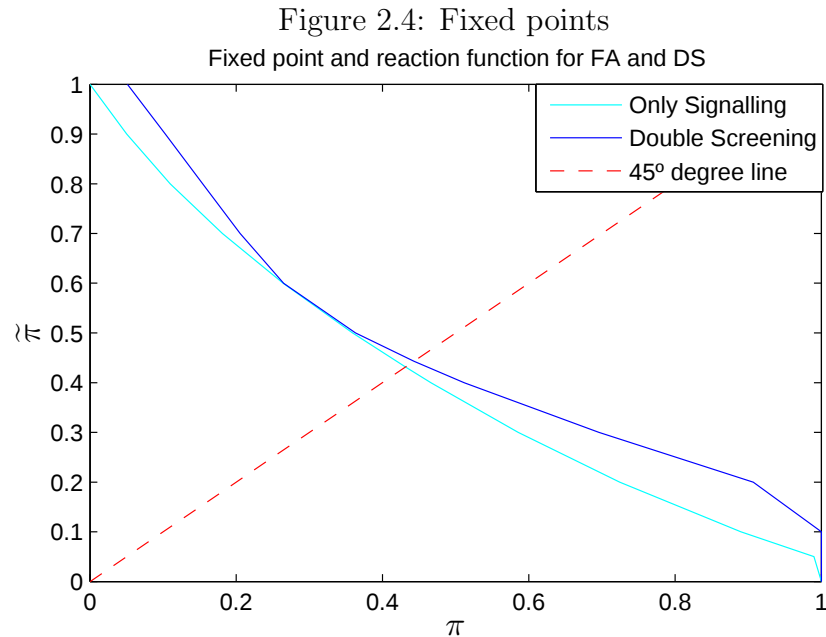
One of the main features of the DS model is the firm's ability to choose between two screening mechanisms. The double screening system makes use of the hiring thresholds and the contracts. The thresholds screen the workers thanks to the stochastic dominance of the signals obtained by the high skilled workers and through the posterior probability of investing. Since the posterior $p(\theta)$ increases monotonically in θ 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 otherwise³.

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.

To show the firm's screening choice I simulate an economy for a vector of signals' quality. Signals quality is defined as $S_q = 1 - OVC$, 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

³For instance if the ex-ante wage were flat $w^a(\theta) = w^c$.



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 fools 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.

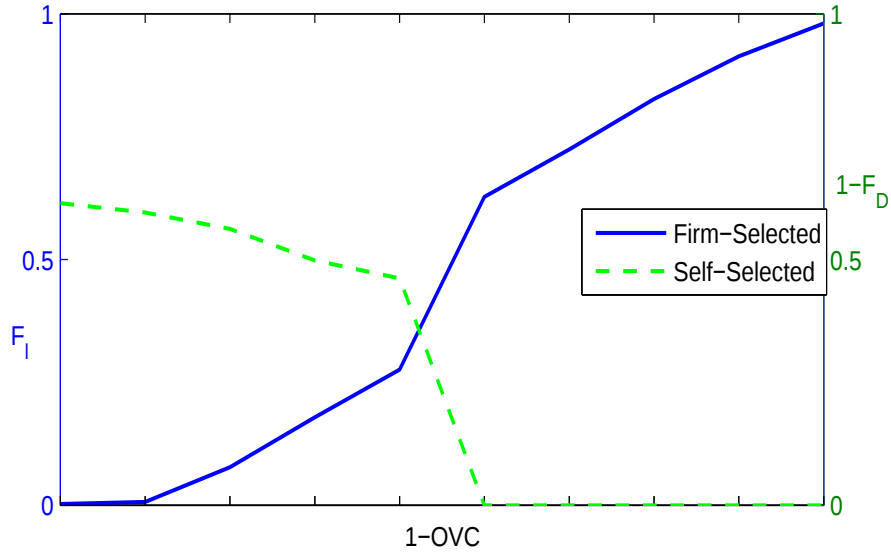
In Figure 2.6 I also show the wages offered by the firm because they offer valuable information on how firms set the contracts. Because the quality of the signals alters the pattern of human capital accumulation⁴ the interpretation of this figure is more subtle.

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

⁴This is further expanded in section 2.4.

Figure 2.5: Firm's usage of double screening.



about deceivers self-selecting 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 $p(\theta, \tilde{\pi}^g) = 1$) 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.

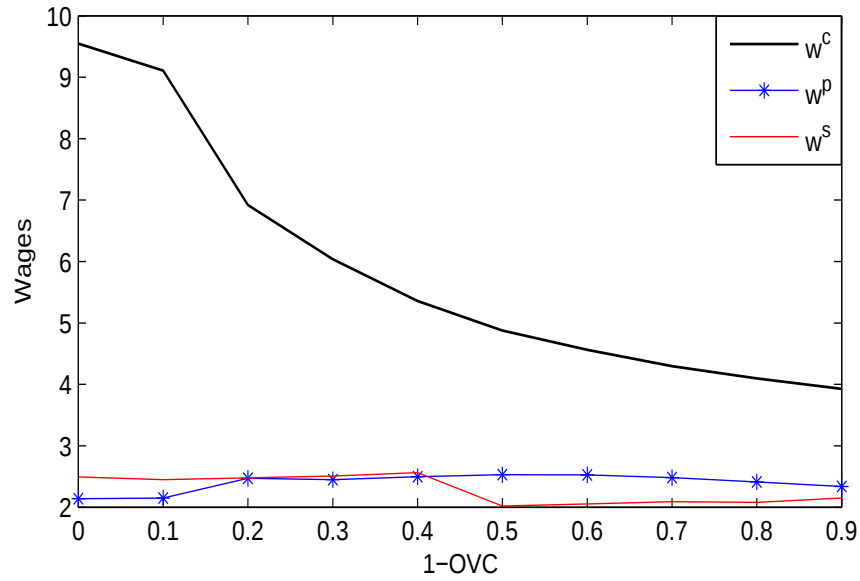
2.4 Human capital accumulation

I develop a benevolent planner⁵ 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, in this and the other chapters of the thesis, 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.

⁵Details of the problem can be found in appendix E.

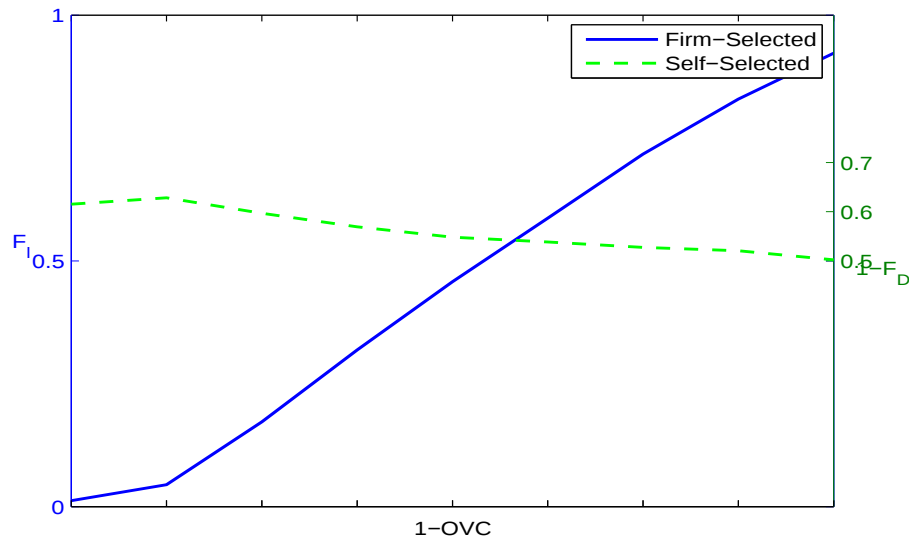
Figure 2.6: Firm's usage of double screening.



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 planner keeps using self-selection irrespective of the signals' quality. This can be seen in figure 2.7.

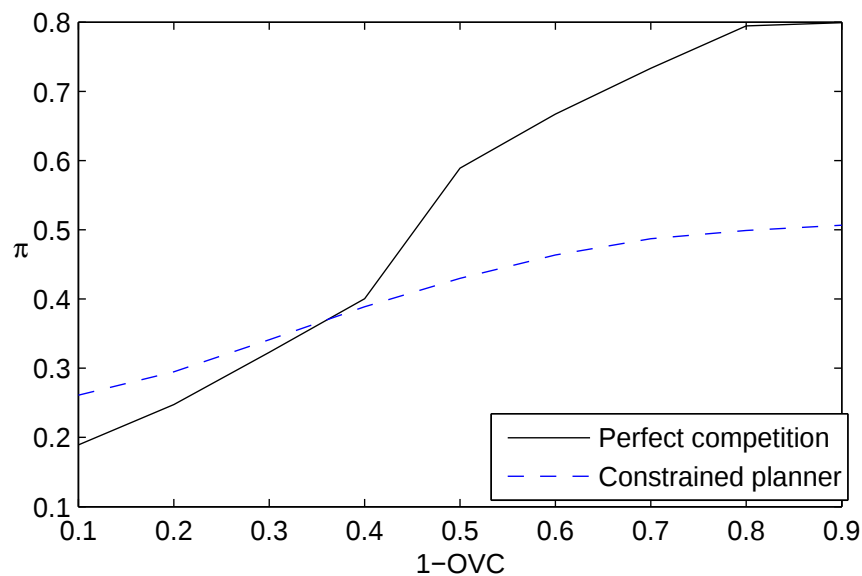
Figure 2.7: Planner's usage of double screening.



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, π_{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.

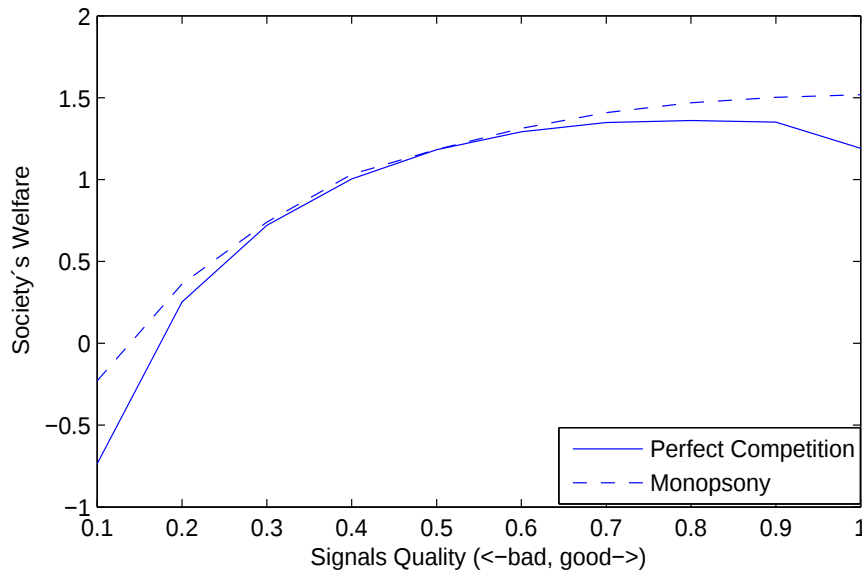
Figure 2.8: Investment rate. Perfect competition vs. constrained planner.



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

Figure 2.9: Welfare, perfect competition vs. monopsony.



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 workers⁶.

To make the analysis I build a zero profit monopsonist as an upper bound for welfare⁷. 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 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).

⁶Note the planner's feasibility condition is a zero profit constraint.

⁷Please refer to appendix G for detailed definition of welfare.

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.

2.5 Conclusion

In this chapter I use the model of adverse selection developed in chapter 1 assuming only one group populates the economy. I only use one group 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 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.

Chapter 3

Applications of the model with 2 groups

3.1 Introduction

In this chapter I use the model developed in chapter one assuming the existence of two groups, W and B . These groups of workers are perfectly distinguishable from each other by a characteristic that is unrelated to productivity. This characteristic can be thought of being race, ethnicity, language or dressing customs. We model two groups because statistical discrimination might arise if firms' beliefs about the two groups differ from each other.

The model with two groups is intended as a continuation of the papers written by Coate & Loury (1993) and Moro & Norman (2004). I extend this part of the literature on statistical discrimination by including performance pay, concave utility and wages that might differ from the marginal products.

Statistical discrimination is, alongside taste for discrimination¹, one of the prevailing explanations for treating individuals with similar skill sets differently². In practice, statistical discrimination arises whenever an easy to observe worker's characteristic (i.e. race) is correlated with a hard to observe variable such as worker's ability. An employer in this situation is tempted to use race as a cheap source of information and thus discriminate on an individual basis, even though the group itself is not being discriminated against.

¹See Becker (1953) and Arrow (1973) for an economic analysis on the consequences of employers' distaste for one of the groups in the economy.

²As arises from the surveys of Aigner & Cain (1977) or Fang & Moro (2010)

Statistical discrimination is not the only theory used to explain the existence of earnings differentials between races. Becker's (1953) theory on taste for discrimination argues that employers pay less to blacks than to whites because it makes them happier. I believe though, that explaining discrimination as coming from a profit maximising behaviour is more appealing and sound.

The literature on statistical discrimination started with Arrow (1973) and Phelps (1974). Firms in their setting use a worker's race during the hiring process insofar as it is correlated with his true productivity. Arrow's paper, in particular lays out the foundations for the literature in statistical discrimination and it seems like the source from where later papers took off. Apart from explaining the basic idea behind statistical discrimination he suggests endogenising the human capital decision and the idea that negative beliefs might be self-fulfilling. Those ideas have been used by several papers including include Borjas & Goldberg (1978), Coate & Loury (1993), Moro & Norman (2004) and Kim & Coate (2009) among others. In particular our research follows the line set by Coate & Loury and Moro & Norman. The former paper offers a model of signaling with human capital accumulation with exogenous wages. Moro & Norman improved Coate & Loury's model by allowing for general equilibrium effects and thus endogenising wages. Both papers offer two extreme equilibria, one where all groups' beliefs converge to the same value and another one in which the disfavoured group's investment rate goes to 0.

Performance pay: One concern with respect to these two models is that they are one shot games where employers set wages once depending on the workers' signals. These games' discriminatory equilibria would unravel were the employers able to observe, ex-post, the actual ability of workers and condition all wages on that observation. Firms can pay the whole wage ex-post because in these models consumers' utility is linear. I do not think linear utility is a reasonable description of reality and thus I assume concave utility.

Kim & Loury (2009) offer, in fact, a dynamic version of Coate & Loury (1993) although discriminatory equilibria could still be unravelled by performance pay. I think our dynamic model adds value with respect to their paper as I allow for pre-market differences between races and we allow firms to fully condition contracts on race among other distinctive features.

The aforementioned lack of performance pay does not fit well with current research on firm behaviour. According to empirical studies by Altonji & Pierret (2001) or Lang & Siniver (2011) I see 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 suggests that firms are capable of learning a worker's ability over time and using it. One of the results in this chapter is that the screening usage (the intensity with which the firm uses each screening method) varies with the group. In addition, I show that allowing firms to choose different contracts for different groups proves to be more effective in closing the racial wage gap than the use of the same contracts for both groups.

Policy advice: More importantly the model suggests that the use by the firms of statistical discrimination is indeed beneficial for the discriminated group in terms of earnings. This is because performance pay acts as a remedial compensation for those who actually possess high ability and encourages human capital accumulation in the disfavoured group. The extent to which statistical discrimination is actually beneficial for the disfavoured group depends on the firms' ability to condition payments on the workers' performance. I know though, that firms might have a hard time finding out a worker's actual productivity and therefore I allow for an imperfect detection of workers' performance in the model.

Another measure for the success of a group in the economy is given by the prevalence of its members in high skilled jobs. We follow Coate & Loury (1993) and Moro & Norman (2004) by assuming the existence of firms' initial beliefs with regards to the average productivity of each of the groups in the economy. These beliefs are crucial -they influence the mass of workers of each group hired for a high skilled position- but the current literature has only focused on their long-run value. This chapter offers, in addition, beliefs' evolutionary paths and a new, more flexible type of steady state.

The last section of this chapter is based on explaining the evolution of the US racial wage gap. To do so we allow for pre-market differences between races. This delves into Phelps (1974) as, to the best of my knowledge, he was the first to put together statistical discrimination and ex-ante different groups. His analysis was, though, criticised by Dennis Aigner & Glen Cain (1977), who seemed 'disconcerted' about Phelps' assumption of pre-market differences. They argued discrimination should be assessed as differences in earnings given the initial conditions are the same. But discrimination might also exist even if the initial setup is favourable to one of the groups, by making the result even *more* favourable for the members of that group. I, therefore, allow for pre-market differences as well as in beliefs.

The results of the two groups' model have consequences for how to prove the empirical existence of statistical discrimination itself. This model shows that two opposite

effects, self-selection and statistical discrimination might co-exist in a labour market with asymmetric information. If this is the case and firms hold a negative belief for one of the groups, members of the disadvantaged group will self-select more often than members of the favoured group³. A higher degree of self-selection implies that the signals of those from the disfavoured group doing complex tasks are higher than those of the members coming from the favoured group. Thus, if statistical discrimination and self-selection co-exist no difference in wages might be observed between blacks and whites even though blacks should be earning more.

More importantly explaining wage differentials by conditioning either on schooling or occupation may fail to get rid of the self-selection bias. Indeed, a regression of such kind might potentially yield no difference in wages between races while masking potential usage of race by employers. This argument can be applied in Altonji & Pierret (2001). They suggest statistical discrimination, if at all, has low prevalence because the race coefficient in their Mincer equation does not show correlation with ex-ante wages. Their result is in contrast, though, to the research by Bertrand & Mullainathan (2004), where it is argued that employers might actively use race when hiring.

3.2 Discriminatory beliefs

This section can be thought of as a continuation of a literature on statistical discrimination, in particular the models developed by Coate & Loury (1993) and Moro & Norman (2004). An important concern shared with those models is whether an initial negative belief towards one of the groups might be self-fulfilling or not. This is, indeed, a recurrent theme in the literature as stems from the review by Moro & Fang (2010).

Throughout this section I consider the economy is in an initial scenario where, maybe due to some historical reasons, firms hold a discriminatory belief towards the B group.

Definition A discriminatory belief is formed by a pair of firm's beliefs $\{\tilde{\pi}^w, \tilde{\pi}^b\}$ such that $\tilde{\pi}^w > \tilde{\pi}^b$.

I choose this particular setting because it seems to capture well the situations of, for instance, blacks and women. We can, thus, wonder whether an economy where one group is initially disadvantaged (in the sense of facing a discriminatory belief) is able

³This occurs for two reasons, first members of the favoured group invest more so fewer can self-select themselves out of complex tasks. Second given the optimal screening, firms tend to induce more self-selection from the disfavoured group.

to reach a race-neutral steady state without help. Firms' beliefs converge and a race-neutral steady is achieved whenever beliefs are equal to each other. A convergent steady state is defined as follows:

Definition A convergent steady state is formed by firms' beliefs $\{\tilde{\pi}_{ss}^w, \tilde{\pi}_{ss}^b\}$ such that $\tilde{\pi}_{ss}^w = \tilde{\pi}_{ss}^b$.

Likewise, firms' beliefs diverge and the economy never enters into a race-neutral steady state whenever firms' belief about the disadvantaged group is 0. A divergent steady state is defined as:

Definition A divergent steady state is formed by firms' beliefs $\{\tilde{\pi}_{ss}^w, \tilde{\pi}_{ss}^b\}$ such that $\tilde{\pi}_{ss}^w \in [0, 1]$ and $\tilde{\pi}_{ss}^b = 0$.

Belief divergence implies no B group workers is hired to perform the complex task. I consider that this equilibrium, even though used by other papers, is too extreme to be seen in reality and, therefore, not as informative as I would want. I fix this issue in the next section where I add a third steady state type to allow for non-zero firms' beliefs but different from each other.

The reader should keep in mind that simulations in this section assume that B and W workers obtain, on average, the same advantages and face the same costs from investing in human capital,

$$A_h^w = A_h^b \quad F_h^w = F_h^b \quad \text{and} \quad F_\phi^w = F_\phi^b. \quad (3.1)$$

Simulations are provided under two economic environments, one where firms have to pay the same wages $\{w^c, w^p, w^s\}$ to both groups (as it is assumed in the literature) and another where I allow complete contract freedom. For comparability with past papers in the area we are interested in equilibrium where no ex-ante differences exist between the groups.

3.2.1 Without contract freedom

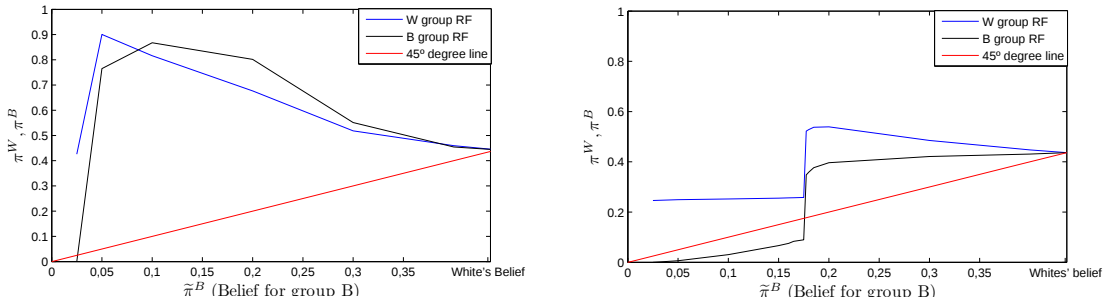
The purpose of this numerical experiment is to compare the effect of performance pay in a setup as close as possible to past papers in the area. Simulations in this subsection are performed assuming firms' ability to choose different wages for each group is removed. This means

$$w^{c,w} = w^{c,b}, \quad w^{s,w} = w^{s,b} \quad \text{and} \quad w^{p,w} = w^{p,b}. \quad (3.2)$$

Still firms are able to use group membership in the Bayesian posterior and to use different hiring thresholds as I assume they cannot be stopped from doing that (i.e. such behaviour would be legally difficult to prove).

The experiment consists on simulating two economies, one where firms have access to both screening methods (the model I propose in this thesis) and another one where firms can only use ex-ante signals to screen the workers (as it is done in past papers). Then, the experiment is done in two steps; first, for each economy, I obtain the colour-blind steady state (right hand side of each graph at Figure 3.1). Then I shock the economy with an array of firms' belief about the B group such that they are below the one for the W group under the colour-blind steady state, $\tilde{\pi}^b \in (0, \tilde{\pi}_{ss}^w)$.

Figure 3.1: *Convergence. Double screening vs. Firm's Selection (Signaling)*



The graph on the left of Figure 3.1 shows an economy where firms can use both screening methods while the one on the right shows an economy where only ex-ante signals can be used as a method of screening. The way to interpret the graph is the following, points above the 45 degree line are convergent (i.e. if I let the firm update its beliefs we arrive at a convergent steady state). Points below the 45 degree line are divergent. A divergent equilibrium is more likely the more the firm rely on group W workers, thus reducing the benefit to invest of group B. We can observe that the addition of performance pay significantly reduces the set of initial beliefs for the B group that lead to a divergent equilibrium.

In sum Figure 3.1 shows performance pay helps to reduce the set of initial beliefs where a divergent equilibrium -in the sense described in this section- occurs. Even though I cannot state the necessary conditions for convergence to be the prevailing situation independently of the initial belief, I can say the existence of a convergent equilibrium is positively associated with the quality of the detection technology. That is, in an economy where performance pay exists, the possibility of being in a divergent equilibrium is less likely to occur.

3.2.2 With contract freedom

The objective of the second numerical experiment is to show the effect of contract freedom on convergence in beliefs. Contract freedom is translated into the model as a permission to firms to pay different wages on the basis of group membership and it is, arguably, a policy controversial enough to be of interest.

The experiment consists of simulating economies with contract freedom to compare with my results from subsection 3.2. The first step to run the numerical simulation assumes a colour-blind steady state where firms' beliefs are such that $\tilde{\pi}^w = \tilde{\pi}^b$. Then we ask what would happen if the belief with respect to the B-group investment rate deteriorates, $\tilde{\pi}^b \in (0, \tilde{\pi}_{ss}^w)$.

Figure 3.2: Investment reaction functions.

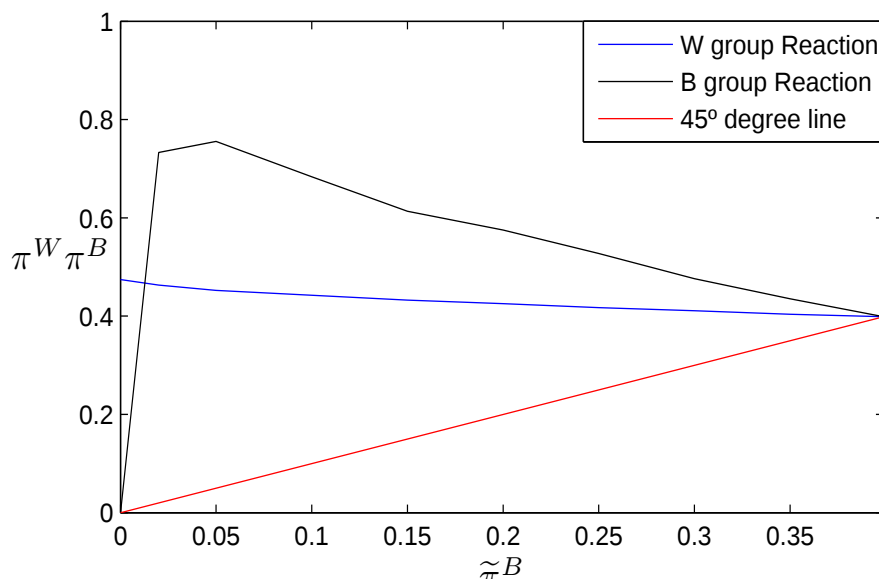


Figure 3.2 shows reaction functions of the B and W groups' investment rates for a range of firms' beliefs about group B. I fix the firms' belief about group W group at the colour-blind steady state level, π_{ss} . As in Figure 3.1 reactions above the 45 degree line are associated with a convergent steady state, below it are associated with a divergent steady state. Note that a reaction above the 45 degree line means that, if I keep iterating the model, the economy will enter the colour-blind steady state. The contrary is also true and a reaction below the 45 degree line implies that the belief about the discriminated against group is 0 in steady state.

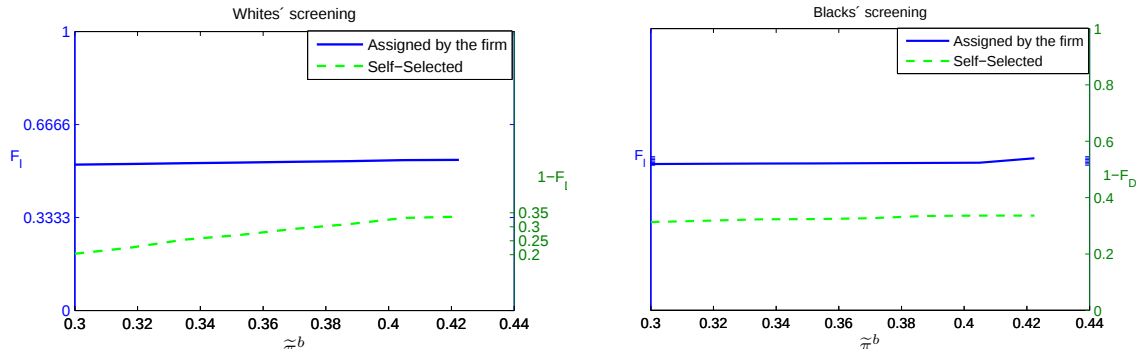
The graph in Figure 3.2 implies that no matter how low is the belief for the B group there exists convergence between the beliefs of the two groups. This result does not depend on the existence of performance pay though, rather, it depends on allowing the

firm to choose different contracts for different groups. As a result firms' hiring policies under contract freedom tend to promote the B group investment in human capital.

Next I turn to the screening behaviour of the firm when the belief towards group B becomes discriminatory. The subsequent discussion is based on a basic idea; when firms' belief about the B group deteriorates they prefer to change the payments instead of the hiring threshold. This behaviour that might seem innocuous at first is the main driver for the result shown in Figure 3.2, that the Divergent steady state does not exist under contract freedom.

Payments are changed for the B group such that the wages B workers would receive in the complex task go up while the one in the simple task goes down, relative to the W group levels. This modification in the contract increases the incentive to invest for the B group members, which explains the reaction function shown in figure 3.2.

Figure 3.3: *Screening policy with different initial belief.*



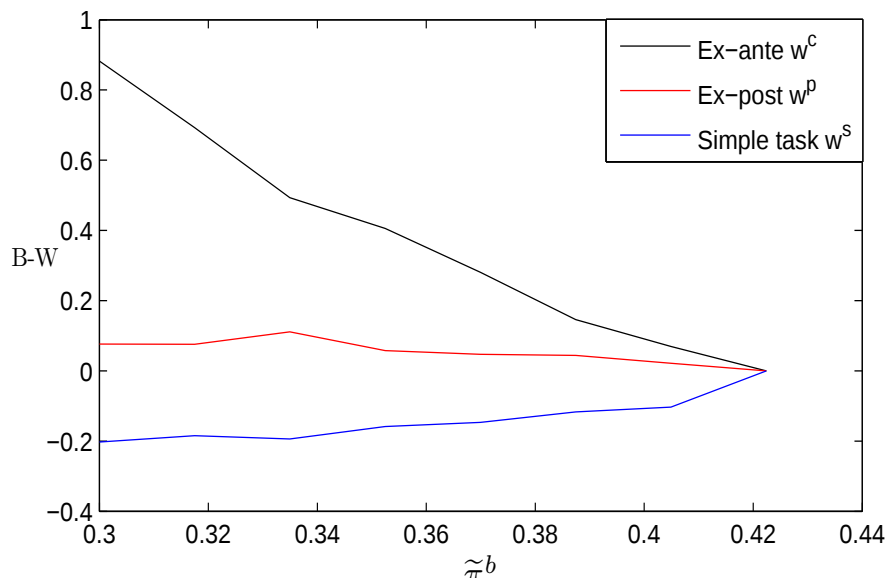
Figures 3.3 and 3.4 helps us to understand the change in firms' hiring behaviour. Figure 3.3 has two graphs containing the firms' screening behaviour for the W group (left) and the B group (right). On the right side of each graph I show the proportion of low skilled workers that are correctly sent to perform simple tasks, I call this screening mechanism Firm's Selection. On the left side of the graphs I show the proportion of low skilled that is *invited* to perform complex tasks but choose to opt out and perform the simple task; this mechanism is called Self-Selection. A range of beliefs for the B group is represented on the x-axis. Note I do not use beliefs, $\tilde{\pi}^b$, close to 0 because the little knowledge it adds does not compensate the computational demands it requires.

As a result I observe that the screening intensity is maintained for both groups with one exception. The firm relaxes the amount of W workers that are Self-Selected. This is due to a technological issue. Since the firm thinks there are less workers who invest in the economy (from group B) the employer decides to re-address the issue by allowing

more workers from the W group to join the complex task.

Then in figure 3.4 I show the differences between group B and group W's wages. We can see the reaction of the firm is to make deceiving less appealing for group B (the one perceived to have invested less) by raising the complex task ex-ante base payment, w^c , and performance pay w^p . Notice the fact that the firm, given total freedom to use the wage mechanism, chooses wages that maintain almost constant the screening intensity for both groups. This is remarkable because in versions of the model where firms are forced to set $w^b = w^w$ we observe a different screening behaviour for each group. In particular we observe that the firm relies more on self-selection for group B and on Firm-Selection for the W group.

Figure 3.4: Firms' wage policy with varying beliefs.



Two lessons are worth remembering from this section. The first one is that when the firm is allowed to set different wages for each group, the divergent equilibrium does not exist. The model will never have a B group that does not invest in human capital at all. The second lesson is that, even when the firm is not allowed to set different wages, the addition of performance pay shrinks the set of group B beliefs that lead to a divergent equilibrium. This can be generalised to shocks similar in effect to an improvement in the detection technology. These include positive developments in financial markets that allow workers to better endure performance payments.

3.3 The case of the US

In subsection 3.2.2, where we assume that both groups are ex-ante equal and that contract freedom exists, convergence between the two groups does not take longer than 2 to 3 generations to occur. This subsection starts by wondering why we do not see this in countries like the US, where the racial wage gap has been persistent since the 19th century and, for many years, firms were free to pay different wages to different races.

My focus to explain the persistence of the US racial wage gap is on the role of pre-market factors. Since these are factors differ between races in the model, I can ask not only how policies affect employment outcomes conditional on pre-market factors but also how pre-market factors respond to policies.

Indeed, past models of statistical discrimination, like Moro & Norman (2004) and Blume (2006), can be criticised for lacking an explanation of the observed racial wage gap persistence. This criticism, though, could be judged as unreasonable. These models assume both groups are ex-ante identical and just cover the labour market interactions. Research papers like Heckman (1998) or Neal & Johnson (1996) suggest that different groups do not enjoy the same opportunities before they start working. Their conclusions imply the aforementioned models are just incapable of providing a satisfactory explanation of the race gap's persistence.

In fact, Neil & Johnson argue that most of the wage differential between races could be explained by a measure of actual ability. Heckman (1998) also delves into the same argument when explaining that the use of certain variables in Mincer Equations is misleading when assessing differences in earnings. An example of Heckman's line of thought relates to the use high school graduation to assess differences in return between blacks and whites. The problem in this case arises because the census data uses as GED (General Equivalence Degree) and graduation from high school as equivalents while the signal sent by each holder is arguably different. Since more blacks than whites hold GED certificates, a standard regression using the generic variable 'High school graduate' would wrongly point to black high school graduates being *discriminated* against. On a more general level Heckman argues that if we could control for what the firm sees (not just what the dataset tells us) the explained portion of the wage would rise from 20%-30% to 60%-80%. To sum up, it can be argued that assuming the residuals from a Mincer Equation can represent discrimination that seems to be far from correct.

My model of double screening allows me to take this argument even further though, by arguing that not only could such variables as years of schooling or occupation be misleading when assessing the existence of discrimination, but variables such as the

AFQT⁴ could too. This is because the share of black workers performing simple tasks might contain a higher proportion of self-selected individuals in the sense described in this thesis. Thus, finding that both races earn the same (even after controlling for some variables) is not conclusive in terms of rejecting the existence of discrimination.

In light of the results from Neal & Johnson and Heckman it could be argued that pre-market differences might be an important cause of labour market differences between races. They can be classified into two broad groups, schooling opportunities and family background. What I do in this section is extend the analysis of section 3.2 by allowing for pre-market differences between the two groups. We focus on the inter-temporal evolution of the workers pre-market investments in education as an explanation for the race gap persistence.

The last part of this chapter is focused on convergence in earnings between blacks and whites. I compare the speed of convergence under two policy regimes, a colour-neutral hiring policy that forbids the usage of race and one that allows firms to use race when hiring. The ban for using race when hiring is associated with the one implied by the title VII of the 1964 Civil Rights Act. My results suggest that allowing firms to discriminate racially in setting wages and in allocating jobs speeds up convergence, and that this result is stronger the more firms rely on performance pay. The result occurs because the incentives to invest for the disfavoured group are stronger under the unconstrained regime than under the one banning the use of race.

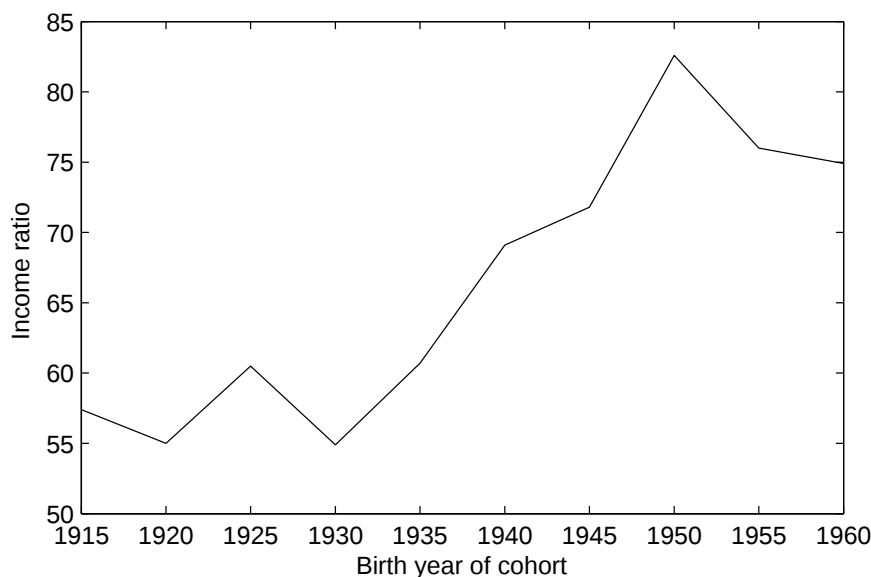
3.3.1 The racial gap in the US

It was not before the 1950's that the racial wage gap significantly started to decrease. This is shown in figure 3.5 where I provide black and white income ratios of each birth cohort. We see that the (1931-1936) cohort was the first generation who saw a decrease in the income racial gap as pointed out in Smith (1984). Smith's idea of showing the data by cohort serves to illustrate that something happened after the 1940's that helped blacks' income convergence. In particular he supports the idea that convergence in schooling opportunities due to historical reasons were a main driving force.

Indeed, Section 3.2's prediction of a fast catch up of black's income with whites' coupled with the actual persistence of the racial income gap seem to point to the existence of certain pre-market differences between groups. These pre-market differences would be the ones preventing full convergence from happening. The importance of these differences in causing the racial wage gap can be found in the work of Heckman (1998)

⁴Armed Forces Qualification test.

Figure 3.5: Black/White income gap by birth cohort



and Neil & Johnson(1996). Their papers mention differences in family background as well as in schooling opportunities⁵ as being able to explain a large part of the wage differential.

According to Heckman (1998) schooling opportunities and family support explain how a worker is perceived by his employer and his actual level of skill which, in turn, is one of the main determinants of a worker's wage.

I adapt the model to allow for pre-market factors. In doing so I translate the different distributions of family backgrounds as the cost of investment. I assume workers with worse family background will have a harder time to remain enrolled, find time to study and pay the tuition fees. In addition differences in schooling are translated into the model via the signals and the complex task productivity. On the one hand we assume a worker's signal depend partially on the school's name or the networks built while in school as pointed in Borjas (1999). On the other hand, I assume some schools and universities prepare workers better for the labour market than other. In the remainder of the section I work on the assumption that differences between blacks and whites might exist in pre-market factors.

⁵See Espenshade & Radford (2009) for a review of how families affect university admissions of their children.

3.3.2 The effect of schooling differences

Educational opportunities for blacks improved slowly during the decades before 1940 as pointed out in Bowen & Bok (1998). It was not before the 1940's that the school attendance rates rose as a result of migrational movements of southern blacks to the industrialised north, as is pointed out in Lang (2007) and Bowen & Bok (1998). The movement was due to the higher labour demand that was prevailing at the beginning of the 2nd World War.

The generation that was born during these years enjoyed an improvement of their schooling conditions. Schooling opportunities were better in the north for blacks. Lighter racial bigotry coupled with a higher proportion of blacks now living in cities helped enrolment rates. By 1960 the length of the school term and teachers salaries had almost converged between black and white schools. In addition median years of schooling for blacks aged 25-29 rose from 7 years at the beginning of the 1940's to 10.5 at the begining of the 1960's. Bowen & Bok show data on graduation rates. The proportion of blacks that graduated from high school rose from 12% to 38% while the proportion who graduated from college shifted from 1.6% to 5.4%. It was a result of this new migrational pattern that southern states started to improve the 1950's blacks' schooling conditions in an attempt to stop the flow of workers to the north.

The above facts describe two effects, higher labour demand and better schooling opportunities. Even though arguably the former could have helped blacks' earnings potential it was the improvements in schooling quality and ex-ante signals that started the virtuous circle as argued in Smith (1988)

A Divergence exercise

In this subsection I run a numerical experiment to see whether or not a reduction in education quality produces an earnings gap. The first step is to simulate a colour-blind steady state where firms hold the same belief about both groups, $\{\tilde{\pi}_{ss}^w = \tilde{\pi}_{ss}^b\}$. Then I lower the average ex-ante signal of black high skilled workers and let firms update their beliefs until a new steady state is achieved. The objective is to show that the model can move from a colour blind steady state to a steady state where one group is disadvantaged by creating pre-market differences.

In contrast to the economy shown in section 3.2, the model might be in an steady state where firms' belief about blacks is neither zero nor the same as that about whites. I refer to this steady state as 'interior'.

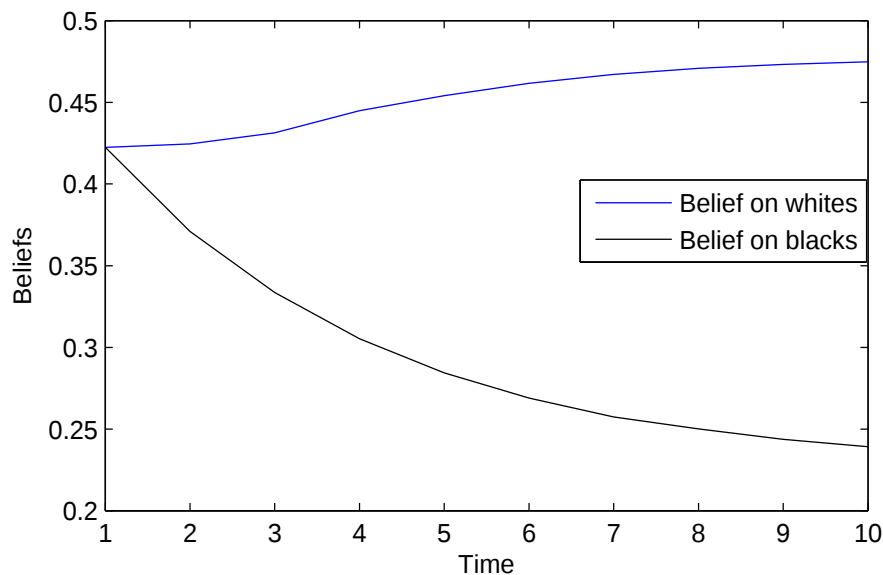
Definition An interior equilibrium is characterize by firms' beliefs $\{\tilde{\pi}_{ss}^w, \tilde{\pi}_{ss}^b\}$ such that

$\tilde{\pi}_s s^w \in [0, 1]$ and $0 < \tilde{\pi}_{ss}^b < \tilde{\pi}_{ss}^w$.

A clarification is due with regards to the relationship among the interior steady state and the other two steady states defined so far. The interior steady state belongs to the same ‘family’ as the convergent steady state -defined in Section 3.2- in the sense that if I were to remove the differences in pre-market factors beliefs, the model new steady state would show $\tilde{\pi}_{ss}^w = \tilde{\pi}_{ss}^b$ and a convergent steady state would have been obtained. The existence of the interior steady state allows us to perform quantitatively accurate assessments of a diversity of policies. Now I am not only measuring whether earnings converge or not, but also by how much and how fast.

Figure 3.6 shows firms’ beliefs over time, for blacks and whites. The x-axis represents time (each unit of time is thought to represent 25 years) and the y-axis firms’ beliefs. In the graph the beliefs diverge from each other until they enter the new -interior- steady state. This implies that in the presence of pre-market differences between groups, the steady state black/white income ratio can be shifted below 1.

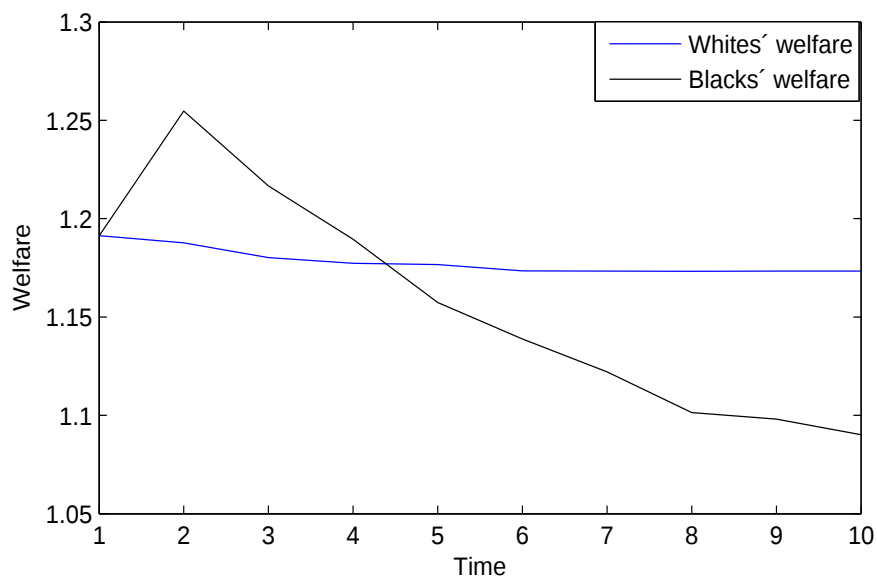
Figure 3.6: Beliefs converging to a non-colour-blind steady state.



Black workers also suffer in terms of welfare. This can be seen in figure 3.7. In this figure the y-axis represents aggregate welfare⁶. The reason why at the beginning blacks are better off is because the firm is being fooled. The employer thinks blacks are more productive than they actually are, therefore paying and hiring more than it would have chosen.

⁶See appendix G for the calculations.

Figure 3.7: Black and white's welfare.



Finally table 3.1 shows the behaviour of the firm in the new interior steady state. The representative firm chooses different screening mechanisms for each race. It uses the hiring threshold to drive out white low skilled workers (F-S, firm selection) and relies on self-selection (S-S) to get rid of the blacks deceivers.

Table 3.1: Steady state. Discriminatory steady-state.

Variables	F-S	S-S	w^c	w^p	w^s
Solutions W	0.5180	0.1585	5.4384	2.6873	2.8936
Solutions B	0.1025	0.7694	8.8930	3.0624	3.7174

As a summary, I have shown, by means of a simulation, that pre-market differences may play an important role in stopping the convergence of earnings between whites and blacks.

3.3.3 The effect of banning race from hiring

Now we move forward to 1964, the year when the first CRA powerful enough to deter discriminatory behaviour was enacted. In particular I focus on the effects that its title VII brought to the labour market. Title VII declares it unlawful ‘to fail or refuse to hire or to discharge any individual, or otherwise to discriminate against any individual with respect to her compensation, terms, conditions, or privileges of employment, because of such individual’s race, color, religion, sex, or national origin; or, the usage of the

worker's race when hiring or deciding the compensation'.

This content is translated into the model's language as a ban for different hiring thresholds, use of race in the Bayesian posterior and use of different base payments for blacks and whites. Because checking firms' hiring thresholds might be hard, I follow Coate & Loury (1993) by allowing firms to use different ones even after the ban. With regards to the Bayesian posterior, the firm uses the average belief when computing it,

$$p(\theta, \tilde{\pi}^{Ave}) = \frac{\tilde{\pi}^{Ave} f_h(\theta)}{\tilde{\pi}^{Ave} f_h(\theta) + (1 - \tilde{\pi}^{Ave}) f_l(\theta)} \quad (3.3)$$

where the average belief is equal to $\tilde{\pi}^{Ave} = 0.5\lambda^w \tilde{\pi}^w + 0.5\lambda^b \tilde{\pi}^b$. Finally, I assume the ban imposes $w^{c,w} = w^{c,b}$, $w^{p,w} = w^{p,b}$ and $w^{s,w} = w^{s,b}$.

Did the ban help to close the black/white income gap? To answer that question I go back to the 1940's and simulate two economies. These economies start off from the same discriminatory steady state where blacks are believed to be less productive than whites doing the complex task $\{A_h^b = 0.7, A_h^w = 1\}$, to obtain worse signals $\{\mu_h^b = 1, \mu_h^w = 2.098\}$ and to endure higher investment costs $\{\mu_\phi^b = 1.8, \mu_\phi^w = 1.2\}$. Then I shock both economies by shutting the gap in signals, $\{\mu_h^b = \mu_h^w = 2.098\}$, and the complex task productivity, $\{A_h^b = A_h^w = 1\}$. This is like assuming that all enjoy the same schooling opportunities. In addition, I decrease the gap in investment costs proportionally to the increase in the share of high skilled workers of the last generation. The idea is to associate smarter parents with lower costs of investments to the point that if both groups invest in human capital at the same rate, the investment costs are the same. Table 3.2 shows parameters regarding the preferences and the technology used for the calibration of the initial steady state. The income share of the complex task, α , is calibrated following data from US on high-skilled/low-skilled earnings; β is set so that there is yearly discounting of 4%.

Table 3.2: *Preferences and technology*

<i>Parameter</i>	<i>Value</i>	<i>Parameter</i>	<i>Value</i>
β	0.6	α	0.66
ϵ	0.99	γ	0.01
A	10	A_h^b	0.7
A_h^w	1	A_l	0.2

In table 3.3 I show the parameters used for detection technology, signals and investment costs for each race in the initial steady state. The detection technology is assumed

to be mild (the average fooling probability is 0.5) as well as the signals distributions. The investment costs' lower bound is set to 0 to match the domain of the incentives to invest $\mathbb{R}_+$, as allowing for negative investment costs would artificially create investors. In addition I assume the proportion of blacks in the economy is $\lambda^b = 20\%$.

Table 3.3: *Detection tech., ex-ante signals dist. and investment costs.*

<i>Parameter</i>	<i>Value</i>	<i>Parameter</i>	<i>Value</i>
μ_η	0.5	σ_ϕ	2
σ_η	0.4	ϕ^h	2.6
η^h	1	ϕ^l	0
η^l	0	μ_h^w	2.098
μ_ϕ^w	1.2	μ_h^b	1
μ_ϕ^b	1.8	$\sigma_h = \sigma_l$	2

One of the economies is called the unconstrained model, as the firm can use statistical discrimination freely. The other economy is the constrained model that mimics the enactment of the Civil Rights Act with respect to the use of race in the hiring process. The constraint means firms have to use an average belief in the formation of the Bayesian posterior and have to pay the base payments to both groups. Still we assume firms in the constrained model can set different hiring thresholds as their existence is hard to prove.

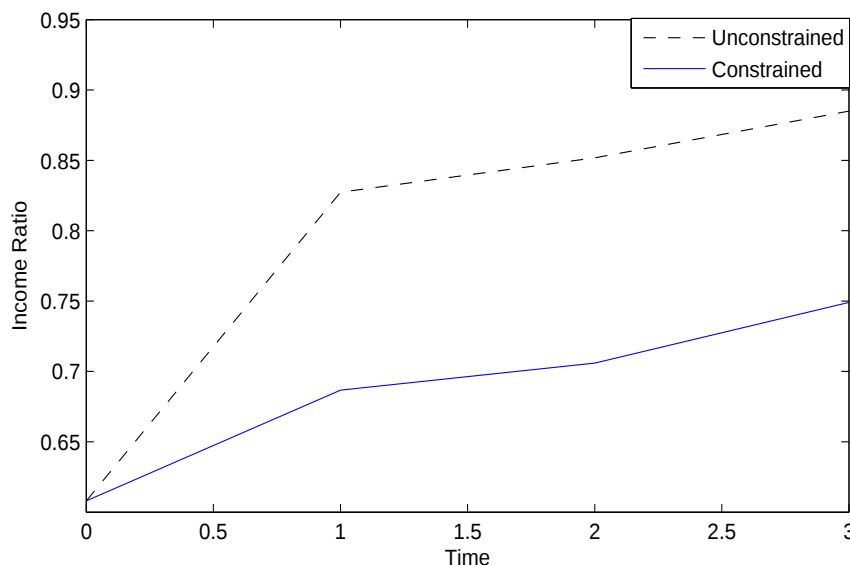
Results are shown in figure 3.8, where time units represent 25 years (a generation). The graph shows faster convergence in earnings when firms statistically discriminate than when they cannot. The simulated steady state implies a high degree of segregation in the job market. Approximately 2% of the workers doing complex tasks are black. It should be noted blacks represent 20% of the total population in this economy.

3.4 Conclusion

The first part of this chapter can be thought of as the continuation of a literature on statistical discrimination, in particular the models developed by Coate & Loury and Moro & Norman (2004).

The model we propose differs from the papers just mentioned because we allow for performance pay, assume concave utility and use a different market structure and equilibrium concept; but more importantly we allow the firm to pay different wages w^c, w^p, w^s for each group.

Figure 3.8: Black/White simulated income ratio



Two lessons are worth remember. The first one is that when allow the firm to set different wages for each group the divergent equilibrium does not exist. We will not have a B group without investing in human capital at all. The second lesson is that, had not we allowed the firm to set different wages, the addition of performance pay shrinks the set of B group beliefs that lead to a divergent equilibrium. This can be generalized to shocks similar in effect to an improvement in the detection technology. These include positive developments in financial markets that allow workers to better endure performance payments.

The last part of the chapter is a policy orientated one. I start by showing the persistence of the differences in black/white income ratio. I follow James Heckman and James Smith when arguing these differences in income might be due to pre-market differences between the races.

I allow the model to have pre-market differeces in schooling and family background. The objective is to show my model is able to mimic a situation like the one occurring in the 1940's in the US. The exercise I carry out with the latter aim is even stronger than I would need to suggest pre-market differences can sustain an income gap. I start an economy which is in a colour-blind steady state to find out that a racial wage gap is created once we allow for differences in signals for the investors in human capital.

Lastly I carry out an exercise to test whether it is preferable to ban the usage of race from the labour market or, rather it is better to defend race-neutral policies. In order to make this computational exercise we calibrate an economy to mimic the black/white income ratio prevailing in the 1960's and I let firms update their beliefs

for several periods. The result is that the economy where race can be used shows a faster convergence rate of blacks' earnings than the economy where the usage of race is banned.

Conclusion

This thesis studies the role of asymmetric information in the labour from a variety of perspectives. In chapter one we lay down the foundations of the research by building a model of adverse selection with endogenous human capital, double screening and two groups potentially populating the economy.

In the second chapter I show applications of the model with one group. I start with a discussion on the optimal screening policy of the firm and how the addition of a detection technology makes the theory to fit better with the employers' learning theory. Second I show the increase in performance pay as a plausible explanation for the rise in the US income inequality. At last I focus on human capital accumulation. I remark that perfect competition does not lead to efficient human capital accumulation. In fact, market structures with fewer firms and therefore less competition might offer more efficient outcomes.

The third chapter contributes to the literature on statistical discrimination by adding a second group to the model. Statistical discrimination arises naturally in this setting with asymmetric information and identifiable groups. This chapter's contribution is twofold:

First I focus on earnings convergence between groups. To this respect, the addition of ex-post signals on top of ex-ante signals offer firms the capacity to choose different screening methods for each group in the economy. The increase in choice facilitates convergence in earnings between the two groups because the benefit to invest in human capital grows when wage inequality increases.

Finally, and based on the model's results, I argue that the persistence of the racial wage gap in the US cannot be explained without pre-market differences in ability. In addition I argue that allowing firms to statistically discriminate might speed up the earnings convergence between groups. The latter has a stronger effect the easier it is for firms to detect and to condition payments on actual ability.

Appendix A

Joint CDF and expected value

In chapter 1 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

$$F_D = \int_{\theta^{min}}^{\theta^{max}} \int_{\eta^{min}(\theta)}^{\eta^{max}} f_l(\theta, \bar{\theta}, \infty) f_\eta(\eta, \eta^l, \eta^h) d\eta d\theta, \quad (\text{A.1})$$

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. $\theta^{max} = \infty^1$, $\eta^{max} = \eta^h$. On the contrary the calculations for the lower bounds are more subtle. On the one hand the lower bound of the outer integral θ^{min} might depend on the upper bound of the inner integral η^{max} in the following way:²:

$$\theta^{min} = \max \left\{ \frac{V^s}{(V^{l,c}(\eta^{max}))^{-1}}, \bar{\theta} \right\} \quad (\text{A.2})$$

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 θ, η satisfy the incentive compatibility constraint for a low skilled to accept a complex task.

$$\eta^{min} = \frac{V^s}{(V^{l,c}(\theta))^{-1}} \quad (\text{A.3})$$

¹ f_l is not truncated from above

²Provided the inverse of $V^{l,c}$ exists.

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,

$$E[\eta|D = 1] = \int_{\theta^{min}}^{\theta^{max}} \int_{\eta^{min}(\theta)}^{\eta^{max}} \eta f_t(\theta, \theta^{min}, \infty) f_\eta(\eta, \eta^{min}, \eta^h) d\eta d\theta, \quad (\text{A.4})$$

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 θ , then η) 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 $\tilde{\pi}$, the representative firm solves,

$$\text{Max}_{\bar{\theta}, w^c, w^s} P = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w) \quad (\text{B.1})$$

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 $C = A_h \tilde{\pi}(1 - F_h) + A_l(1 - \tilde{\pi})(1 - F_l)$ and $S = \tilde{\pi}F_h + (1 - \tilde{\pi})F_l$. The payroll function also suffer some modifications,

$$\omega(\bar{\theta}, w) = w^c (E[p(\theta)]_{I=1} H^c + E[p(\theta)]_{I=0} L^c) - w^s N^s \quad (\text{B.2})$$

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

$$\bar{U}^H \leq [1 - F_h(\bar{\theta})]V_E^{h,c} + F_h(\bar{\theta})V^s \quad IR.H \quad (\text{B.3})$$

$$\bar{U}^L \leq [1 - F_l(\bar{\theta})]V_E^{l,c} + F_l(\bar{\theta})V^s \quad IR.L \quad (\text{B.4})$$

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

$$V^{h,c} \geq V^s \quad \quad \quad IC.H \quad \quad \quad (B.5)$$

$$V^s \geq V^{l,c}(\theta, \eta), \quad \text{broken } \forall(\eta, \theta) \quad \quad \quad IC.L \quad \quad \quad (B.6)$$

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 $E[\eta]_{D=1} = \mu_{eta}$.

Appendix C

Model with only Self-Selection

This appendix describes a model with only ex-post signals. In this setting firms set a menu of contracts so that low skilled workers invited to perform the complex task reveal their skill level and opt to take on the simple task. To achieve their aim, firms condition part of the complex task wage on actual productivity that is paid ex-post. Firms, through this mechanism, are able to find out, with a certain degree of accuracy, workers' productivity.

There are two ways to build this specification both yielding the same result. One is assuming ex-ante signals still exist but are meaningless (i.e. both ex-ante signals cdf's have the same distribution), in which case every worker has a Bayesian posterior probability of having invested equal to 0.5. The second one, and the one I follow, simplifies even further the problem and assumes ex-ante signals do not exist at all. This forces firms to invite a random sample of the applicants to the complex task.

Given outside options $\{\bar{U}^H, \bar{U}^L\}$ and a belief with respect the workers' investment rate in human capital, $\tilde{\pi}$, the firm solves,

$$\text{Max}_{\bar{\theta}, w^c, w^s} P = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w) \quad (\text{C.1})$$

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 $C = A_h \tilde{\pi} \bar{\theta} + A_l \bar{\theta} (1 - \tilde{\pi}) F_D$ and $S = \tilde{\pi} (1 - \bar{\theta}) + (1 - \tilde{\pi}) (1 - \bar{\theta}) + \bar{\theta} (1 - \tilde{\pi}) [1 - F_D]$. The payroll function also suffers some modifications,

$$\omega(\bar{\theta}, w) = w^c N^c - w^p (H^c + E[\eta | D = 1] L^c) - w^s N^s \quad (\text{C.2})$$

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

$$\bar{U}^H \leq \bar{\theta}V^{h,c} + (1 - \bar{\theta})V^s \quad \text{IR.H} \quad (\text{C.3})$$

$$\bar{U}^L \leq \bar{\theta}V_E^{l,c}[1 - F_\eta(\bar{\eta})] + \bar{\theta}V^s F_\eta(\bar{\eta}) + (1 - \bar{\theta})V^s \quad \text{IR.L} \quad (\text{C.4})$$

A key difference arises in comparison to the FS specification with regards to the IC.L constraint. While in the FS specification all low skilled workers invited to perform a complex task accept it, in the SS specification some reveal their skill level and opt out to perform the simple task.

$$V^{h,c} \geq V^s \quad \text{IC.H} \quad (\text{C.5})$$

$$V^s = V^{l,c}(\bar{\eta}) \quad \text{IC.L} \quad (\text{C.6})$$

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,

$$\bar{\eta} = \left(\frac{V^s}{V^{l,c}} \right)^{-1} \quad (\text{C.7})$$

provided the inverse of $V^{l,c}$ exists.

Appendix D

Model with perfect information

This appendix shows how to work out the model presented in chapter 1 if firms were able to observe with certainty, and ex-ante, whether a worker has invested in his human capital or not. Note this extra information clears out the informational asymmetry from the model.

The maximization problem that firms face varies slightly. One of the novelties lies on firms not worrying about low skilled workers potentially performing the complex task . In addition, I simplify the model further by eliminating the performance payment, w^p , because firms know ex-ante the performance of every worker and there is no reason to withhold payments.

Given outside options $\{\bar{U}^H, \bar{U}^L\}$ and a belief with respect the workers' investment rate in human capital, $\tilde{\pi}$, the firm solves,

$$\text{Max}_{\bar{\theta}, w^c, w^s} P = Y(\bar{\theta}, w) - \omega(\bar{\theta}, w), \quad (\text{D.1})$$

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 $C = A_h \tilde{\pi} \bar{\theta}$ and $S = \tilde{\pi}(1 - \bar{\theta}) + (1 - \tilde{\pi})$. The payroll simplifies to just

$$\omega(\bar{\theta}, w) = w^c N^c - w^s N^s. \quad (\text{D.2})$$

Constraints Both Incentive Compatibility constraint are now missing because workers cannot pretend to be a different type. I am left with the 2 Individual rationality constraints.

$$\bar{U}^H \leq V^{h,c} \quad \text{IR.H} \quad (\text{D.3})$$

$$\bar{U}^L \leq V^s \quad \text{IR.L} \quad (\text{D.4})$$

Appendix E

Constrained planner

I build 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.

$$\text{Max}_{\bar{\theta}, w^c, w^p, w^s, \pi} W = W^{h,c} H^c + W^{l,c} L^c + W^s N^s - \int_{\phi^l}^{\xi} F_{\phi}(\phi) d\phi \quad (\text{E.1})$$

where total welfare W is constructed as the addition of the high skilled and low skilled individual welfare¹ 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, ξ , which is the highest investment cost a worker would pay.

First the planner faces a feasibility constraint. The planner must ensure that the earnings² of the workers do not exceed what it is produced.

$$H^c E^{h,c} + L^c E^{l,c} + N^s E^s \leq Y \quad (\text{E.2})$$

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 only add a set of IC constraints because the IR's are substituted by equation E.5,

$$V^{h,c} \geq V^s \quad \text{IC.H, and} \quad (\text{E.3})$$

¹Calculations for the welfare measure is shown in appendix G.

²Calculations for the earnings measure is shown in appendix G.

$$V^s \geq V^{l,c}(\eta, \theta) \quad \text{for some } \theta, \eta \quad \text{IC.L.} \quad (\text{E.4})$$

Lastly I add an equation that states the mass of workers who would invest in their education,

$$\pi = F_\phi(\xi). \quad (\text{E.5})$$

How to construct the incentive to invest, ξ , is explained in chapter 1. Finally I also refer the reader to chapter 1 to check how the deceiving decisions are calculated.

Appendix F

Workers' utility maximization problem.

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

F.1 Utility under certain payments

The utility obtained by high skilled workers in the complex task, with contract $\{p(\theta)w^c, w^p, C\}$, and any worker in the simple task, with contract $\{w^s, 0, S\}$, is known with certainty.

To solve for the optimal allocation, workers maximize the following function:

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

subject to

$$c^a \leq w^a \quad (\text{F.2})$$

$$c^a + c^o = w^a + w^o \quad (\text{F.3})$$

The utility function evaluated at the optimum takes two values:

$$V = \begin{cases} \frac{w^{a^{1-\epsilon}} - 1}{1-\epsilon} + \beta \frac{w^{o^{1-\epsilon}} - 1}{1-\epsilon} & \text{if } w^a \leq \frac{w^a + w^o}{1+\beta^{1/\epsilon}} \\ \frac{\left(\frac{w^a + w^o}{1+\beta^{1/\epsilon}}\right)^{1-\epsilon} - 1}{1-\epsilon} + \beta \frac{\left(\frac{w^a + w^o}{1+\beta^{-1/\epsilon}}\right)^{1-\epsilon} - 1}{1-\epsilon} & \text{if } w^a > \frac{w^a + w^o}{1+\beta^{1/\epsilon}} \end{cases} \quad (\text{F.4})$$

F.2 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 as insurance against the risk of not having anything when old. The maximization program for a given contract $\{p(\theta)w^c, w^p|y = p, C\}$ is,

$$\text{Max } U(c^a, c^{o,h}, c^{o,l}) = \frac{c^{a^{1-\epsilon}} - 1}{1-\epsilon} + \beta \left[\eta \frac{c^{o,h^{1-\epsilon}} - 1}{1-\epsilon} + (1-\eta) \frac{c^{o,l^{1-\epsilon}} - 1}{1-\epsilon} \right] \quad (\text{F.5})$$

subject to

$$c^{o,h} = w^a + w^o - c^a \quad (\text{F.6})$$

$$c^{o,l} = w^a - c^a \quad (\text{F.7})$$

An analytical expression for the value function is not easy to find. Here I provide the first order condition that gives the optimum:

$$c^{a^{-\epsilon}} = \beta \left(\zeta (w^a + w^o - c^a)^{-\epsilon} + (1-\zeta)(w^a - c^a)^{-\epsilon} \right) \quad (\text{F.8})$$

Once c^a is known I can pin down future consumption using the constraints of the problem. The value function is calculated as 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 defines the measures of welfare and earnings used in chapter 1 and appendix D. First I define welfare by workers' occupation and skill. Then I turn to do the same for the earnings measure.

G.1 Welfare

Using some previously defined objects I define average welfare for high skilled and low skilled workers, at each, the simple and the complex task. These calculation are for the double screening specification. For high skilled workers doing the complex task, welfare (net of investment costs) is:

$$W^{h,c} = \int_{\bar{\theta}}^{\infty} V^{h,c}(p(\theta)w^c, w^p) f_h(\theta, \bar{\theta}, \infty) d\theta. \quad (\text{G.1})$$

For the low skilled workers who deceive the firm I need to take into account the bias produced by workers self-selecting:¹

$$W^{l,c} = \int_{\theta^{min}}^{\infty} \int_{\eta^{min}(\theta)}^1 V^{l,c}(p(\theta)w^c, w^p, \eta) f_l(\theta, \theta^{min}, \infty) f_{\eta}(\eta, \eta^{min}, \eta^h) d\eta d\theta \quad (\text{G.2})$$

The average welfare obtained by high or low skilled workers performing the simple task is given by:

$$W^{h,s} = V^s, \quad (\text{G.3})$$

¹Refer to Appendix A for information about the bounds of this integral.

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 workers,

$$W = H^c W^{h,c} + L^c W^{l,c} + N^s W^s - \int_{\phi^l}^{\xi} F_{\phi} d\phi. \quad (\text{G.4})$$

G.2 Earnings

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

$$E^{h,c} = \int_{\bar{\theta}}^{\infty} p(\theta) w^c f_h(\theta, \bar{\theta}, \infty) d\theta + w^p. \quad (\text{G.5})$$

For low skilled workers who deceive the firm I need to take into account the bias produced by the self-selection² of workers:

$$E^{l,c} = \int_{\theta^{\min}}^{\infty} \int_{\eta^{\min}(\theta)}^{\eta^h} \eta w^p f_l f_{\eta} d\eta d\theta + \int_{\eta^{\min}}^{\eta^h} \int_{\theta^{\min}(\eta)}^{\infty} w^c p(\theta) f_l f_{\eta} d\eta d\theta. \quad (\text{G.6})$$

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

$$E^{h,s} = w^s, \quad (\text{G.7})$$

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,

$$E = H^c E^{h,c} + L^c E^{l,c} + N^s E^s. \quad (\text{G.8})$$

²Refer to Appendix A for information about the bounds of this integral.

Appendix H

Identification of outside options for competitive equilibria

H.1 Model with one group

To identify an equilibrium in the any one group model I need to find the outside options such that firms' profits are zero. The strategy I follow is to set $\bar{U}^L = -\infty$ and search for the $\bar{U}^H$ that makes profit equal to zero. In equilibrium, the value of $\bar{U}^L$ is determined endogenously.

H.2 Model with 2 groups

There is an extra problem when solving for the competitive equilibrium whenever there is more than 1 group. Even after I get rid of the IR.L constraints for low skilled¹ I still have to iterate over two numbers, $\bar{U}^{H,w}, \bar{U}^{H,b}$, until expected profit is 0.

The identification procedure goes as follows:

- Find the pair $\{\bar{U}_{eq}^{H,w}, \bar{U}_{eq}^{H,b}\}$ such that 0 profit is obtained in a model where firms' beliefs are the same for both groups.
- Set a new, discriminatory belief, $\tilde{\pi}^w > \tilde{\pi}^b$.
- If the resulting profit is negative make $\bar{U}^{H,b}$ smaller until $P=0$, holding $\bar{U}^{H,w} = \bar{U}_{eq}^{H,w}$.
- If the resulting profit is positive make $\bar{U}^{H,w}$ higher until $P=0$, holding $\bar{U}^{H,b} = \bar{U}_{eq}^{H,b}$.

¹Because they are not binding.

Bibliography

- [1] Akerlof, G. (1970). *The Market for ‘Lemons’: Quality Uncertainty and the Market Mechanism*. The Quarterly Journal of Economics, Vol. 84, No. 3, pp. 488-500.
- [2] Altonji, J. & Pierret, C. (2001). *Employer Learning and Statistical Discrimination*. The Quarterly Journal of Economics, Vol. 116, No. 1. (Feb., 2001), pp. 313-350.
- [3] Arrow, K. (1971). *The Theory of Discrimination*. Working Papers 403, Princeton University, Department of Economics, (1971) Industrial Relations Section.
- [4] Becker, G. (1957). *The Economics of Discrimination*, 1st ed. (Chicago: University of Chicago Press).
- [5] Becker, G. (1975). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. 2d ed. New York: Columbia University Press.
- [6] Bénabou, R. & Tirole, J. (2013). *Bonus Culture: Competitive Pay, Screening and Multitasking*. NBER working paper 18936.
- [7] Bertrand, M. & Mullainathan, S. (2003). *Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination*. The American Economic Review, vol.94, No 4, 991-1013.
- [8] Blume, L. (2006). *The Dynamics of Statistical Discrimination*. Economic Journal, Royal Economic Society, vol.116(515), (2006), pages F480-F498, November.
- [9] Borjas, G. & Goldberg, M. (1978). *Biased Screening and Discrimination in the Labor Market*. The American Economic Review, vol.68, No 5, 918-922.

- [10] Borjas, G. (1999). *Heaven's door: immigration policy and the American economy*, Princeton University Press.
- [11] Bowen, W. & Bok, D. (1998). *The shape of the river*. Princeton University Press.
- [12] Coate, S. & Loury, G. (1993). *Will Affirmative-Action Policies Eliminate Negative Stereotypes?*. The American Economic Review, vol.83, No 5 (December 1993), 1220-1240.
- [13] Espenshade, T. & Radford, A. (2008). *No longer separate not yet equal*. Princeton University Press.
- [14] Fang, H. & Moro, A. (2010). *Theories of Statistical Discrimination and Affirmative Action: A survey*. NBER Working Paper 15860.
- [15] Gall, T. & Legros, T. & Andrew Newman (2006). *The timing of education*. Journal of the European Economic Association, MIT Press, vol. 4(2-3), pages 427-435, 04-05.
- [16] Heckman, J. (1998). *Detecting Discrimination*. The Journal of Economic Perspectives, Vol. 12, No. 2, pp. 101-116.
- [17] Heywood, J. & Parent, D. (2012). *Performance Pay and the White-Black Wage Gap*. Journal of Labor Economics, Vol. 30, No. 2, pp. 249 - 290.
- [18] Lang, K. (2007). *Poverty and Discrimination*. Princeton University Press.
- [19] Lang, K. & Siniver, E. (2011). *Why is an elite undergraduate education valuable? Evidence from Israel*. Labour Economics, Elsevier, vol. 18(6), pages 767-777.
- [20] Moro, A. & Norman, P. (2003). *Affirmative action in a competitive economy*. Journal of Public Economics, 87,(2003) pp. 567-594.
- [21] Moro, A. & Norman, P. (2004). *A general equilibrium model of statistical discrimination*. Journal of Economic Theory, vol. 114 (2004), pages 1-30.

- [22] Neal, D. & Johnson, W. (1996). *The Role of Premarket Factors in Black-White Wage Differences*. Journal of Political Economy, Vol. 104, No. 5, pp. 869-895.
- [23] Riley, J. (2001). *Silver Signals: Twenty-Five Years of Screening and Signaling*. Journal of Economic Literature, Vol.39, pp. 432-478.
- [24] Rothschild, M. & Stiglitz, J. (1976). *Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information*. The Quarterly Journal of Economics, Vol. 90, No. 4, pp. 629-649.
- [25] Sala I Martín, X. (2006). *The World Distribution of Income: Falling Poverty and... Convergence, Period*. The Quarterly Journal of Economics, Vol. 121, No. 2, pp. 351-397.
- [26] Smith, J. (1984). *Race and Human Capital*, American Economic Review, 74, p. 685-698.
- [27] Spence, M. (1973). *Job Market Signalling*. The Quarterly Journal of Economics, Vol. 87, No.3, pp. 355-374.