The Impact of Performance Pay on Racial Discrimination

ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Department of Economics

Supervisor: **Dr. M. Bøg**

Name: **Remco Mocking**

Exam number: **295055**

E-mail address: **295055rm@student.eur.nl**

The Impact of Performance Pay on Racial Discrimination

Remco Mocking

Erasmus University Rotterdam

August, 2009

~ Abstract ~

In this study we analyze the impact of a switch to performance pay on negative stereotypes in a model of statistical discrimination. We introduce effort incentives in the model by Coate and Loury (1993) and analyze discriminatory outcomes under different wage regimes. We show that negative stereotypes are possibly eradicated when performance pay is introduced. However, we find that a weak relationship between worker effort and firm output might result in an increase in the absolute level of discrimination when performance pay is introduced.

Keywords: statistical discrimination, negative stereotypes, performance pay

1. Introduction

Traditionally, the literature in the field of personnel economics focuses on the effects of providing incentives on organizational performance. The effects of performance pay (e.g. Lazear, 2000), tournaments (e.g. Lazear and Rosen, 1981) and non-monetary-incentives (e.g. Akerlof and Kranton, 2005) are evaluated in a bunch of studies. The literature on racial discrimination mainly focuses on the mechanisms that account for inequalities between races. Two main models have been developed: models of statistical discrimination (e.g. Arrow, 1972; Phelps, 1972) and models of taste-based discrimination (e.g. Becker, 1957; Welch 1967). In models of taste-based discrimination, racial differentials are caused by personal racial preferences of employers, co-workers or customers. On the contrary, in models of statistical discrimination, racial differentials arise due to negative beliefs of employers about a certain group of workers. In this study we build upon the model of statistical discrimination developed by Coate and Loury (1993).

Despite the fact that both performance pay and racial discrimination are widely discussed topics in modern microeconomics, there is not much (theoretical) literature on the link between the two issues. Some empirical studies (e.g. Belman and Heywood, 1988; Fang and Heywood, 2006; Heywood and O'Halloran, 2005) found evidence for a reduced racial wage gap in companies where wages were based on performance of workers. Furthermore, Heywood and O'Halloran (2006) present a theoretical explanation for the reduced racial wage gap when workers are paid according to their output. The studies mentioned focus on racial earnings differentials within firms. These studies form an interesting starting point for our study.

However, we are not particularly interested in the link between performance pay and racial earnings differentials. We are interested in discriminatory practices at the appointment or promotion stage of the labour market. Currently, only a few studies have focused on this effect. For example, Elvira and Town (2001) have found empirical evidence for differences in supervisors' performance evaluations dependent on the race of their subordinates. The focus in our study is on the link between performance related pay and negative stereotypes held by employers. We ask ourselves the question whether the introduction of performance related pay can eradicate negative stereotypes. Starting from a situation with flat wages and negative stereotypes we analyze the impact of a switch to a performance pay regime on discrimination.

We are the first to compare appointment rates of different identifiable groups of workers under different wage regimes in a model of statistical discrimination. In such a model employer beliefs play an important role. When an employer believes that a worker from group A is less likely to be productive than a worker from group B, it is said that the employer holds negative stereotypes against group A. The model on statistical discrimination by Coate and Loury (1993) forms the basis for the analysis on the impact of performance pay on negative stereotypes.

The Coate and Loury model describes a job-assignment game, in which workers and employers are randomly matched with each other. Employers have to decide whether they assign the worker to a high-skilled job or to a low-skilled job. Workers have to decide whether or not to make a costly investment which makes them productive in the high-skilled job. The employer observes the group identity of the worker, though he does not observe the productivity of the worker. In addition, the employer observes a noisy signal which gives an indication about the likelihood that the worker is qualified. Qualified workers are more likely to emit a high signal. Employers form beliefs about the productivity of the different groups based on prior experiences. In equilibrium, these beliefs must be confirmed. That is, the fraction of investors in a particular group must equal the employer's belief about the investment rate in that group.

In a discriminatory equilibrium employers believe workers from one group to be less productive than workers from another group. In case of such negative stereotypes it is less likely that an employer assigns a worker from the disadvantaged group to the high-skilled job. This, in turn, reduces the investment incentives of workers from the discriminated group. It is possible that a discriminatory equilibrium persist even when the different groups are ex ante identical.

We extent the model of Coate and Loury by incorporating effort incentives. This makes it possible to analyze different wage regimes in the high-skilled job. We use a standard principal-agent model to derive optimal contracts under a flat wages regime and a regime characterized by performance pay. From the optimal contracts we can specify employer and worker payoffs associated with assigning a qualified or an unqualified worker to the high-skilled job. We show that these payoffs vary under the different wage regimes.

In order to be able to analyze the impact of performance pay on negative stereotypes, we start in a situation with flat wages and discrimination against one group of workers. We argue that the switch from flat wages to performance related pay decreases the likelihood that a discriminatory equilibrium persists. This can be explained by the changes in the employer and worker payoffs. Under performance pay, employer profits associated with assigning a qualified worker to the high-skilled job are higher. This leads them to be less strict in their *assignment standards*. Workers get extra investment incentives under performance pay, since only qualified workers are eligible for an extra bonus payment. The combination of these two effects increases the likelihood that negative stereotypes are eradicated under performance pay.

The effects of the introduction of performance pay described above might sound obvious. Employers become less strict in their assignment standards and workers become qualified at a higher rate because of increased monetary incentives. We show that negative stereotypes are more likely to be eradicated when the relationship between effort and output is stronger. A strong relationship between the effort exerted by the worker and the output of the firm results in higher monetary incentives. Next to that employer profits of assigning a qualified worker increase in the strength of the effort-output relationship. As a result, the decrease in employer standards and the increase in the worker investment rate are more severe the stronger the effort-output relationship. The combination of these two effects makes it more likely that negative stereotypes are eliminated when the effort-output relationship is relatively strong.

However, the mechanism described above is not as clear-cut as it seems. We show that the effect of introducing performance pay on negative stereotypes becomes uncertain when the effort-output relationship is relatively weak. In that case the possibility that negative stereotypes persist exists. Besides that it is even possible that the absolute level of discrimination increases. That is, the fraction of investors increases in both groups but the investment rate of workers from the advantaged group increases more than the investment rate in the discriminated group. A policy implication following from this finding is that the stimulation of performance pay programs is not per definition a good policy when the objective is a reduction of discrimination at the appointment stage of the labor market.

The outline of this study is as follows. Section 2 gives a short literature overview. We discuss several (empirical) articles on the link between performance pay and racial discrimination. Besides that we will explain the contribution of our study to the existing literature. For those who are not familiar with the model of Coate and Loury we start with a review of that model in section 3. We give the assumptions of the model and we derive optimal behavior of employers and workers. After that, we explain the equilibrium concept in the Coate and Loury model. In the second part of section 3 we develop the framework of our study. Starting from the baseline model we introduce effort incentives, which enables us to analyze different wage regimes. We will focus on two different wage regimes, namely flat wages and performance related pay. We derive optimal contracts under the two regimes and we present the payoffs we need for our analysis.

In section 4 we use the payoffs derived in section 3 to study optimal behavior of employers and workers under the different regimes. We formulate the impact of performance related pay on employer and worker behavior in two propositions. Starting from a situation with flat wages and negative stereotypes we look at the impact of performance pay on this discriminatory equilibrium. We show how the introduction of performance pay impacts the equilibrium outcome of the model. Furthermore we analyze the effect of the relationship between worker effort and firm output on the likelihood that negative stereotypes are eradicated. The impact of the introduction of performance pay on negative stereotypes is formulated in two other propositions. At the end of section 4 we look at the social consequences related to the introduction of performance pay and we try to find out whether social interests are always aligned with the interests of the individual firm. Section 5 summarizes the most important results of our analysis. We give an interpretation of the results we obtained and we formulate an answer to question whether the introduction of performance pay can eradicate negative stereotypes.

2. Literature

There is a list of empirical work on the link between performance pay and racial earning differentials. For example Belman and Heywood (1988) analyze a dataset including individual wage and industry data on performance pay use in the US. They find that incentive pay tends to have (at best) a very small effect on the earnings of whites, but a major positive effect on the earnings of blacks. This reduces the part of the wage gap that can be explained by discrimination. Next to that it indicates that it is more difficult to discriminate when pay is based on output measures as compared to the situation where pay is related to subjective evaluations of output.

Fang and Heywood (2006) are the first to do research on the link between the payment method and ethnical wage differentials in Canada. Their hypothesis is that performance pay must reduce ethnic wage differentials for two reasons. First of all the intensity of prejudice must be greater for an employer to discriminate under performance pay than when effort is evaluated subjectively. Secondly, when productivity is measured using a standardized norm, the practice of discrimination becomes more transparent. This will result in an increased probability of detection of discriminatory practices, which increases the cost of discrimination and reduces its extent. Using data from the Canadian Workplace and Employee Survey they find evidence for a significant ethnic wage differential for those paid by time rates. However, they did not find evidence for an ethnic wage differential for those receiving wages dependent on output.

In a same vein, Heywood and O'Halloran (2005) analyze US data from the National Longitudinal Survey of Youth. They compare the earnings of blacks and Hispanics to the earnings of non-blacks and non-Hispanics. They find a significant racial earnings gap for those receiving time rates. For those receiving output-based pay they did not find evidence for a racial earnings difference. An interesting additional finding is that for those receiving individual bonus payments the racial earnings gap is larger than for those receiving time rates. This can be explained by the fact that bonus payments are based on the subjective evaluation of a supervisor. This analysis is more or less comparable to the two studies we described above. They focus on earnings differentials between two groups and the relation with the method of pay. In addition to the previous studies Heywood and O'Halloran present a theoretical model which gives a possible explanation for the findings. They assume a model of taste-based discrimination in which the utility of the employer increases with the extent of discrimination. On the other hand, the utility of the employer decreases in the probability that the discriminatory practices of the employer are detected. Under performance pay the detection probability increases which results in a lower optimal level of discrimination.

The studies described above are all related to the link between performance pay and racial wage differences. The literature on the link between performance pay and discrimination at the appointment or promotion stage is more limited. For instance, Elvira and Town (2001) investigate the role of the racial composition of the employee-supervisor pair and worker productivity on (subjective) performance evaluations within a large US company. They find that the subjective performance evaluation significantly depends on the race of the worker. Controlling for actual productivity which is measured using an objective performance measure, they find that blacks receive lower ratings than whites. The subjective performance evaluation directly influences wages and career opportunities.

It is even more interesting to see that they find different outcomes for different employee-supervisor pairs. A white worker working for a white boss receives significantly higher ratings than a black worker working for a white supervisor. The same holds when they compare black workers working for a black boss and white worker working for a black boss. That is, they find that the evaluation of a workers depends on the races of the supervisor and the worker.

In the study by Elvira and Town, wages depend on subjective performance measures executed by the supervisor. They investigate the effect of the racial composition of employee-supervisor pairs on discriminatory outcomes. In our study we exclude the effect of racial differences between supervisors and subordinates by assuming that an objective performance measure is used to evaluate the performance of the worker. We contribute to the existing literature by studying the effect of different wage regimes on appointment or promotion rates of minority workers. In the section below we will start with an overview of the model by Coate and Loury (1993) which forms the framework for our analysis.

3. Model

In this section we will build the framework in which we analyze the impact of performance pay on negative stereotypes. The framework is largely based on the model of statistical discrimination developed by Coate and Loury (1993). The Coate and Loury model is a job-assignment model in which employers are randomly matched with workers from a large population. Employers offer two types of jobs: a low-skilled job and a high-skilled job. Employers do not observe the productivity of the worker, they only observe the group identity of the worker. This leads to a situation in which employers form beliefs about the productivity of certain groups. In an equilibrium of the model the beliefs of the employer are confirmed. Employers have negative stereotypes about a certain group when they believe that group to be less productive on average. An employer holding negative stereotypes about a certain group of workers is less likely to assign workers of that group to the high-skilled job. Workers have to decide whether or not to make a costly investment. The investment makes them productive in the high-skilled job. In case of negative stereotypes about a group of workers, it is less likely that workers from that group make the investment. After all, even when they make the investment, employers are less likely to assign them to the high-skilled job.

We start this section with a formal overview of the basic assumptions and the workings of the Coate and Loury model. We will show how workers and employers react in the basic model and we will explain the equilibrium concept. In the last part of this section we explain how we change the baseline model in order to analyze different wage regimes. We present a variant of the Coate and Loury model in which effort incentives are included. This version of the Coate and Loury model helps us to examine the impact of performance pay on negative stereotypes in section 4.

3.1 Coate and Loury Model

The Coate and Loury model is a job-assignment model with a large number of risk-neutral employers, offering two types of jobs: a high-skilled job and a low-skilled job. The model can also be used to evaluate promotion rates instead of appointment rates. In that case one should imagine that workers are contracted for a probationary period. After that period the employer has to decide whether or not to assign the worker to the high-skilled job. Each employer is randomly matched with a lot of workers from a large population. The population consists of two identifiable groups of workers, W's (fraction λ) and B's (fraction $1-\lambda$). Once matched with a worker, the employer has to assign the worker to one of the two jobs.

Workers need to decide whether or not to make a costly investment, which makes them valuable in the high-skilled task. The costs of this investment are distributed identically in both groups according to some CDF $G(c)$. Both qualified and unqualified workers want to get assigned to the high-skilled job, which yield a gross return equal to w. Employers want to assign as many qualified workers to the high-skilled job as possible, since this yields a net return of $x_q > 0$. Assigning an unqualified worker to the high-skilled job yield a payoff equal to $-x_u < 0$. The ratio r is defined as x_q/x_u. The gross return for a worker in the low-skilled job is normalized to zero. The net return for the employer of assigning a worker to the low-skilled job is normalized to zero as well.

The problem is that the employer is not able to observe the investment decision of the worker at the assignment stage. The only information observed by the employer is the group membership of the worker (W or B) and a noisy signal ($\theta \in [0,1]$). The signal gives an indication about the likelihood that the worker is qualified. One could imagine the signal emitted by the worker as the outcome of a job interview or a period of intensive monitoring in a probationary period. The employer is still not sure whether the worker is qualified for the high-skilled job, but after the interview or the probationary period he has an indication about the likelihood that the worker is qualified.

The Monotone Likelihood Ratio Property (MLRP) ensures that higher signal are more likely to be emitted by qualified workers. The distribution from which the signal θ is drawn depends on the investment decision of the worker. When the worker has made the investment $F_q(\theta)$ represents the probability that the signal does not exceed θ and $f_q(\theta)$ is the related probability density function. In case of no investment $F_u(\theta)$ gives the probability that the signal does not exceed θ and $f_u(\theta)$ is the related probabilty density function. The likelihood ratio at θ is defined as the ratio of the two probability density functions: $\varphi(\theta) \equiv f_u(\theta)/f_q(\theta)$. It is assumed that $\varphi(\theta)$ is non-increasing in θ on the interval $[0,1]$. Under that assumption it must hold that $F_q(\theta) \leq F_u(\theta)$ for all values of the signal θ. In words, higher values of the signal are more likely to be emitted by qualified workers. The implication is that workers who made the investment are more likely to emit a high signal. This is important because we will see below that employers base their assignment decision on the signal emitted by the worker. Workers who emit a higher signal are more likely to get assigned to the high-skilled job than workers sending a low signal.

Under these conditions, employers optimally choose to set threshold standards of the signal θ. These standards depend on the prior belief the employer has about a certain group. The employer assigns a worker from group i to the high-skilled job, only when the worker meets the threshold standard applicable for group i. Workers have to decide whether they become qualified or not. In order to become qualified they need to make a costly investment. The fraction of workers with investment costs below some $\bar{c}$, $G(\bar{c})$, becomes qualified. The level of investment costs for which it is worthwhile to become qualified depends on two factors: the gross return from being assigned to the high-skilled job and the increased probability of assignment to the high-skilled job after making the investment.

3.1.1 Optimal Standards

Knowing the assumptions of the Coate and Loury model, we are now able to derive optimal employer behavior. As described above employers optimally set a threshold standard $s^*(\pi^i)$, where π^i is the prior belief held by the employer about the proportion of qualified workers in group i. After observing the signal θ emitted by the worker, the employer updates his beliefs according to Bayes' rule. The posterior beliefs are given by:

$$\xi(\pi^i, \theta) \equiv \frac{\pi^i f_q(\theta)}{\pi^i f_q(\theta) + (1-\pi^i) f_u(\theta)} = \frac{1}{1 + [(1-\pi^i)/\pi^i]\varphi(\theta)} \quad (1)$$

After observing the worker's group identity and his signal, the employer is able to determine the expected payoff from assigning the worker to the high-skilled task. The employer decides to assign the worker to the high-skilled job when the benefit of doing so exceeds the benefit from assigning him to the low-skilled job. This policy can be expressed by the following inequality:

$$\xi(\pi^i,\theta)x_q - \left(1-\xi(\pi^i,\theta)\right)x_u \geq 0 \qquad (2)$$

Rearranging terms and using our definition of *r* results in the expression for the optimal standards given by equation (3). It can be proven that $s^*(\pi^i)$ is decreasing in π^i. This implies that more optimistic prior beliefs result in less strict standards. The intuition behind this is that employers do not attach any value to the signal when they believe (almost) every worker to be qualified. In contrast, when employers believe that a very low percentage of the workers becomes qualified, they will only assign workers emitting very high signals. The result is that standards are very low for high prior beliefs, while standards become very strict for low prior beliefs about worker productivity.

Next to that employers set lower standards for any belief π^i, when the ratio *r* increases. When the gains from assigning a qualified worker are relatively high compared to the losses associated with assigning an unqualified worker, standards become easier.

$$s^*(\pi^i) = \min\left\{\theta \in [0,1] | r \geq \left(\frac{1-\pi^i}{\pi^i}\right)\varphi(\theta)\right\} \qquad (3)$$

3.1.2 Worker Investment

Now that we know the standards set by the employer, we can derive optimal worker behavior. A worker will invest if and only if the expected benefit of doing so exceeds the expected benefit of not investing. The expected benefit of investing is determined by the probability of meeting the standard, the gross return from assignment and the investment costs. The expected benefit of not investing is solely determined by the probability of assignment and the gross return of getting assigned since the worker does not have any costs of investment. The probability of assignment to the high-skilled job is higher for qualified workers as compared to unqualified workers as a result of our assumptions on the signaling technology. Signals for qualified workers are drawn from another distribution than signals for unqualified workers. As explained above, higher signals are more likely to be emitted by qualified workers. This, in turn, increases the probability of assignment. The probability of assignment equals $1-F_u(s)$ for unqualified workers and $1-F_q(s)$ for qualified workers. Therefore, the expected benefit of investing equals $\left[1-F_q(s)\right]w - c_i$. The expected benefit of not making the investment equals $[1 - F_u(s)]w$. This leads us to conclude that a worker will make the investment if and only if $c_i < [F_u(s) - F_q(s)]w$. The second part of this inequality is defined as the expected benefit from investing excluding investment costs for any worker, given standard s:

$$E[B(s)] \equiv w[F_u(s) - F_q(s)] \qquad (4)$$

Given the assumptions about the signaling technology, it can be proven that equation (4) is a single-peaked function of *s*. The benefits of investing increase when $\varphi(s)>1$, while the benefits are decreasing when $\varphi(s)<1$. Together with the fact that $\varphi(s)$ is non-increasing in *s* on the interval [0,1], we can conclude that $E(B(s))$ is single-peaked in *s*. When standards are very high or very low, the benefit of investing tends to equal zero. For very low standards the explanation is that a worker has a high chance of meeting the standards, even when he does not make the investment. In contrast when standards are very high, the chance of meeting the standard is very low independent on the investment decision.

Using equation (4) and the fact that the costs of investment are distributed according to some CDF G(c), it is possible to derive the proportion of investors. It is important to notice that, under assumption, investment costs are distributed identically in both groups. This implies that both W- and B-workers react in the same way for a given standard. The individual worker will become qualified when $c^i \leq E(B(s))$. The proportion of workers becoming qualified might differ among the two identifiable groups, according to equation (5) (where s^i denotes the standard faced by group *i*). Furthermore we assume that there are no workers with zero investment costs, such that the proportion of investors tends to go to zero when standards become very high or very low. The worker investment curve described by equation (5) is single-peaked as well.

$$\pi^i = G(E[B(s^i)]) \qquad (5)$$

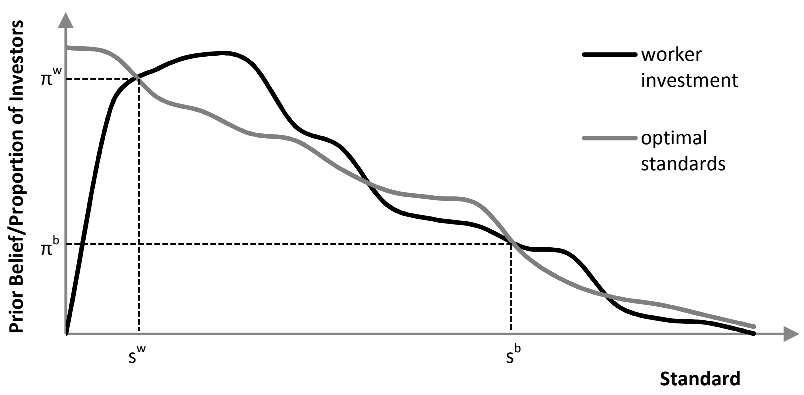
3.1.3 Equilibrium

In a self-confirming equilibrium, the prior belief held by the employer is exactly equal to the investment rate of the workers. In other words, in equilibrium the employer sets a standard consistent with his prior beliefs or experience about the proportion of qualified workers in group *i*. By setting the standard at the optimal level, the employer induces the workers from group *i* to become qualified at a rate exactly equal to his prior belief about the proportion of qualified workers. An equilibrium is therefore defined as a pair of beliefs (π^w, π^b) for which it holds that:

$$\pi^i = G\left(E\left[B\left(s^*(\pi^i)\right)\right]\right) \quad i{=}b{,}w \qquad (6)$$

It is helpful to draw the optimal standard curve and the worker investment curve in a (π,s)-diagram. Figure 1 shows the downward-sloping optimal standards curve and the single-peaked worker investment curve. According to the definition in (6), a self-confirming equilibrium arises when the optimal standard curve and the worker investment curve intersect. A discriminatory equilibrium, or negative stereotypes, might occur when there are multiple intersections between the two curves as in figure 1. As is shown in the figure employers believe that the proportion of investors among B's is lower than among W's ($\pi^w > \pi^b$). For that reason, employers set a stricter standard for B-workers ($s^b > s^w$). This reduces the expected benefit of investing for B-workers which results in a lower proportion of investors among B-workers compared with the investment rate among W-workers.

Figure 1: Equilibrium in the Coate and Loury model



In figure 1 four solutions of equation (6) can be distinguished, since the optimal standards curve and the worker investment curve intersect four times. However, only the equilibrium (π^w, π^b) lifted out in the analysis above is a discriminatory *locally stable* equilibrium. Local stability is an important concept in the analysis of negative stereotypes. Suppose that the employer's prior belief in period 0 ($\pi^{b,t=0}$) slightly deviates from the equilibrium belief (π^{b^*}). This deviation, in turn, leads to a slightly different standard and investment rate in period 0. This impacts the prior belief of the employer in period 1 ($\pi^{b,t=1}$)

etc. When the equilibrium is locally stable, the old equilibrium (π^{b^*}) is reached by the adjustment process described above. An equilibrium is locally stable under the condition that the absolute value of the slope of the optimal standards curve exceeds that of the worker investment curve. Local stability is important in our analysis, since only locally stable equilibria are likely to persist over time. When an equilibrium is not locally stable a small error made by the employer can completely break down a discriminatory equilibrium. Such non-stable equilibria are therefore much more likely to disappear over time. That is the reason why we will only focus on locally stable equilibria in the remaining part of the analysis.

3.2 A Model with Effort Incentives

The model of Coate and Loury described above forms the basis for our analysis on the impact of performance pay on negative stereotypes. Nevertheless, we need to change the model to some extent in order to be able to analyze the impact of different wage regimes. The Coate and Loury model simply assumes a gross gain (w) for each worker assigned to the high-skilled job. In addition the employer is assumed to earn x_q when assigning a qualified worker to the high-skilled job and to incur a loss equal to x_u when assigning an unqualified worker to the more demanding job.

We will make more specific assumptions on the gains and losses of workers and employers by incorporating a principal-agent model in the Coate and Loury model. The principal-agent model allows us to analyze two different wage regimes: a flat wage regime and a performance pay regime. We will distinguish between different situations and corresponding payoffs. Below we will start with an overview of the principal-agent model we incorporate in the model. After that, we will specify the optimal contracts offered under the different regimes. From the optimal contracts we are able to derive the payoffs received by workers and employers under the different wage regimes. These payoffs form the basis of the analysis in section 4.

3.2.1 Employers

We make the assumption that employers maximize their profits. The profits of assigning a worker to the high-skilled job are determined by price p, output y and the wage (w) paid to the qualified worker.

$$\Pi = py - w \qquad (7)$$

In addition to that we assume a simple linear relationship between output y and the effort exerted by the (qualified) worker: $y = \gamma e$. In this simple case the factor γ determines to what extent worker effort and output are related. High values of γ correspond to a very direct relationship, whereas low values of γ indicate that output is not necessarily the result of high worker effort. In the latter case, high worker effort does not necessarily result in high output. It might be the case that factors beyond the control of the worker, such as fortune or market conditions, play an important role in the output realization.

When the employer assigns a worker to the high-skilled job, profits do also depend on the qualifications of the worker. When the worker is unqualified, the effort exerted by the worker does not contribute to the profits of the employer. In that case, the employer faces a loss equal to the wage paid to the qualified worker. All in all we can distinguish between three different cases with corresponding payoffs. Payoffs depend on the job to which the worker is assigned and the investment decision of the worker.

1. The employer assigns a qualified worker to the high-skilled task. In this case, the effort exerted by the worker contributes to the profits of the employer:

$$\Pi_q^h = p\gamma e - w \qquad (8)$$

2. The employer assigns an unqualified worker to the high-skilled task. In this case, the effort exerted by the worker does not contribute to employer profits:

$$\Pi_u^h = -w \qquad (9)$$

3. The employer assigns a worker to the low-skilled task. The profits associated with this action are normalized to zero:

$$\Pi^l = 0 \qquad (10)$$

3.2.2 Workers

Workers are considered to be utility maximizing agents. The worker utility depends on the wage the worker receives (w), the effort he exerts (e) and the costs of the investment needed to become valuable in the high-skilled job (c^i). In the first place we recognize that exerting effort is costly to the worker. For that reason we assume a standard convex cost of effort function, which is the same for both qualified and unqualified workers. Next to that we assume that qualified workers in the high-skilled job are to some extent intrinsically motivated to exert effort.

There are clear analytical reasons for assuming intrinsic motivation among qualified workers. In a set-up without intrinsic motivation, workers are not willing to exert any effort under a flat wage regime. In that case a comparison between a flat wage and a performance pay regime would not be very interesting. Nevertheless, the concept of intrinsic motivation is not only introduced for analytical reasons. A situation in which none of the workers is willing to exert effort does not seem to be very realistic. One can imagine that a qualified worker actually likes to work in the more demanding task. It could for instance give him the feeling that he is important and it could give rise to feelings of self-esteem. In contrast, unqualified workers are less likely to love working in the qualified job. The job might be too demanding, which demotivates them to exert some effort.

Even when feelings of importance and self-esteem do not play a role, workers are always facing an indirect monetary incentive, even when they work under a flat wage regime. Qualified workers choose to exert effort under a flat wage regime, since they fear that they get fired when exerting zero effort. For unqualified workers this does not hold. Even when they exert effort, it does not contribute to the profits of the employer. For that reasons we assume that only qualified workers have some intrinsic motivation to exert effort.

We model this intrinsic motivation by assuming a certain level of effort at which the qualified worker feels 'comfortable'. Effort below or above this level is costly. The intuition behind this way of modeling is twofold. At one hand, when effort is too low workers fear that they get fired and are willing to increase effort. At the other hand, when effort comes above a certain level, the cost-of-effort-effect takes over and exerting extra effort becomes costly.

Now we can distinguish between four different cases with corresponding payoffs. Again, payoffs depend on the job to which the worker is assigned and the investment decision made by the worker.

1. A qualified worker gets assigned to the high-skilled job. This worker is to some extent intrinsically motivated to exert effort:

$$U_q^h = w - \frac{1}{2}(e-1)^2 - c_i \qquad (11)$$

2. A qualified worker gets assigned to the low-skilled job. The gross utility of getting assigned to the low-skilled job is normalized to zero. However, the worker still faces the investment costs, c_i:

$$U_q^l = -c_i \qquad (12)$$

3. An unqualified worker gets assigned to the high-skilled job. For this worker effort is just costly and there is no intrinsic motivation to exert effort:

$$U_u^h = w - \frac{1}{2}e^2 \qquad (13)$$

4. An unqualified worker gets assigned to the low-skilled job:

$$U_u^l = 0 \qquad (14)$$

3.2.3 Optimal Contracts

Now that we know the profit and utility functions, we can determine the optimal contracts under the different wage regimes. Before we are able to do so, we need to define the different wage regimes. We distinguish between two different wage regimes: a flat wage regime and a performance pay regime. Under a flat wage regime workers in the high-skilled job earn a flat wage, which does not depend on the level of effort they exert. Workers determine their optimal level of effort and employers determine the optimal level of the wage. We will see that the optimal level of effort under a flat wage regime will not depend on the wage the worker receives.

Under a performance pay regime, workers in the high-skilled job are (partly) paid according to their productivity. Workers receive a base wage, α, which is independent on the level of effort they exert. In addition workers can earn a bonus which depends on the level of effort they exert. Again we will show how workers determine their optimal level of effort and how the employers respond by defining the optimal contract.

Flat wage regime

The optimal levels of effort under a flat wage regime are given by maximizing equations (11) and (13) with respect to *e*. The optimal level of effort depends on the investment decision of the worker:

$$e_q^* = 1 \qquad (15)$$

$$e_u^* = 0 \qquad (16)$$

We see that qualified workers are motivated to exert some positive level of effort under a flat wage regime. This is caused by the assumption we made about the intrinsic motivation of qualified workers. Unqualified workers are not inclined to exert some positive amount of effort.

In this set up the level of effort exerted by the workers is not affected by the level of the flat wage, w. As a consequence the employer optimally chooses to set the wage as low as possible. In that case the profits of assigning a qualified worker to the high-skilled job (8) are maximized and the loss of assigning an unqualified worker to the high-skilled job (9) is minimized. In fact, this means that the employer will set the wage at the legally determined minimum level or at the level determined by collective labour agreements. In such a case we can state that the limited liability constraint is binding. In the analysis below, $\overline{w}$ indicates the optimal level of the flat wage. To be sure that the employer makes a profit on assigning a qualified worker to the high-skilled job, we assume that $p\gamma - \overline{w} > 0$.

Performance pay regime

We first need to specify the wages which qualified and unqualified workers can earn in the high-skilled job under a performance pay regime. Both qualified and unqualified workers receive the base salary, α. However, only qualified workers are eligible for the effort dependent bonus. Once assigned to the high-skilled job, employers evaluate the performance of the worker using some objective performance measure. This enables us to abstract from discriminatory practices at the performance evaluation stage.

$$w_q^h = \alpha + \beta e \qquad (17)$$

$$w_u^h = \alpha \qquad (18)$$

Inserting equations (17) and (18) into equations (11) and (13) and maximizing with respect to e yields the optimal effort levels for qualified and unqualified workers in the high-skilled job under performance pay:

$$e_q^* = 1 + \beta \qquad (19)$$

$$e_u^* = 0 \qquad (20)$$

The optimal level of effort for the qualified worker increases in the level of the incentive intensity parameter β. The higher monetary incentives, the higher the level of effort exerted by the qualified worker. Next to that we see that the optimal level of effort equals 1 in absence of monetary incentives. This corresponds to the flat wage case, in which monetary incentives do not play a role. The optimal level of effort for the unqualified worker does not change compared to the flat wage case. Unqualified workers are not influenced by monetary incentives since their effort is not valued by the employer. For that reason, they are not able to earn the bonus provided under the performance pay regime.

Under a performance pay regime the employer is able to affect the level of effort exerted by the qualified worker. The employer can use the incentive intensity parameter β as a tool to influence that level of effort. By inserting the optimal level of effort (19) and the formula for the wage contract (17) into equation (8), the employer is able to determine the optimal level of β.

$$\beta^* = \frac{p\gamma - 1}{2} \qquad (21)$$

The first conclusion we can make is that not all firms are willing to provide positive monetary incentives. For some firms, the effort exerted as a result of the intrinsic worker motivation is sufficient. In some cases firms even want to restrict worker effort by providing negative monetary incentives. If we fix price p at p^*, we observe that only those firms will provide positive monetary incentives for which it holds that:

$$\gamma > \frac{1}{p^*} \qquad (22)$$

When the relationship between effort and output is strong enough, employers optimally provide a bonus related to the effort exerted by the (qualified) worker. That is, when γ is high enough, employers are able to increase profits by providing a bonus to qualified workers. In this study we are interested in the effects of *performance pay* on negative stereotypes. For that reason we limit the analysis to firms for which equation (22) holds. Only when incentives are positive, we can talk about *performance pay*. In the next part, in which the two wage regimes are compared, we will prove that only when equation (22) holds firms are willing to switch to performance pay. We believe that the effect of the effort-output relationship on negative stereotypes is the most interesting effect. Therefore we will continue to fix the price at p^* to lift out the effect of the parameter γ in subsequent sections.

In order to formulate the optimal contract, we also need to make a statement about the optimal base salary $\bar{\alpha}$. For the employers, there is no reason to set $\bar{\alpha}$ at a level above the minimum wage to be paid under a flat wage regime. Therefore we state that $\bar{\alpha} \leq \bar{w}$. In the remaining part of the analysis we will assume that it is legally not allowed to set the base salary under performance pay at a level below the legally or collectively determined minimum wage. In that case the employer optimally sets $\bar{\alpha}$ at the same level as $\bar{w}$.

Comparison

With the optimal contracts under the two wage regimes in the back of our minds, we can make a comparison. From the previous part we know that firms switching to performance pay will set $\beta^*>0$. This results in higher worker effort under performance pay as compared to the flat wage regime. As a result of higher worker effort, employer revenue (price times output) increases. However, wages paid to qualified workers in the high-skilled job do increase as well. We assumed that $\bar{\alpha}$ is set at the same level as $\bar{w}$ and that qualified worker can earn an extra bonus equal to $\beta^* e_q^*$. This makes the total effect on employer profits uncertain.

In Appendix A we prove that employer profits associated with assigning a qualified worker to the high-skilled job are strictly higher under a performance pay regime as compared to the flat wage situation. This result holds under the condition given by equation (22). Only when γ is high enough, firms are able to increase profits when they switch to performance pay. For other firms, switching to performance pay is not a serious option, since it will result in a decline in profits. The employer loss associated with assigning an unqualified worker to the high-skilled job does not change, since $\bar{\alpha} = \bar{w}$. This is an important result for the analysis in section 4.

For qualified workers payoffs are different as well under the different wage regimes. Under performance pay, qualified workers earn higher wages in the high-skilled job. However, they exert a higher level of effort as well. Since effort is, according to equation (11), costly above some specified level, we need to check which of the two effects dominates. In Appendix B it is proven that the utility of qualified workers working in the high-skilled job increases under performance pay. Since $\bar{\alpha} = \bar{w}$, the utility of unqualified worker assigned to the high-skilled job remains unchanged. They earn the same wage and optimally exert the same level of effort.

4. Analysis

In this section we use the results derived from our model with effort incentives to study the effects of performance pay on negative stereotypes. The model with effort incentives enables us to analyze different wage regimes within the standard Coate and Loury model. This section is divided into five subsections.

The first two subsections are on optimal standards and worker investment. These two subsections are divided into three parts. We start with the results we obtain under the two different wage regimes. We use the optimal contracts under both wage regimes and the corresponding payoffs to find expressions for optimal standards and worker investment in our model with effort incentives. At the end of each subsection we compare the outcomes of our model to the outcomes of the Coate and Loury model. More importantly, we analyze the differences between the two wage regimes and we show how performance pay impacts negative stereotypes. These two subsections end with an proposition about the impact of performance pay on optimal standards and worker investment.

The third subsection is on the equilibrium of our model with effort incentives. In this subsection we take a discriminatory equilibrium with flat wages as the starting point. The reason is that we are interested in the impact of a switch from flat wages to performance pay on negative stereotypes. Given our starting point, we analyze possible impacts of performance pay on negative stereotypes. We start with the sole effect of optimal standards, followed by the effect of worker investment. We end with the combined equilibrium effect of the employer and worker decision on negative stereotypes. Next to that we take a closer look at the impact of the effort-output relationship (γ) on the likelihood that negative stereotypes are eventually eliminated under performance pay. This subsection ends with two propositions about the impact of γ on the likelihood that negative stereotypes are eliminated under performance pay.

The fourth subsection is about the social consequences of switching to performance pay. Is a switch from flat wages to performance pay socially optimal? Even when negative stereotypes are reduced or eliminated, we need to check whether the overall investment rate in society is maintained. In the last subsection we try to answer the question whether social interests are always aligned with the interests of the individual firm. We show that a switch to performance pay might be optimal from a social point of view, while it might not be optimal from the employer's perspective.

4.1 Optimal standards

Flat wage

Recall from the previous section that employers pay a flat wage to all workers who get assigned to the high-skilled job under a flat wage regime. From equations (8) and (15) and the assumption on the wages paid to the worker we can determine the profits obtained by the employer when he assigns a qualified worker to the high-skilled job. The loss associated with assigning an unqualified worker to the high-skilled job equals the wage paid to the worker. In terms of the Coate and Loury model we can state that $x_q = p\gamma - \bar{w}$ and $-x_u = \bar{w}$. r^f is defined as the ratio x_q/x_u when workers are paid a flat wage. From equation (3) we are able to derive the following expression for the optimal standards under a flat wage regime:

$$s^*(\pi^i) = \min\left\{\theta \in [0,1] | r^f = \frac{p^*\gamma - \bar{w}}{\bar{w}} \geq \left(\frac{1-\pi^i}{\pi^i}\right)\varphi(\theta)\right\} \tag{23}$$

Performance pay

In the same vein we can check how employers react under a performance pay regime. From the previous section we know the employer profits (losses) associated with assigning a qualified (unqualified) worker to the high-skilled job. In this case $x_q = p^*\gamma(1+\beta^*) - [\bar{\alpha} + \beta^*(1+\beta^*)]$ and $-x_u = \bar{\alpha}$. We define r^f as the ratio x_q/x_u under performance pay and substitute this in equation (3) to find the following expression for optimal standards under performance pay:

$$s^*(\pi^i) = \min\left\{\theta \in [0,1] | r^p = \frac{p^*\gamma(1+\beta^*) - [\bar{\alpha} + \beta^*(1+\beta^*)]}{\bar{\alpha}} \geq \left(\frac{1-\pi^i}{\pi^i}\right)\varphi(\theta)\right\} \tag{24}$$

Comparison

We observe that the general shape of the optimal standards curve is identical to the downward sloping optimal standards curve of the standards Coate and Loury model. The only thing that changes is that we have specified the level of the payoffs under different wage regimes. Instead of fixing payoffs at x_q and x_u, we found payoffs under the two wage regimes dependent on the parameters of the model (p, γ, $\overline{w}$, $\bar{\alpha}$ and β).

We are able to compare the optimal standards curves under the two wage regimes using the results from the previous section. We know that both revenue (price times output) and wage costs increase under performance pay. But we also know from Appendix A that revenue increases more than wage costs. This implies that the gain associated with assigning a qualified worker to the high-skilled task is higher under performance pay as compared to that gain under a flat wage regime. Next to that, the loss of assigning an unqualified worker to the high-skilled job is equal under both wage regimes. This is under the assumption that $\bar{\alpha} = \overline{w}$. These results have their impact on the ratio x_q/x_u under the two regimes. r^p, the ratio x_q/x_u under performance pay is strictly greater than r^f, the same ratio under a flat wage regime.

We still have to check the impact of the different wage regimes on optimal standards. From equation (3) it follows that the higher the ratio r, the lower optimal standards for any prior belief π^i. This means that optimal standards are strictly lower under performance pay than under a flat wage regime for any prior belief π^i. Optimal employer behavior is described by the optimal standards curve derived in section 3 (figure 1). Comparing the expressions for the optimal standards under the two wage regimes, we can conclude that the optimal standards curve under a performance pay regime lies strictly below the flat wage optimal standards curve. When an employer shifts from flat wages to performance pay, the optimal standards curve shifts inward.

The magnitude of this inward shift is determined by the effort-output relationship γ. In Appendix C it is proven that the ratio r^p is increasing in γ. This is completely explained by the fact that the gain of assigning a qualified worker to the high-skilled job increases in γ. The stronger the effort-output relationship, the higher the gains, in terms of profits, from assigning a qualified worker to the high-skilled job. However, the loss from assigning an unqualified worker to the high-skilled job is not affected by γ. Consequently, optimal standards are decreasing in γ for any prior belief π^i. The stronger the effort-output relationship, the more severe the inward shift of the optimal standards curve when the employer switches from a flat wage regime to performance related pay. This leads to the following proposition about the difference in employer behavior under the two wage regimes.

PROPOSITION 1: *Under performance pay optimal standards* $s^*(\pi^i)$ *are lower for any prior belief* π^i *compared to optimal standards under flat wages, given that* $\bar{\alpha} \leq \bar{w}$ *and* $p^*\gamma > 1$. *In addition, higher values of* γ *correspond to larger decreases in optimal standards for any prior belief* π^i.

4.2 Worker investment

Flat wage

In section 3 we explained how the worker investment decision is related to optimal standards and the expected benefit from investing. Under a flat wage regime the worker can earn the wage $\bar{w}$ in the high-skilled job. Optimal effort levels for qualified and unqualified workers are given by equations (15) and (16). Inserting the wage and the levels of effort into the worker utility functions yields $U_q^h = \bar{w} - c_i$ and $U_u^h = \bar{w}$. Using the utility levels from the optimal contracts we are able to derive the investment decision of the worker. As explained in section 3 equation (25) gives the expected benefit from investing, excluding investment costs. The proportion of workers becoming qualified can be derived from equation (25) using equation (5) in the previous section.

$$E[B(s)] \equiv \bar{w}\left[F_u(s) - F_q(s)\right] \qquad (25)$$

Performance pay

Under performance pay all workers who get assigned to the high-skilled job receive the base salary $\bar{\alpha}$. However, only qualified workers are able to earn the bonus which depends on the level of effort exerted by the worker. The optimal levels of effort under this type of contract are given by equation (19) and (20). The corresponding utility levels are equal to $U_q^h = \bar{\alpha} + \beta^*(1 + \beta^*) - \frac{1}{2}\beta^{*2} - c_i$ for qualified workers and $U_u^h = \bar{\alpha}$ for unqualified workers. The worker is willing to invest when the expected utility associated with investing exceeds the utility level he expects to obtain when he does not make the investment. Again we derive the expected benefit from the investment, excluding the investment costs (equation 26). From equation (26) we can derive the proportion of workers becoming qualified using equation (5).

$$E[B(s)] \equiv \bar{\alpha}\left[F_u(s) - F_q(s)\right] + \left[1 - F_q(s)\right]\left(\beta^* + \frac{1}{2}\beta^{*2}\right) \qquad (26)$$

Comparison

Basically, the worker investment decision under flat wages is similar to the investment decision in the standard Coate and Loury model. Under the assumptions about the signaling technology, the worker investment curve will have the same shape as the one in figure 1. When performance pay is introduced, the proportion of investors changes drastically. We made the assumption that the base salary (α) under performance pay equals the flat wage $\bar{w}$. That is, the first part of equation (26) is identical to equation (25), which points to an equal proportion of investors for all standards.

However, under performance pay there is a wage difference between qualified and unqualified workers. Qualified workers are able to earn a bonus once they are assigned to the high-skilled job. In contrast, the effort exerted by an unqualified worker is not valued by the employer which ensures that unqualified workers never receive the bonus. Under performance pay workers have an additional reason to become qualified. It is not only the increased probability of getting assigned, it is also the bonus payment which induces workers to become qualified. Since only qualified workers are able to earn the bonus, the expected benefit from becoming qualified goes up. This is represented by the second part of equation (26). This part of the equation shows that the benefit from investing (excluding investment costs) gets an extra boost, only for qualified workers.

We are now able to consider the impact of the different wage regimes on the proportion of investors. Under flat wages the proportion of investors follows a similar pattern as in the standard Coate and Loury model. Both at very low and at very high standards the proportion of investors tends to equal zero. At low standards, the probability of assignment is always high, even when the worker is unqualified. At very high standards, the probability of assignment is very low, even when the worker made the investment.

Under performance pay the situation is very different. Now some workers are willing to invest, even when the standard equals zero. The second part of equation (26) is strictly positive at any standard. The probability that a worker meets the standard $(1 - F_q(s))$ tends to equal 1 when standards are low. So when the worker makes the investment, it is almost certain that he gets assigned to the high-skilled job and that he earns the bonus. Now consider the situation in which the worker does not make the investment. At low standards it is still very likely that he gets assigned to the high-skilled job. However, in this situation the worker does not earn the bonus because he is unqualified and his effort is not valued by the employer. At very high standards, the probability that the worker gets assigned after making the investment tends to go to zero. In that case, it is very uncertain whether the worker receives the bonus after making the investment and this reduces investment incentives. However, since we assumed that $\bar{\alpha} = \bar{w}$, the proportion of investors under a performance pay regime would never fall below the investment rate under a flat wage regime.

The worker investment curve will shift outward when an employer switches from flat wages to performance pay. For any standard, more workers are willing to make the investment as a result of the monetary incentive provided to qualified workers. In Appendix C we check how the effort-output relationship γ affects the magnitude of this outward shift. We find a positive relationship between γ and investment incentives. The stronger the relationship between effort and output, the higher monetary incentives. This, in turn, leads to an increase in investment incentives. The bonus to be gained becomes higher which induces extra workers, with higher investment costs, to make the investment. All in all we can formulate the following proposition:

PROPOSITION 2: *Under performance pay the proportion of workers making the investment $G\left(E\left[B\left(s^i\right)\right]\right)$ is higher for any standard $s\left(\pi^i\right)$ than under a flat wage regime, given that $\bar{\alpha} = \bar{w}$ and $p^*\gamma > 1$. In addition, higher values of γ correspond with stronger increases in the proportion of investors for any standard $s\left(\pi^i\right)$.*

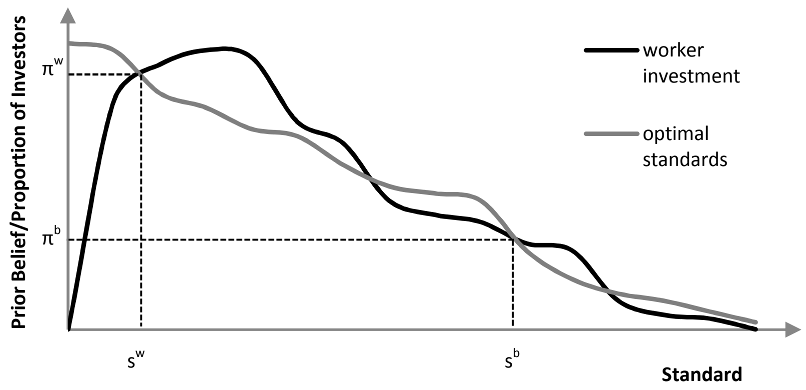
4.3 Equilibrium

4.3.1 Flat wage

What do the results derived above mean for equilibrium? We want to check the impact of performance pay on negative stereotypes within the Coate and Loury model on statistical discrimination. The only way to answer this question is assuming a baseline situation characterized by a wage regime different than performance pay and discrimination in equilibrium. We take the flat wage regime analyzed above as the starting point for our analysis and we assume negative stereotypes in equilibrium. Starting from that situation we analyze the impact of performance pay on optimal standards and worker investment. From the movements of the optimal standards curve and the worker investment curve we are able to derive the relationship between performance pay and negative stereotypes.

We found that the expressions for optimal standards and worker investment under a flat wage regime do not differ conceptually from those in the standard Coate and Loury model. Therefore we take figure 2 below, which is an copy of figure 1 in the previous section, as the starting point for the analysis. In figure 2 we observe the standard downward sloping optimal standards curve. Next to that we see that the proportion of investors tends to equal zero both at very high and very low standards. In equilibrium B's face higher standards than W's, which induces B's to become qualified at a lower rate than W's. The employer holds negative stereotypes against B's and his beliefs are confirmed in equilibrium.

Figure 2: Equilibrium under a flat wage regime



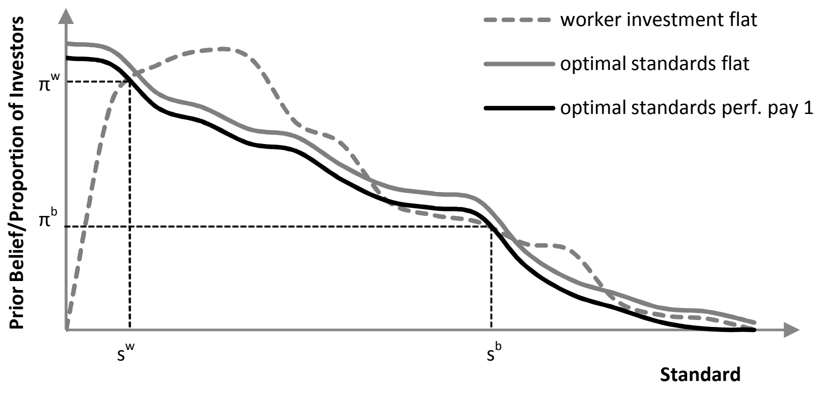
4.3.2 Performance pay

Optimal standards

We know from proposition 1 that optimal standards under performance pay are lower for any prior belief π^i. The magnitude of the inward shift of the optimal standards curve is determined by the effort-output relationship. We start analyzing the impact of a relatively minor inward shift of the optimal standards curve, holding the worker investment decision constant for the moment. Figure 3 shows that the investment rates of B's and W's converge compared to the flat wage situation. Both B's and W's face lower standards in the new situation. This results in a higher investment rate among B's, while W's invest at a lower rate.

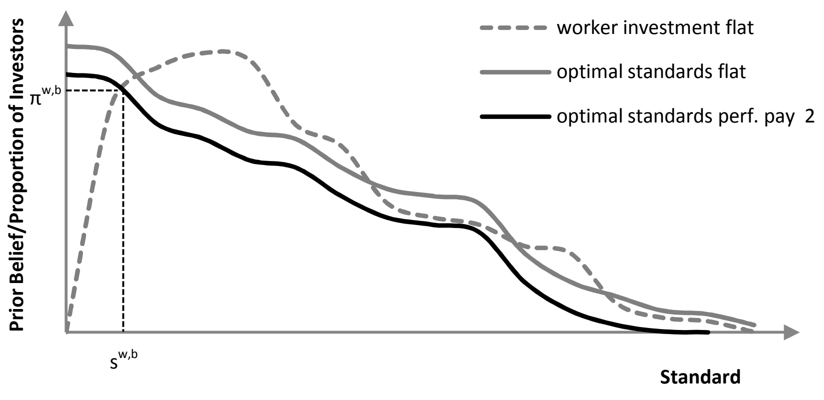
Negative stereotypes are not eliminated in this situation. Employers continue to hold negative stereotypes against B's. Since the relationship between effort and output is not very strong, employers are not able to gain very much in terms of profit from the introduction of performance pay. This induces employers to set standards relatively close to the standards under a flat wage regime. However, the absolute level of discrimination ($\pi^w - \pi^b$) reduces since investment rates and employer beliefs about the two groups converge.

Figure 3: optimal standards under performance pay



When the relationship between effort and output is relatively strong, we know from proposition 1 that we can expect a more severe inward shift of the optimal standards curve. This situation is shown in figure 4. In this case only one locally stable equilibrium persists. This means that negative stereotypes are completely eliminated as a result of the introduction of performance pay, holding the worker investment decision constant. Employers hold one and the same belief about the proportion of investors in both groups, resulting in equal standards for the two types of workers ($s^b = s^w$). Workers from both groups confirm this belief in equilibrium by choosing to become qualified at the same rate ($\pi^b = \pi^w$).

Figure 4: optimal standards under performance pay

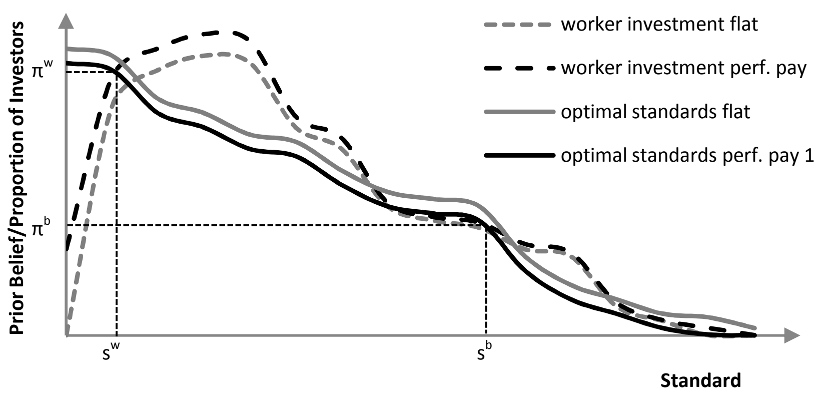


It is not likely that the new equilibrium is reached immediately. Employer beliefs about the proportion of investors in the two groups do not change instantly after the introduction of performance pay. However, these beliefs are no longer confirmed by the investment rates of the workers. For example, the investment rate of B's will be higher than expected by the employer. As a result standards will change, which induces worker to invest at another rate etc. Following this logical adjustment process, the new equilibrium will be reached.

Worker investment

Recall from section 4.2 that worker investment rates will be higher under performance pay for any standard $s(\pi^i)$. The outward shift of the worker investment curve is shown graphically in figure 5. Next to that we know that some workers are willing to invests, even when the standard set by the employer equals zero. In figure 5 we assume a relatively weak relationship between effort and output, resulting in a moderate inward shift of the optimal standard curve. This is basically the same situation as in figure 3.

Figure 5: optimal standards under performance pay



Starting from the discriminatory equilibrium in figure 3 we adjusted the worker investment curve to the new situation. Investment incentives go up for any standard, because of the bonus a qualified worker receives in the high-skilled job under the performance pay regime. The effect of the increased investment incentives is intuitive. Workers from both groups become qualified at a higher rate.

The overall effect on discrimination is uncertain in this case. Negative stereotypes continue to exist when the relationship between effort and output (γ) is relatively weak. As a result of lower standards, the investment rate of W's decreases while that of B's increases. However, the increase in investment incentives leads to a higher investment rate among both groups. When the effort-output relationship is relatively weak, the introduction of performance pay certainly leads to a higher investment rate among B's. The overall effect on the investment rate among W's is uncertain. Therefore it is uncertain whether or not absolute discrimination ($\pi^w - \pi^b$) is reduced under performance pay for low values of γ. These results might be important for policy makers. When the introduction of performance pay is stimulated as a means to reduce discrimination it is important to realize that performance pay does not necessarily decrease or eliminate discriminatory practices. It might even backfire in the sense that investment gap between two groups of workers increases.

But what happens when the effort-output relationship γ is a little bit stronger? We know the effect of γ on the optimal standards curve and the worker investment curve. A stronger relationship between effort and output leads to an inward shift of the optimal standards curve, while the worker investment curve shifts outward. These two effects together work in the same direction, namely the complete elimination of negative stereotypes. When γ is large enough we get a situation identical to the situation depicted in figure 4. All locally stable equilibria but the one on the upward sloping part of the worker investment curve are eliminated. The new fraction of investors is the same in both groups and might be above the old π^w as well as in between the old π^w and π^b. That is, the investment rate of B's goes up, while the effect on the fraction of W's becoming qualified is uncertain. We summarize the results we find above in the following propositions:

PROPOSITION 3: *Starting from a discriminatory equilibrium under flat wages, there exists a* $\gamma^* > \frac{1}{p^*}$*, such that if* $\gamma > \gamma^*$ *negative stereotypes are completely eliminated when the firm moves to performance pay.*

PROPOSITION 4: *Starting from a discriminatory equilibrium under flat wages, with* $\gamma^* > \gamma > \frac{1}{p^*}$ *negative stereotypes persist when the firm moves to performance pay. The investment rate of the disadvantaged group increases, while the total effect on the fraction of investors in the advantaged group is uncertain. Therefore, the absolute level of discrimination* ($\pi^w - \pi^b$) *might either increase or decrease.*

The mechanism behind this relationship between γ and the likelihood that the negative stereotypes are eliminated is intuitive. The stronger the relationship between effort and output, the more an employer gains from the introduction of performance pay. After all, the effort of the worker is more closely related to the profits of the employer, which induces the employer to choose for higher monetary incentives. Despite the higher wage costs of hiring a qualified worker, we have proven that the profits associated with assigning a qualified worker increase in γ. Holding the loss of assigning an unqualified worker constant, optimal standards decrease in γ for any prior belief.

In this model employers are a priori color-blind. However, based on experience they (correctly) suspect a correlation between group identity and productivity. The fact that payoffs change when switching from flat wages to performance pay, induces them to change optimal behavior. This, in turn, impacts worker investment decisions. Workers face lower standards for given prior beliefs. Next to that, the bonus structure under performance pay provides an extra incentive for worker to become qualified. The higher γ, the higher the bonus payment to be earned and the more severe the increase in investment incentives.

Total elimination of negative stereotypes becomes more likely when the relationship between effort and output is stronger for the reasons described above. Employers are able to gain more from assigning a qualified worker while the loss associated with assigning an unqualified worker does not change. Similarly, workers are able to gain more from getting assigned to the high-skilled job. Both effects increase in γ and contribute to the elimination of negative stereotypes.

4.4 Social optimality

In the analysis above we did not consider the social consequences of the introduction of performance pay. A discriminatory equilibrium as shown in figure 2 is socially inefficient. To see this we compare the two self-confirming beliefs π^w and π^b. We see that $\pi^w > \pi^b$, which implies that $s^*(\pi^w) < s^*(\pi^b)$. Comparing π^w and π^b we can conclude that both workers and employers are better off in π^w. Workers have a higher chance of getting assigned to the high-skilled job and employers hire workers from a pool with a higher percentage of qualified workers.

We define a switch to performance pay to be socially efficient when the average proportion of investors in society increases. If that is the case, both workers and employers are on average better off. Suppose that the relationship between effort and output is relatively weak. In that case we know that negative stereotypes are not completely eliminated. The introduction of performance pay is socially efficient if and only if equation (27) holds. We know that the proportion of investors among B's increases as a result of performance pay ($\pi_p^b > \pi_f^b$). The effect on the investment rate among W's is uncertain, but suppose that W's become qualified at a lower rate than before ($\pi_p^w < \pi_f^w$). It follows from (27) that the introduction of performance pay is only efficient when the proportion of B's in the population

$(1-\lambda)$ is large enough. In case that the investment rate among W's increases as well, the introduction of performance pay is always socially efficient.

$$\lambda\pi_p^w + (1-\lambda)\pi_p^b > \lambda\pi_f^w + (1-\lambda)\pi_f^b \qquad (27)$$

Now we turn to the other scenario we described above, with a strong relationship between effort and output. In that case all locally stable equilibria are eliminated except the one on the upward sloping part of the worker investment curve. In that case it is uncertain how the new investment rate relates to the old investment rate in society. Suppose that the new investment rate for worker from both groups lies in between the two old rates ($\pi_f^w > \pi_p^{w,b} > \pi_f^b$). Equation (28) gives the condition under which this situation is socially efficient. Again, social efficiency is only possible when the share of B's in the population $(1-\lambda)$ is large enough. In the other possible situation with $\pi_p^{w,b} > \pi_f^w$, the introduction of performance pay is always social efficient.

$$\pi_p^{w,b} > \lambda\pi_f^w + (1-\lambda)\pi_f^b \qquad (28)$$

4.5 The Individual Firm

It is interesting to see that a socially optimal outcome is not always optimal for the individual firm or employer. In this subsection we look at the change in profits of the employer when moving from flat wages to performance pay and we show that there is a *composition effect*. In Appendix D we derive the employer profits in equilibrium. Optimal standards for the two groups of workers as well as the investment rates among the two groups influence profits in equilibrium. Next to that the profits of assigning a qualified worker and the loss associated with assigning an unqualified worker play a role. The last term that determines equilibrium profits of an individual employer is the composition of the workforce of the firm.

We have seen that different outcomes are possible when moving from a flat wage regime to performance related pay. These outcomes are formulated in propositions 3 and 4. In all possible cases the investment rate of B's goes up, while the effect on the investment rate of W's is uncertain. From equation (D3) in Appendix D we know the factors that determine equilibrium profits. In Table 1 we summarize in which direction these factors change when switching from flat wages to performance pay.

$\left(1 - F_q\left(s^*(\pi^w)\right)\right)$	$\left(1 - F_u\left(s^*(\pi^w)\right)\right)$	$\left(1 - F_q\left(s^*(\pi^b)\right)\right)$	$\left(1 - F_u\left(s^*(\pi^b)\right)\right)$	π^w	π^b	x_q	x_u
+	+	+	+	+ / -	+	+	=

Table 1: the change in parameters defining profits when switching to performance pay.

From Table 1 we are able to analyze the impact of a switch to performance pay on equilibrium profits. We can divide the workers into four different groups based on group identity (W or B) and the investment decision (qualified or unqualified). Next to that we take into account that the proportion of W's becoming qualified might either increase or decrease.

For instance, take a look at the group of qualified W-workers and assume that the proportion W's becoming qualified decreases after switching to performance pay. Since standards are in any case lower under performance pay, the probability that a qualified worker gets assigned to the high-skilled job increases. Next to that, the profit associated with assigning a qualified worker is higher under performance pay than under a flat wage regime. These two effects have a positive effect on equilibrium profits. However, the fact that W's become qualified at a lower rate has a negative impact on equilibrium profits. Therefore, the total effect of qualified W-workers on equilibrium profits is uncertain. In case of an increase in the proportion of W's becoming qualified after moving to performance pay, the total effect of qualified W-workers is strictly positive.

Following the same analysis for the three other groups we find the contribution of these groups to equilibrium profits. A '+' ['-'] in Table 2 means that the contribution of the group to employer profits increases [decreases] when switching to performance pay. The only certain contribution to profits is that of the qualified B-workers. After moving to performance pay B's face a lower standard, they become qualified at a higher rate and the payoff associated with assigning a qualified B-worker increases.

$\Delta\pi^w$	Qualified W-workers	Unqualified W-workers	Qualified B-workers	Unqualified B-workers
> 0	+	+ / -	+	+ / -
< 0	+ / -	-	+	+ / -

Table 2: the impact on profits after switching to performance pay per group of workers.

Suppose that a switch to performance pay leads to a decrease in the proportion of W's becoming qualified, but that this is offset by an increase in the investment rate of B's. That is, the introduction of performance pay is optimal from a social perspective according to equation (27). We can ask ourselves the question whether the shift to performance pay is also optimal from the perspective of the individual employer. Remember that workers and employers are randomly matched with each other. With a large number of workers and firms we know that the average or representative firm hires a proportion of W's equal to λ. We consider a situation in which the representative firm observes an increase in equilibrium profits after moving to performance pay.

As a result of the random matching process there is some dispersion in the composition of workforces of the firms. We consider a firm with a relatively high fraction of W's in its workforce (λ in equation (D2) is close to 1) and compare this firm to the representative firm. From table two it follows that the contribution of qualified W's to equilibrium profits either increases or decreases when moving to performance pay, given that the proportion of W's becoming qualified decreases. The contribution of unqualified W's decreases when it is assumed that the proportion of investors among W's decreases. The intuition behind this result is that unqualified W's get more easily assigned because of lower standards, while the pool of unqualified W's increases.

Suppose furthermore that the total effect on W-workers is negative. That is, under performance pay, the contribution of W's to equilibrium profits is lower than under flat wages. Since we assumed that the representative firm observes an increase in profits when moving to performance pay, we know that the contribution of B's to profits must be positive. The implication of this results is that the profits of the firm with a relatively high fraction of W's are strictly lower than the profits of the representative firm under performance pay. It might even be the case that a firm with a relatively high fraction of W's faces a decrease in profits when moving to performance pay. This is what we call the composition effect, in some cases it depends on the composition of the firm's workforce whether or not profits increase after the introduction of performance pay.

All in all we can conclude that there might exist a situation in which social interests are not aligned with the interests of all individual firms. We explained how the workforce composition of an individual might differ from the workforce composition of the representative firm. In our example we showed that a firm with a relatively high fraction of W's might observe a decrease in profits when moving to performance pay, while the representative firm's profits increase. The composition effect accounts for this result.

5. Conclusion

We asked ourselves the question whether the introduction of performance pay could reduce or eliminate discrimination at the assignment stage of the labor market. This question is analyzed in a model of statistical discrimination comparable to that of Coate and Loury (1993). Employers in these type of models do not dislike a certain group of workers per se. Discrimination arises because of negative beliefs held by employers about the productivity of an identifiable group of workers. We hypothesized that performance pay might help to reduce or eliminate discrimination. The reasoning behind this hypothesis was that employers would set less strict standards, while workers investment incentives are higher under a performance pay regime. The combination of these two effects would decrease the likelihood that a discriminatory equilibrium persists under performance pay.

In section 3 we started with a review of the standard Coate and Loury model. We discussed employer and worker behavior and we explained the equilibrium concept of the model. The basic assumptions of the model do not change when we introduced effort incentives using a standard principal-agent model. We specified employer profits and worker utility. We assumed that employer profits depend on the price, output and wage costs. Output, in turn, is affected by the effort-output relationship. Worker utility depends on the wage, the cost of effort and investment costs. In addition we assumed some form of intrinsic motivation which guarantees that qualified workers are willing to exert effort even in absence of direct monetary incentives. Employers maximize profits, whereas workers maximize utility in our model.

The principal-agent model enabled us to analyze different wage regimes. We derived optimal contracts under a flat wage regime and a performance pay regime. The payoffs resulting from the optimal contracts are used to analyze the impact of performance pay on negative stereotypes. This is where section 4 starts. We derived expressions for optimal standards and for worker investment under the two wage regimes. These conditions are dependent on the payoffs derived in section 3. We analyzed the differences between the expressions under the two wage regimes. The results are summarized in two propositions. Under performance pay employers set lower standards for any prior belief about the productivity of a group of workers. Next to that, the investment incentives of the workers went up since they can earn an extra bonus under performance pay once they become qualified.

We continue with the effect of the changes in optimal employer and worker behavior on equilibrium outcomes of the model. We are interested in the question whether performance pay can lead to elimination of negative stereotypes. For that reason we presented a discriminatory equilibrium characterized by flat wages as the starting point of our analysis. We showed how the introduction of performance pay impacts the equilibrium outcome of the model. The less stringent assignment standards in combination with higher investment incentives among workers leads to a new equilibrium. In the new equilibrium a reduction in the absolute level of discrimination is possible, but uncertain.

The total effect depends on the parameters of the model. We are particularly interested in the effect of the effort-output relationship on the likelihood that negative stereotypes are eliminated. We found that a sufficiently strong relationship between effort and output goes hand in hand with total eradication of negative stereotypes. This result is formulated in the third proposition and it shows that total eradication of negative stereotypes is possible when switching to performance pay. However, when the link between effort and output is not that strong, negative stereotypes persist. This is formulated in the fourth proposition which states that discrimination persist and might even get worse when the link between effort and output is relatively weak. Additionally we give the conditions for social optimality and we show that social interests are not always aligned with the interests of all individual employers.

It is very interesting to take a closer look at the link between the effort-output relationship (γ) and the likelihood that negative stereotypes are eliminated. We expect that discrimination at the appointment stage is less likely to occur in industries characterized by performance related pay and a strong relationship between effort and output. A strong effort-output relationship is likely to be found in relatively simple (manufacturing) tasks where the input of the worker directly results in output. In that case it is very easy to measure the input of the worker and performance pay is meaningful in such jobs. Both workers and employers gain relatively much from the introduction of performance pay. A reduced level of discrimination or total elimination of negative stereotypes is conceivable when performance pay is introduced in such jobs.

When the task becomes more complex and the output depends on more factors than just the effort exerted by the worker, there seems to be less space for performance pay. The employer optimally chooses to provide lower monetary incentives when the effort-output relationship is weak. The result is that worker incentives do not differ very much from the incentives under a flat wage regime, while employers do not gain very much from the introduction of performance pay as well. In that case the introduction of performance related pay is not a suitable way to reduce or eliminate discrimination at the appointment stage. These findings are important from a policy making point of view. Stimulation of performance pay programs with the objective to reduce discriminatory practices does only make sense when the link between effort and output is sufficiently strong. Otherwise it might backfire and the investment gap between two groups might increase.

We can compare our findings to the results found by Elvira and Town (2001). They found that the supervisor's performance evaluation in more complex tasks is influenced by the race of the subordinate he evaluates. We can contribute to this result by stating that discrimination at the appointment or promotion stage is more likely to occur in more complex tasks, even when the supervisor or employer does not dislike minority workers per se. That is, discrimination might occur as a result of beliefs about group productivity as well.

Our study is limited to a theoretical model describing the link between performance pay and negative stereotypes in a model of statistical discrimination. This gives rise to some empirical question related to this topic. For further research, it would be interesting to focus on the link between γ and the likelihood that discrimination occurs in an empirical study. Is discrimination at the appointment stage more likely to occur in industries with a weak effort-output relationship? In a same vein it would be interesting to find out whether a self-selection effect exists. Do minority workers self-select into firms or industries characterized by performance pay and a strong effort-output relationship?

Appendix A

In section 3 we derived the optimal levels of effort and the optimal contracts offered by the employer. Using this information we are able to find the profits of assigning a qualified worker to the high-skilled task. The loss associated with assigning an unqualified worker to the high-skilled job is the same under the two regimes, since we assume that $\bar{\alpha} = \bar{w}$. Under a flat wage regime, fixing p at p^*, profits equal:

$$\Pi_q^h = p^*\gamma - \bar{w} \qquad \text{(A1)}$$

Under performance pay the profits of assigning a qualified worker to the high-skilled job are given by:

$$\Pi_q^h = p^*\gamma(1+\beta) - \left(\bar{\alpha} + \beta(1+\beta)\right) \qquad \text{(A2)}$$

Given that $p^*\gamma > 1$ and that $\bar{\alpha} = \bar{w}$, we see that both revenue (price times output) and wage costs increase under performance pay. The question is which of the two effects dominates. Therefore we compare the increase in revenue (Δr) to the increase in wage costs (Δw). By substituting the expression for β^* we find that:

$$\Delta r = p^*\gamma\left(\frac{p^*\gamma-1}{2}\right) \qquad \text{(A3)}$$

$$\Delta w = \left(\frac{p^*\gamma-1}{2}\right)\left[1+\left(\frac{p^*\gamma-1}{2}\right)\right] \qquad \text{(A4)}$$

We solve for $\Delta r > \Delta w$ and we find that this inequality holds if and only if $p^*\gamma > 1$. This means that employer profits under performance pay are strictly higher than under a flat wage, given that γ is high enough (see equation (22) in the main text). Only when γ is high enough, employers will consider a switch to performance pay. That is, for low values of γ the introduction of performance pay will result in strictly lower profits.

Appendix B

Using the optimal levels of effort and the optimal contracts derived in section 3 we are able to find the equilibrium utility levels for qualified workers in the high-skilled job.

Flat wage regime:

$$U_q^h = \overline{w} - c_i \qquad \text{(B1)}$$

Performance pay regime:

$$U_q^h = \bar{\alpha} + \beta^*(1+\beta^*) - \frac{1}{2}\beta^{*2} - c_i \qquad \text{(B2)}$$

Given that $p\gamma > 1$, we know that β^* is positive. We also know that $\bar{\alpha} = \overline{w}$. This implies that the utility of a qualified worker working in the high-skilled job is strictly higher than the utility of the same worker under a flat wage regime. Next to that, the utility of an unqualified worker assigned to the high-skilled job is the same under both regimes.

Appendix C

The parameter γ affects optimal standards through the ratio $r^p \equiv x_q/x_u$:

$$r^p = \frac{p^*\gamma(1+\beta^*)-[\bar{\alpha}+\beta^*(1+\beta^*)]}{\bar{\alpha}} \qquad \text{(C1)}$$

Substituting for β^* and taking the derivative w.r.t. γ yields:

$$\frac{\delta r^p}{\delta\gamma} = \frac{1}{2}p^2\gamma + \frac{1}{2}p > 0 \qquad \text{(C2)}$$

From equation (C2) it follows that the ratio $r^p \equiv x_q/x_u$ increases in γ. This is completely explained by the fact that profits are increasing in γ. In addition we check whether the increase in profits increases in γ by taking the second derivative w.r.t. γ:

$$\frac{\delta^2 r^p}{\delta\gamma^2} = \frac{1}{2}p^2 > 0 \qquad \text{(C3)}$$

From equation (C3) we can conclude that the increase in profits is increasing in γ.

The parameter γ affects worker investment as well. Substituting for β^* yields the following expression for the benefit from investing, excluding investment costs:

$$E[B(s)] \equiv \bar{\alpha}\left[F_u(s) - F_q(s)\right] + \left[1 - F_q(s)\right]\left(\frac{p^*\gamma}{2} + \frac{1}{2}\left(\frac{p^*\gamma}{2}\right)^2\right) \qquad \text{(C4)}$$

Taking the first derivative of $E[B(s)]$ w.r.t. γ yields:

$$\frac{\delta E[B(s)]}{\delta\gamma} = \left[1 - F_q(s)\right]\left(\frac{1}{2}p^* + \frac{1}{4}p^{*2}\gamma\right) > 0 \qquad \text{(C5)}$$

Equation (C5) shows that the benefit from investing excluding investment costs increases in the parameter γ.

Appendix D

In this appendix we derive the total profits earned by the employer in equilibrium. We assume that the workforce of the firm is standardized to 1 and consists of two groups of workers. A fraction λ of the workers belongs to the W's and a fraction $1-\lambda$ belongs to the B's. First of all we calculate the total number of qualified workers assigned to the high-skilled job (N_q^h):

$$N_q^h = \lambda\left[\left(1-F_q\left(s^*(\pi^w)\right)\right)\pi^w\right] + (1-\lambda)\left[\left(1-F_q\left(s^*(\pi^b)\right)\right)\pi^b\right] \quad \text{(D1)}$$

The total number of unqualified workers assigned to the high-skilled job N_q^h equals:

$$N_u^h = \lambda\left[\left(1-F_u\left(s^*(\pi^w)\right)\right)(1-\pi^w)\right] + (1-\lambda)\left[\left(1-F_u\left(s^*(\pi^b)\right)\right)(1-\pi^b)\right] \quad \text{(D2)}$$

Recall from the main text that the employer earns $x_q > 0$ when assigning a qualified worker to the high-skilled task, while he earns $-x_u < 0$ when assigning an unqualified worker to the high-skilled task. Using these payoffs it is easy to calculate employer profits in equilibrium from equations (D1) and (D2):

$$\begin{aligned}\Pi^* = {} & \lambda\left[\left(1-F_q\left(s^*(\pi^w)\right)\right)\pi^w x_q - \left(1-F_u\left(s^*(\pi^w)\right)\right)(1-\pi^w)x_u\right] \\ & +(1-\lambda)\left[\left(1-F_q\left(s^*(\pi^b)\right)\right)\pi^b x_q - \left(1-F_u\left(s^*(\pi^b)\right)\right)(1-\pi^b)x_u\right] \quad \text{(D3)}\end{aligned}$$

References

Akerlof, George A., and Rachel E. Kranton. 2005. "Identity and the Economics of Organizations," *Journal of Economic Perspectives* 19 1:9-32.

Arrow, Kenneth J. 1972. "Models of Job Discrimination," In A.H. Pascal, ed. *Racial Discrimination in Economic Life*. Lexington, MA: D.C. Heath:83-102.

Becker, Gary. 1957. "The Economics of Discrimination," *Chicago: University of Chicago Press*.

Belman, Dale, and John S. Heywood. 1988. "Incentive Schemes and Racial Wage Discrimination," *Review of Black Political Economy* 17 (summer):47-56.

Coate, Stephen, and Glenn C. Loury. 1993. "Will Affirmative-Action Policies Eliminate Negative Stereotypes?," *The American Economic Review* 83 5:1220-1240.

Elvira, Marta, and Robert Town. 2001. "The Effects of Race and Worker Productivity on Performance Evaluations," *Industrial Relations* 40 4:571-590.

Fang, Tony, and John S. Heywood. 2006. "Output Pay and Ethnic Wage Differentials: Canadian Evidence," *Industrial Relations* 45 2:173-194.

Heywood, John S., and Patrick L. O'Halloran. 2006. "Market Structure, Payment Methods and Racial Earnings Differences," In J.S. Heywood and J. Peoples, *Product Market Structure and Labor Market Discrimination*, Albany, NY: SUNY Press:187-208.

---. 2005. "Racial Earnings Differentials and Performance Pay," *Journal of Human Resources*, 40 2:435-452.

Lazear, Edward P. 2000. "Performance Pay and Productivity," *American Economic Review* 90 5:1346-1361.

Lazear, Edward P., and Sherwin Rosen. 1981. "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy* 89 5:841-864.

Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism," *The American Economic Review* 62 4: 659-661.

Welch, Finis. 1967. "Labor-Market Discrimination: An Interpretation of Income Differences in the Rural South," *The Journal of Political Economy* 75 3:225-240.