

Subtle Discrimination*

Elena S. Pikulina

Daniel Ferreira

July 2023

Abstract

We propose a theory of *subtle discrimination*, defined as biased acts that cannot be objectively ascertained as discriminatory. We present a model in which candidates compete for a promotion. When choosing among similarly qualified candidates, the principal subtly discriminates by favoring candidates from a particular group. Subtle discrimination matters because it affects decisions to invest in human capital. The model predicts that discriminated agents perform better in low-stakes careers while favored agents perform better in high-stakes careers. In equilibrium, firms are polarized: high-productivity firms strive to be “progressive” and have diverse top management teams, while low-productivity firms prefer to be “conservative” and have little diversity at the top.

Keywords: Discrimination, human capital, firm-specific skills, promotion

JEL Classification: M51, J71, J31

*Pikulina: elena.pikulina@sauder.ubc.ca, Finance Division, Sauder School of Business, University of British Columbia. Ferreira: d.ferreira@lse.ac.uk, Department of Finance, London School of Economics and Political Science, CEPR and ECGI. We thank Murray Carlson, Hans Degryse, Jack Favilukis, Guido Friebel, Paolo Fulghieri, Lorenzo Garlappi, Will Gornall, Sebastian Gryglewicz, Isaac Hacamo, Alex Imas, Alesio Piccolo, Maddalena Ronchi, and Edward Van Wesep for valuable feedback and, on occasion, lively discussions. We also thank seminar and conference participants at Schulich School of Business (York University), Sauder Business School (University of British Columbia), London School of Economics and Political Science, New Economic School, the University of Naples, the University of Hong Kong, Erasmus University Rotterdam, Bayes Business School, Goethe University Frankfurt, 2022 Annual Pacific Northwest Finance Conference, 2022 North-American Economic Science Association Conference, 2023 American Finance Association Meetings, 2023 FMA Napa/Sonoma Conference, ESCP ESG Worskshop, 2023 Corporate NBER Meetings, 2023 Adam Smith Conference, and CSEF Conference on Finance, Labor and Inequality for valuable comments.

1 Introduction

Often today the bias is just subtler, the attitudes more hidden, the rationalization more nuanced. Exclusions show up in forms that are harder to prove but continue to keep workplaces homogeneous. It's often so subtle that those in power find it hard to see, harder to acknowledge, and impossible to fix, in spite of all the stories, the data, and the research making it clear that the problem is very real. (Pao 2017, p. 9).

Economists traditionally classify discriminatory acts based on their source. Some view discriminatory acts as a consequence of rational statistical discrimination (Phelps (1972); Arrow (1973)). A second view is that discrimination is caused by biases in preferences or tastes (Becker (1957)), beliefs (Bordalo et al. (2016); Bohren et al. (2019)), or incentives (Dobbie et al. (2021)). Such a classification is undoubtedly useful, but not the only one possible. Social and organizational psychologists classify discriminatory acts based on their transparency. These scholars define *subtle discrimination* as acts that are ambiguous in intent to harm, ex-post rationalizable, difficult to verify, and often (but not always) unintentional.¹ In the workplace, examples include asking female employees to perform menial tasks, not praising the performance of minority employees, and overpromoting men to managerial positions when choosing among equally-qualified candidates.

It is hard to substantiate claims of subtle discrimination. For a discriminatory act to be prosecuted in the United States, the discriminated party must provide direct evidence of intent to harm or deny rights, or prove a clear pattern of adverse events unexplainable on grounds other than discrimination. Partly due to the threat of legal action, acts of overt discrimination have become relatively rare. By contrast, nondiscriminatory narratives and plausible deniability may cloak subtle discrimination and make it a common but

¹See, for example, Dovidio and Gaertner (1986, 2000), Essed (1991), Deitch et al. (2003), Dipboye and Halverson (2004), Noh et al. (2007), Hebl et al. (2008), Van Laer and Janssens (2011), Jones et al. (2017), Dhanani et al. (2018), and Hebl et al. (2020). While studies often use alternative terms, such as “modern discrimination,” “aversive discrimination,” “everyday discrimination,” “ambivalent discrimination” and “covert discrimination,” they all contrast subtle discrimination with “old-time” overt discrimination and emphasize its ambiguous, hard-to-detect and yet pernicious nature.

invisible phenomenon.

Despite its prevalence, the impact of subtle discrimination on workers and firms has received little attention in the economics literature. Our paper is a first attempt at formalizing the notion of subtle discrimination. We define subtle discrimination as biased acts that cannot be objectively ascertained as discriminatory. The defining property of subtle discrimination is *plausible deniability*. When a biased decision-maker acts in a discriminatory way, the decision-maker may offer a narrative to defend his acts. When such a narrative is plausible, we say that the decision-maker is *subtly biased*. For example, a manager who asks a female subordinate to perform a menial task could argue that he sometimes asks men to perform such duties and that the frequency of such requests is low and unrelated to gender. Likewise, if performance is subjectively assessed, a manager accused of not praising someone's performance could claim that they did not perform adequately. Finally, to conceal a small bias towards a specific group, a manager may say that he uses subjective criteria (e.g., "potential") to select between two candidates with the same objective qualifications and performance records.

We apply our notion of subtle discrimination to a model of promotions. In the model, two ex-ante identical agents with labels "blue" and "red" compete for promotion by investing in human capital. Labels are payoff-irrelevant; profit depends only on the promoted agent's acquired skills. The decision-maker (the *principal*) has a small bias in favor of the blue agent. Because the bias is small, when the red agent is objectively more qualified than the blue candidate, the principal prefers to promote the former. That is, the principal does not *overtly discriminate*. However, when both candidates are similarly qualified, the principal is more likely to promote the blue candidate. Here, discrimination is subtle because one cannot observe the "tie-breaking" rule that the principal uses. Thus, in the context of the model, subtle discrimination is an inability or unwillingness to break ties fairly.

Our analysis relies on the fact that, in many circumstances, choosing between two

candidates with similar objective qualifications is difficult. In such a case, the principal is likely to use his subjective assessment to separate the candidates. Managerial discretion allows biases to affect decisions. In an employment setting, Hoffman et al. (2018) show that biases rather than information explain most of the cases in which managers use their discretion for selecting candidates. Using internal data from a large bank, Bircan et al. (2022) find that supervisors' discretion in job assignments explains most of the gender promotion gap, implying the existence of "*subtle mechanisms that disadvantage women*" (p.3).

Our paper makes three primary contributions. First, we propose a classification of discriminatory acts into two categories: subtle and overt. Second, in a model of promotions, we show that subtle discrimination and overt discrimination have different empirical predictions. Third, our model of subtle discrimination in promotions generates a rich set of empirical predictions relating firm characteristics to the performance of different groups of workers, the diversity of top management teams, and firms' choices of anti-discrimination policies.

We distinguish subtle discrimination from overt discrimination based on the ease of objectively ascertaining particular acts as discriminatory. For example, denying someone a job *because* of their race or making derogatory comments about a person's identity are overtly discriminatory acts because they are clearly based on a person's group identity. In contrast, when two candidates are similar in their qualifications, passing over one candidate for a promotion is not clear evidence of discrimination. In our model, when blue and red agents are similarly skilled, the principal can rationalize (to others or to himself) the act of promoting the blue agent by using his subjective assessment to estimate the agents' "potential." In contrast, promoting an unskilled agent from a favored group over a skilled agent from an unfavored group is clear evidence of discrimination; we classify such acts as overt discrimination. While our particular formalization necessarily leaves out important details, it is simple, intuitive and tractable.

Although individual acts of subtle discrimination cannot be observed, in the model, the agents form beliefs about the principal's *subtle bias*. Such beliefs affect agents' decisions to invest in human capital. Because promotions are competitive, subtle biases also affect how agents react to each other's decisions. In equilibrium, agents differ in their investment decisions, which creates an *achievement gap*, i.e., a difference in accumulated human capital and obtained qualifications. Two opposing forces contribute to the achievement gap. On the one hand, unfavored agents are discouraged from investing in human capital because they anticipate a low probability of being promoted.² We call this force the *discouragement effect*. On the other hand, an unfavored agent may choose to overinvest in skills in an attempt to separate herself from the favored agent. We call this force the *overcompensation effect*. Unlike the discouragement effect, the overcompensation effect can dominate only when discrimination is subtle rather than overt.

We show that the sign and the magnitude of the achievement gap depend on the stakes faced by the agents. When the net benefit from promotion is large – a *high-stakes career path* – the discouragement effect dominates, implying that favored (blue) agents invest *more* than unfavored (red) agents. In this case, the achievement gap is positive: favored agents have more visible achievements (e.g., better qualifications and performance records) than unfavored agents. In contrast, when the net benefit from promotion is small – a *low-stakes career path* – the overcompensation effect dominates and, thus, favored agents invest *less* than unfavored agents, leading to a negative achievement gap. We show that these results hinge crucially on discrimination being subtle instead of overt.

These results are helpful when interpreting the evidence on the professional advancement of women. Evidence that women have lower promotion rates in high-skilled occupations can be found in Hospido et al. (2019) for central bankers, Bosquet et al. (2019) for academic economists, Azmat et al. (2020) for lawyers, and Bircan et al. (2022) for bankers. Promotions in such careers are typically associated with large increases in pay and non-

²See Coate and Loury (1993) for an early model of underinvestment in human capital in the context of statistical discrimination.

pecuniary benefits, such as prestige and status. Azmat et al. (2020) show that female associates in law firms invest less in the qualifications required for promotion (e.g., hours billed) than male associates. Hospido et al. (2019) and Bosquet et al. (2019) find that women are less likely to seek promotion in the first place. By contrast, Benson et al. (2021) find that women in management-track careers in retail have better (pre-promotion) performance than men.³ These facts are consistent with our prediction that discriminated groups are discouraged from investing in promotable tasks in high-stakes careers while being over-incentivized to undertake such investments in low-stakes careers.

Our model also predicts that, in high-stakes careers, differences in observable achievements (such as human capital, performance, experience, and effort) explain most of the *promotion gap* (i.e., the difference in promotion rates between groups). Because the promotion gap increases with the expected benefits of promotion, the model can also explain the evidence of increasing promotion gaps at the top of hierarchies, a fact that is known as the “leaky pipe” phenomenon (Lundberg and Stearns (2019); Sherman and Tookes (2022)). Such evidence is at odds with Lazear and Rosen’s (1990) prediction that promotion gaps should decrease as the skill required for a task increases.

To shape their subtle biases, we allow firms to adopt different anti-discrimination (i.e., diversity, equity, and inclusion) policies. Firms can choose to become more progressive (i.e., less biased) or conservative (i.e., more biased). In equilibrium, firms become polarized. On one side, we have high-productivity firms offering high-stakes careers to their employees. Such firms choose to become progressive and, thus, have greater diversity in their top management teams. On the other side, we have low-productivity firms that offer low-stakes careers. Such firms choose to be conservative and, thus, have little diversity at the top. The model predicts that even small differences in firm productivity can account for large differences in corporate culture and top-level diversity. Furthermore, market forces cannot eliminate such differences. Thus, our model provides novel

³Despite women’s better performance, supervisors still consider men to have higher “potential” on average, which leads to higher promotion rates for men.

empirical predictions relating firm quality to observed diversity metrics. Consistent with the model’s predictions, Edmans et al. (2023) find that employees’ perception of diversity, equity, and inclusion is stronger in growing, high-valuation, and financially strong firms. Firm polarization in diversity preferences is also a potential explanation for the evidence that large and well-performing firms have more women on their boards (Adams and Ferreira (2009)).

Our model offers a novel perspective on the costs and benefits of companies focusing on social issues, especially regarding the diversity of their workforce. While some argue that progressive values typically do not conflict with the pursuit of profit (e.g., Edmans (2020)), others claim that some businesses excessively focus on promoting progressive causes to the detriment of profits (see Edgecliffe-Johnson (2022)). Our model illustrates one mechanism through which progressive firms can increase profits: Employees who believe that a company does not discriminate in promotions are encouraged to invest in promotable tasks. Moreover, the model shows that such benefits accrue primarily to firms with high-stakes career paths. By contrast, firms in which investment in human capital has low returns have less to gain from eliminating discrimination in promotions. For such firms, investing in a reputation for progressiveness does not pay off.

In the next section, we begin by defining subtle discrimination in the context of a choice between two candidates for a position. We present our main analysis in Section 3. In Section 4, we review some of the related theoretical literature. We discuss the main empirical implications in Section 5, and conclude. The Internet Appendix presents several variations of our main model.

2 Subtle Discrimination: Definition and Interpretation

In this section, we present the decision-making setup we use to define subtle discrimination. A decision-maker needs to choose one of two candidates $i \in \{b, r\}$, called Blue

and Red. Blue and Red have objective qualifications and skills, summarized by s_b and s_r , which are observed by everyone. Without loss of generality, we set $\Delta s \equiv s_r - s_b \geq 0$. The decision-maker privately observes a (subjective) signal x_i for each agent. The signals are independent and identically distributed random variables with support $[x, \bar{x}]$. Let $F(\cdot)$ denote the cumulative distribution function of $\Delta x \equiv x_r - x_b$. The decision-maker has a mandate to maximize an observable payoff (e.g., profit), π . If the decision-maker chooses agent i , the payoff is $\pi_i = s_i + \omega x_i + u$, where $\omega > 0$ and u is a zero-mean random variable. We assume that everyone knows ω and holds the same beliefs about $F(\cdot)$ and u .

Note that, in the absence of biases, the decision-maker would choose Blue if $\Delta s + \omega \Delta x < 0$. Thus, the probability of an unbiased decision-maker choosing Blue is $F\left(-\frac{\Delta s}{\omega}\right)$. Now, let's consider the case of a biased decision-maker. To keep the analysis general, we model the decision-maker's behavior directly, without specifying payoffs and beliefs. Let $P(\Delta s, \omega)$ denote the probability that the decision-maker chooses Blue given Δs and ω . Our only assumption for the decision-maker's behavior is that $P(\Delta s, \omega)$ is decreasing in Δs . We define the decision-maker's bias towards Blue as

$$b(\Delta s, \omega) = P(\Delta s, \omega) - F\left(-\frac{\Delta s}{\omega}\right). \quad (1)$$

There could be many sources for the bias in (1). For example, the decision-maker may have a preference for blue agents (as in Becker (1957)). Alternatively, the decision-maker may incorrectly believe that blue agents are more productive (Bordalo et al. (2016); Bohren et al. (2019)). Even with correct beliefs, the decision-maker may be better at reading subjective signals from blue agents (e.g., Cornell and Welch (1996); Fershtman and Pavan (2021)). Biases may also come from extraneous sources, such as poorly designed incentive structures (Dobbie et al. (2021)). In what follows, we do not take a stand on whether biases are caused by beliefs or preferences, or whether they are intrinsic or extrinsic. Our model can accommodate most of these interpretations.

Next, we define our notion of subtle bias:

Definition. If $F\left(-\frac{\Delta s}{\omega}\right) > 0$, we say that bias $b(\Delta s, \omega)$ is subtle.

The intuition is as follows. By assumption, Red’s objective qualifications are (weakly) better than Blue’s qualifications ($\Delta s \geq 0$). Nevertheless, the decision-maker may still choose Blue, either because the decision-maker is biased towards Blue or because the decision-maker privately observes $-\omega\Delta x \geq \Delta s$. That is, whenever $F\left(-\frac{\Delta s}{\omega}\right) > 0$, a biased decision-maker can “plausibly deny” being biased towards Blue. *Plausible deniability* is the defining property of subtle discrimination. In our formalization, plausible deniability means that, if everyone observes Δs and ω and has the same belief about $F(\cdot)$, choosing Blue does not constitute conclusive evidence that the decision-maker is biased. *Subtle discrimination* is thus the act of offering favored treatment to an agent or group of agents when the decision-maker can resort to a plausible-deniability defense of such an act.

Our notion of subtle discrimination is also compatible with the decision-maker being unconsciously biased. This interpretation is valid under the assumption that the decision-maker does not directly benefit from choosing a particular type. One interpretation is that the decision-maker always rationalizes his choice (as in Cherepanov et al. (2013)). In practice, the decision-maker might find it difficult to correct the bias (at least in a finite series of decisions) if he believes that his choices are unbiased. Such unconscious biases are most likely to pertain to System 1 thinking, i.e., fast, automatic, and effortless associations (Kahneman (2011)).

Note that subtle discrimination is more likely to occur when objective differences (here summarized by Δs) are small. We can think of $F\left(-\frac{\Delta s}{\omega}\right)$ as a measure of “subtlety” in discrimination. Note that bias subtlety is maximized at $\Delta s = 0$. When agents are observationally equivalent, any particular choice is rationalizable even when the importance of subjective information is minimal (i.e., when $\omega \rightarrow 0$).

In contrast with subtle biases, we now define *overt bias*:

Definition. If $F\left(-\frac{\Delta s}{\omega}\right) = 0$, we say that bias $b(\Delta s, \omega)$ is overt.

Case $F\left(-\frac{\Delta s}{\omega}\right) = 0$ occurs when $\Delta s > \omega(\bar{x} - \underline{x})$. In this case, choosing Blue is incontrovertible evidence of biased decision-making.

According to our taxonomy, a bias is either subtle or overt. While this binary classification is useful, we can also think of biases being more or less subtle (or overt). In particular, without the bounded support assumption, all biases would be subtle. Still, a subtle bias with very small $F\left(-\frac{\Delta s}{\omega}\right)$ is, for all practical purposes, an overt bias, in the sense that one would find it difficult to defend choosing Blue.

Note that, by definition, an individual act of subtle discrimination is not immediately detectable. While outcome realizations (i.e., π) might be informative about underlying biases, they may not be sufficient to reveal subtle biases, because the noise in performance (u) provides cover for subtle discrimination. Still, agents and observers may hold (correct or incorrect) beliefs about the existence of subtle biases.⁴ Thus, subtle biases may matter through their influence on agents' decisions.

It might be possible to estimate subtle biases statistically after observing a very long sequence of identical decisions made by the same decision-maker. In reality, such a way of detecting subtle discrimination faces practical difficulties. First, for very small differences in observables Δs , one would need a very large sample to detect subtle discrimination with a reasonable degree of confidence. Second, again for small differences Δs , even mild disagreements about ω and $F(\cdot)$ can provide sufficient cover for subtle biases. Thus, our notion of subtle discrimination is particularly relevant in situations where (i) differences in objective qualifications between groups are small, (ii) decisions (such as promotions) made by the same decision-maker are infrequent, and (iii) observers disagree about the importance of subjective information for forecasting performance.

In the next section, we present an application of our concept of subtle discrimination to a model of promotions. For simplicity, we consider the limiting case in which the principal's subjective information is negligible for the decision: $\omega \rightarrow 0$. In that case,

⁴Bircan et al. (2022) show evidence of correct belief formation about women's disadvantage in the assignment of roles.

subtle discrimination can occur only when the difference in skills $s_r - s_b = \Delta s$ is zero.

3 A Model of Subtle Discrimination in Promotions

In this section, we present a model of subtle discrimination in promotions. After presenting the setup in Subsection 3.1, in Subsection 3.2, we describe the first-best solution to serve as a benchmark. We then solve the model for an exogenously given compensation contract in Subsection 3.3. In Subsection 3.4, we let firms choose compensation contracts optimally. In Subsection 3.5, we endogenize the subtle bias. In the Internet Appendix, we discuss the robustness of the model with respect to its assumptions.

3.1 Definitions and Model Setup

At Date 0, a firm hires two ex-ante identical agents – b (*Blue*) and r (*Red*) – for an entry-level position (*job 1*). Both vacancies need to be filled. Red and Blue are payoff-irrelevant labels. To save on notation, we normalize the revenue generated by the agents on their entry-level jobs to zero.

We assume that the firm does not (or cannot) discriminate at the hiring stage; thus, the 50/50 split between b and r reflects the composition of the candidate pool in the sector. Explicitly modeling the market for workers does not change the main model implications, as long as candidate pools do not become fully segregated. That is, we implicitly assume that some frictions prevent firms from selecting applicants of a single type.

At Date 1, the agents simultaneously undertake a non-observable investment (or effort), $e_i \in [0, 1]$, $i \in \{b, r\}$, to acquire a “skill.” The skill is an observable but not verifiable binary variable: $s_i \in \{0, 1\}$.⁵ Agent i ’s probability of acquiring the skill is e_i . Both agents are risk-neutral and have the same skill-acquisition cost function, $c(e_i)$, which we assume

⁵Strictly speaking, the model requires only that the principal observes s_i , thus the “non-verifiability” of s_i is not a crucial assumption. However, assuming that all parties observe s_i makes the interpretation more natural.

is strictly increasing and convex. That is, agent i 's utility is $u_i = w_i - c(e_i)$, where w_i is the agent's monetary compensation. Without loss of generality, we set $c(0) = 0$.

We interpret skill broadly as any kind of observable evidence that predicts an agent's future performance.⁶ We assume that the skill is firm-specific in the sense that it is less valuable to agents who leave the firm; see the Internet Appendix for a further discussion of this assumption and an extension in which agents invest in general skills.

At Date 2, the principal can choose one of the agents to fill a top position (*job 2*) in the firm.⁷ Agents who are not promoted remain at the entry-level job. While working in job 1, the agents reveal their managerial abilities to the principal. Formally, with probability μ , agent i generates a signal $m_i = 1$, indicating that the agent is qualified for managerial positions. Promoting an agent who does not have managerial capabilities ($m_i = 0$) results in negative profits. Thus, only those with managerial capabilities will be promoted. Among those with $m_i = 1$, agents can be skilled ($s_i = 1$) or unskilled ($s_i = 0$). Promoting an unskilled agent increases the principal's expected payoff by $l > 0$ while promoting a skilled agent increases the payoff by $l + H$, where $H > 0$ denotes the *productivity gain* upon promotion of a skilled agent. That is, a skilled agent is always more productive than an unskilled one when assigned to job 2.⁸ We interpret H as a firm characteristic. Larger H means that human capital is more important at higher hierarchical levels.

Although the principal cannot offer wages contingent on skill acquisition, the principal can commit to a set of non-negative wages (w_1, w_2) for the holders of jobs 1 and 2, respectively. We call $W \equiv w_2 - w_1$ the *promotion premium*. We describe the compensation contract by a vector $w = (w_1, W)$ representing a basic reward and a promotion premium.

We are interested in the case in which contractual discrimination in promotion decisions is not possible. That is, the principal must offer the same contract w to both agents.

⁶For example, in the legal profession, hours billed to clients and new client revenue raised are the main tools used to assess the performance of associates (Cotterman (2004), Heinz et al. (2005)).

⁷Although we interpret this decision as a promotion, it can also be interpreted as a direct hiring decision for a specific post. See Coate and Loury (1993) for a model that also allows for both interpretations.

⁸More generally, investment in skill may also affect the probability of acquiring managerial capabilities (μ). We chose the current formulation for simplicity.

The principal uses the promotion contest to provide incentives for skill acquisition. Because $H > 0$, when choosing between two candidates with managerial capabilities, an unbiased principal always promotes a skilled agent over an unskilled one. As in Prendergast (1993), the principal can effectively commit to rewarding skill acquisition through promotions. In addition, if $l > W$, it is always in the principal's interest to promote one of the agents, even when both agents are unskilled. As l is a free parameter in the model, we assume that it is sufficiently high so that $l > W$. For the remainder of the paper, to simplify the exposition, we assume that $\mu = 1$, and thus $m_b = m_r = 1$ always, that is, both agents have managerial capabilities. This assumption implies that job 2 is always filled with one of the agents.

As in Section 2, we model the principal's behavior by the probability of choosing Blue given ω and Δs . For simplicity, we consider the limiting case in which the principal's subjective information is negligible for the decision: $\omega \rightarrow 0$. There are three cases for the difference in skills: $s_r - s_b = \Delta s \in \{-1, 0, 1\}$. We assume that there is no overt discrimination: if $\Delta s \neq 0$, the principal chooses the skilled agent (Blue if $\Delta s = -1$ and Red if $\Delta s = 1$). However, if $\Delta s = 0$, the principal chooses Blue with probability $0.5 + \beta$, where $\beta \in (0, 0.5]$ is the principal's subtle bias. Using the notation in Section 2, $\beta = \lim_{\omega \rightarrow 0} \lim_{\Delta s \rightarrow 0} b(\Delta s, \omega)$. Thus, given the principal's behavior, the firm's (expected) profit is $\Pi = l + H(e_b + e_r - e_b e_r) - 2w_1 - W$.

In the context of this model, the principal's behavior can be described as a simple lexicographic heuristic: if $s_b \neq s_r$, choose the skilled agent; if $s_b = s_r$, choose Blue with probability $0.5 + \beta$. While we assume that ties are exact, this is not a necessary condition for subtle discrimination (see Section 2). It is also not necessary for the main results; all we need is that observable skills are sufficiently similar so that subtle discrimination is possible. Assuming exact ties is algebraically convenient for solving the firm's maximization problem as described in Subsections 3.4 and 3.5, but otherwise it is not needed as we show in the Internet Appendix.

3.2 Benchmark: First-best Investment Levels

As a benchmark, we consider the problem of an unbiased social planner who maximizes total surplus. Define (expected) social surplus as $S = \Pi + E[u_b] + E[u_r]$. The planner's problem is

$$\max_{(e_b, e_r) \in [0,1]^2} l + H(e_b + e_r - e_b e_r) - c(e_b) - c(e_r). \quad (2)$$

The planner faces a trade-off between effort duplication and effort sharing. On the one hand, asking both agents to invest in skill acquisition implies that, with positive probability, some acquired skills goes to waste. This waste is the cost of effort duplication. On the other hand, if only one agent invests, her marginal cost of effort is higher than that of the idle agent. Similar to risk sharing under concave utilities, effort sharing (i.e., marginal cost equalization across agents) is efficient under convex costs. The first-best choice thus depends on which of these two effects dominates.

The following proposition formalizes this intuition (all proofs are provided in Appendix A):

Proposition 1. *The first-best investment levels can take one of two forms: (i) $e_b^{FB} = e_r^{FB} = \tilde{e} < 1$ or (ii) $e_b^{FB} > 0$ and $e_r^{FB} = 0$ (or, equivalently, $e_b^{FB} = 0$ and $e_r^{FB} > 0$).*

Proposition 1 says that the first-best outcome can be symmetric or asymmetric. If the benefits from effort sharing are greater than the costs of effort duplication, the social planner forces both agents to choose the same investment level (Case (i)). If the costs of duplication outweigh the benefits from effort sharing, the social planner asks only one agent to invest in skill acquisition (Case (ii)).

Consider the case in which $c(e_i) = \frac{ke_i^2}{2}$. Then, if $H \leq k$, the first-best solution is

$$e_b^{FB} = e_r^{FB} = \tilde{e} = \frac{H}{H+k}. \quad (3)$$

That is, treating both agents equally is socially optimal. If $H > k$, the first-best is a corner

solution, $e_b^{FB} = 1$ and $e_r^{FB} = 0$ (or $e_b^{FB} = 0$ and $e_r^{FB} = 1$). That is, it is better to treat agents asymmetrically and ask only one of them to invest in skill acquisition.

To simplify the exposition, for the rest of the paper, we assume that the cost function is $c(e_i) = \frac{ke_i^2}{2}$. We choose to sacrifice generality to obtain analytical proofs for most results, which help explain the economic forces at play. Nevertheless, none of the main results depends on the cost function being quadratic (see the Internet Appendix).

3.3 Equilibrium under Exogenous Compensation Contracts

Here we describe the agents' investment choices under a fixed contract w . We assume that the contract is individually rational; both Blue and Red accept the contract at Date 0. At Date 1, the agents simultaneously choose their investment (i.e., effort) levels. At Date 2, investment outcomes are realized and the principal decides who to promote to the top position. Both agents anticipate that, at Date 2, the principal's decision will be biased in favor of Blue. That is, in case of a tie, the principal promotes Blue with probability $\frac{1}{2} + \beta$. While we interpret such beliefs as being correct, the results in this section do not depend on whether beliefs are correct or incorrect.

3.3.1 Equilibrium Characterization

We define agent i 's expected utility as:

$$U_i(e, w) \equiv w_1 + W \left[e_i(1 - e_{-i}) + \left(\frac{1}{2} + \beta_i \right) e_i e_{-i} + \left(\frac{1}{2} + \beta_i \right) (1 - e_i)(1 - e_{-i}) \right] - \frac{ke_i^2}{2}. \quad (4)$$

In the agent's expected utility function, the first term is the baseline reward, the second term is the promotion premium times the probability of promotion, and the third term is the skill-acquisition cost. The first term inside the square brackets corresponds to the probability of agent i acquiring the skill when agent $-i$ fails to acquire the skill. In this

case, the principal promotes agent i . The second and third terms correspond to the probability of promotion via a tie-breaking decision.

An agent's problem at Date 1 is to maximize his/her expected utility $U_i(e, w)$ by choosing an investment level e_i taking the contract, w , and the effort of the other agent, e_{-i} , as given:

$$\max_{e_i \in [0,1]} U_i(e, w). \quad (5)$$

Assuming an interior solution,⁹ the agents' reaction functions are

$$e_b = \frac{W}{k} \left(\frac{1}{2} - \beta + 2\beta e_r \right), \quad (6)$$

$$e_r = \frac{W}{k} \left(\frac{1}{2} + \beta - 2\beta e_b \right). \quad (7)$$

Define $\sigma \equiv \frac{W}{k}$, i.e., the ratio of the promotion premium to the cost parameter. We call σ the *premium-cost ratio*. The premium-cost ratio is a reaction function shifter (see (6) and (7)). Higher σ implies a higher net marginal benefit of investment for any given pair (e_b, e_r) . Intuitively, a high premium-cost ratio implies that the gain from promotion, W , is large relative to the cost of investment, which is proportional to k . High σ can thus be interpreted as a "high-stakes" career path, i.e., a case in which there is much to gain from investing in skill acquisition. By contrast, if σ is low, agents benefit little from investing. In this case, we say that the agents are on a low-stakes career path. Thus, we also informally refer to σ as the "stake" of a career path.

In the baseline case with no subtle bias ($\beta = 0$), the reaction functions (6)-(7) are flat, implying that $e_b^* = e_r^* = \frac{\sigma}{2}$ is the dominant strategy. That is, if there is no bias, the agents choose their optimal effort levels as if there were no strategic considerations. The intuition is as follows. If agent i increases her effort by ϵ , her promotion probability in the state where the other agent is skilled increases by $\frac{\epsilon}{2}$ because, in the case of a tie, the winner is chosen with equal probabilities. In the state where the other agent is unskilled, agent

⁹In Subsection 3.4, we show that optimal contracts always imply interior solutions.

i increases her promotion probability by $\epsilon - \frac{\epsilon}{2} = \frac{\epsilon}{2}$, because when she is skilled, she is promoted with certainty and when she is unskilled, she and the other unskilled agent are equally likely to be promoted. In sum, agent i 's promotion probability always increases by half of the increase in her effort, regardless of whether the other agent is skilled or unskilled.¹⁰

The introduction of a bias in favor of the blue agent breaks this symmetry. Now, if Blue increases his effort by ϵ , his promotion probability in the state where Red is skilled increases by $\epsilon(\frac{1}{2} + \beta)$, while his promotion probability in the state where Red is unskilled increases by less: $\epsilon(\frac{1}{2} - \beta)$. Thus, the state now matters for the effort decision: Blue's marginal benefit of effort is larger when Red is more likely to be skilled. That is, Blue's reaction function is positively sloped. For similar reasons, Red's reaction function is negatively sloped. Intuitively, with a subtle bias in favor of blue agents, ties are more valuable to Blue than they are to Red. Thus, Blue wants to imitate Red's behavior, which causes Blue's reaction function to slope upwards. By contrast, Red adopts the opposite strategy in an attempt to avoid ties.

The following proposition characterizes the equilibrium investment choices.

Proposition 2. *A unique equilibrium exists. For any $\beta \in [0, 0.5]$, there exists $\bar{\sigma}(\beta) > 1$ (with $\bar{\sigma}(0.5) = \infty$) such that, if $\sigma \leq \bar{\sigma}(\beta)$, the equilibrium is interior and the investment levels are given by:*

$$e_b^* = \frac{\sigma(0.5 - \beta) + 2\beta\sigma^2(0.5 + \beta)}{1 + 4\beta^2\sigma^2}; \quad (8)$$

$$e_r^* = \frac{\sigma(0.5 + \beta) - 2\beta\sigma^2(0.5 - \beta)}{1 + 4\beta^2\sigma^2}. \quad (9)$$

If $\sigma > \bar{\sigma}(\beta)$, $e_b^* = 1$ and $e_r^* = \min \left\{ \frac{\sigma(1-2\beta)}{2}, 1 \right\}$.

¹⁰The flatness of the reaction functions under $\beta = 0$ is a robust property of the model and not a consequence of the binary setup (see the Internet Appendix for an example). The key assumption here is that low-skill agents are sometimes promoted.

3.3.2 Discouragement versus Overcompensation

Figure 1 shows the equilibrium investment levels as a function of the premium-cost ratio, σ , for two levels of the subtle bias. The figure shows that for low values of σ , Red invests more than Blue, while for high values of σ , it is Blue who invests more.

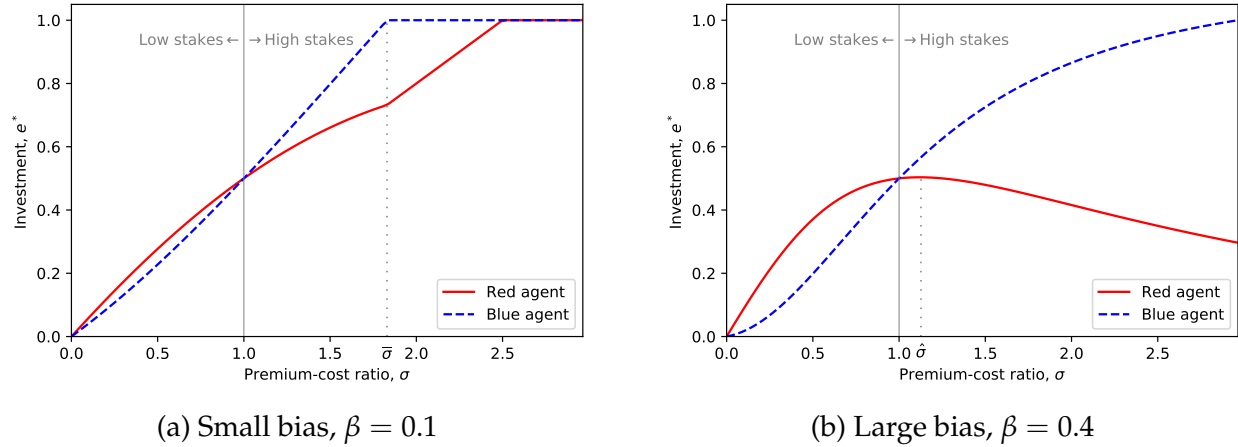


Figure 1: Equilibrium investments, e_b^* and e_r^* , as functions of the premium-cost ratio, σ , for two levels of subtle bias, $\beta_1 = 0.1$ and $\beta_2 = 0.4$.

The following corollary formalizes the comparative statics illustrated in Figure 1. For simplicity of exposition, from now on we assume that the equilibrium is interior.

Corollary 1. *When stakes are low, Red invests more than Blue. When stakes are high, Blue invests more than Red. Formally, $e_r^* \geq e_b^*$ if and only if $\sigma \leq 1$.*

Red’s investment decision is shaped by two opposing forces. On the one hand, Red wants to invest heavily in skill acquisition in an attempt to separate herself from Blue. We call this force the *overcompensation effect*. Overcompensation may occur because the red agent knows she is held to “higher standards:” Unless she is clearly more qualified than Blue, she is viewed less favorably. On the other hand, Red is discouraged from investing because her chances of promotion are slim even if she acquires the necessary skills. We call this force the *discouragement effect*.¹¹ Parameter σ determines what effect dominates in

¹¹See Coate and Loury (1993) and MacLeod (2003) for early models with a similar discouragement effect.

equilibrium. If the stakes are low, Blue exerts low effort. Thus, Red is willing to overcompensate by investing more, both because the probability of separation is high and because the marginal cost of investing is low under a convex cost function. As the stakes increase, Blue chooses higher levels of investment, discouraging Red from investing. At high investment levels, the probability of separation is low while the marginal cost of investing is high.

Remark 1. While the discouragement and overcompensation effects are present in any contest model, for the overcompensation effect to dominate in equilibrium, we need at least one reaction function to slope upwards. If $l < 0$, there might not be a winner to the contest, and both reaction functions slope downwards. In this case, the discouragement effect always dominates the overcompensation effect, and Red invests less than Blue in equilibrium. Focusing on the $l > 0$ case makes the model richer because either effect may dominate in equilibrium.

Remark 2. The interaction between the overcompensation and discouragement effects is robust to situations in which Blue and Red have different beliefs about β . Suppose, for example, that Red believes there is subtle discrimination ($\beta > 0$), but Blue believes that β is zero. Then, we have $e_b^* = \frac{\sigma}{2}$ and $e_r^* = \sigma(\frac{1}{2} + \beta(1 - \sigma))$, implying that Corollary 1 holds.

Remark 3. A potential consequence of the discouragement effect is that a principal who is unaware of his bias (and the strategic interaction it creates between the two agents) might interpret Red's behavior as a lack of interest in high-paying positions. In other words, he might incorrectly "learn" that red and blue agents have different preferences with respect to earned income. Such learning might further reinforce the principal's subtle bias or even result in an explicit bias in favor of blue agents.¹²

¹²In order to do so, the principal could construct a narrative that is not contradicted by his observed data (see, for example, Eliaz and Spiegler (2020)) rather than engage in rational Bayesian learning.

3.3.3 Subtle Discrimination versus Overt Discrimination

To understand how the model's predictions relate to the type of discrimination, here we consider how they would change in the presence of overt discrimination. Define the overt bias as $\delta = \lim_{\omega \rightarrow 0} b(1, \omega)$. Suppose $\delta > 0$. The reaction functions become

$$e_b = \sigma \left[\frac{1}{2} - \beta + (2\beta - \delta)e_r \right], \quad (10)$$

$$e_r = \sigma \left[\frac{1}{2} + \beta - \delta - (2\beta - \delta)e_b \right]. \quad (11)$$

Thus far, we have imposed no structure on β and δ other than $\delta \leq 0.5 + \beta$ (which follows from $P(\Delta s, \omega)$ being decreasing in Δs). Without further structure, we do not know whether the reaction functions in (10) and (11) have positive or negative slopes. To impose further (but minimal) structure, we can think of overt discrimination as a probability over two states of the world: in the first state, which happens with probability δ , the principal has a large bias towards Blue and, thus, chooses Blue with probability one, while in the second state, which happens with probability $1 - \delta$, the principal's bias is small, thus he still chooses Red if $\Delta s = 1$ (Red is skilled and Blue is unskilled). While in State 2 discrimination does not occur when $\Delta s = 1$, it may still occur when $\Delta s = 0$ (both agents have the same skill) because even a small bias might affect decisions in this case. Formally, this implies that δ imposes a lower bound on β : $\frac{1}{2} + \beta \geq \delta + (1 - \delta)\frac{1}{2}$, which implies $2\beta \geq \delta$.

Let $\epsilon \equiv \beta - \frac{\delta}{2} \geq 0$ denote the *excess subtle bias*. If $\epsilon = 0$ (no excess subtle bias), the reaction functions in (10) and (11) are again flat, implying $e_b = e_r = \frac{\sigma(1-\delta)}{2}$ (for $\delta < 1$). That is, if subtle discrimination is fully "explained" by an overt bias, both agents choose the same investment levels in equilibrium. Asymmetric investment levels occur only when subtle discrimination is stronger than overt discrimination (which is the most reasonable scenario). In other words, overt discrimination moderates the effect of subtle discrimination.

We now generalize Corollary 1 for the case in which overt discrimination is present:

Corollary 2. *If $\delta \geq 0$, $e_r^* \geq e_b^*$ if and only if $\sigma \leq \frac{1}{1-\delta}$.*

Corollary 2 shows that overt discrimination makes it less likely that Red invests more than Blue. Intuitively, under overt discrimination, Red benefits less from separating herself from Blue because, even when Red is more skilled than Blue, she is still passed over for promotion with positive probability. That is, the overcompensation effect is weaker when the principal also overtly discriminates.

These results show that subtle discrimination has unique implications. In the context of our model, subtle discrimination creates incentives for separation, while overt discrimination does not. Only in the presence of excess subtle discrimination can the overcompensation effect dominate the discouragement effect. Therefore, our model is particularly relevant in situations where overt discrimination is mild or nonexistent, but subtle discrimination remains.

For the remainder of the paper, we return to the case of “pure” subtle discrimination, i.e., $\delta = 0$. All results are qualitatively unchanged if we interpret β as excess subtle discrimination.

3.3.4 Stakes and Investment in Skills

The next corollary presents further comparative statics results.

Corollary 3. *For $\sigma \leq \bar{\sigma}(\beta)$ (i.e., the equilibrium is interior), we have that*

1. e_b^* is strictly increasing in σ ;
2. There exists $\hat{\sigma}(\beta) \leq \bar{\sigma}(\beta)$ such that e_r^* increases with σ for $\sigma \leq \hat{\sigma}(\beta)$ and decreases with σ for $\sigma > \hat{\sigma}(\beta)$.
3. $\hat{\sigma}(\beta)$ is strictly decreasing in β .

This corollary shows that Blue’s investment in skill acquisition is increasing in the premium-cost ratio (see Part 1). Interestingly, Red’s investment does not always increase

with σ . If the stakes are sufficiently high ($\sigma > \hat{\sigma}$), the discouragement effect dominates and Red's investment declines with the premium-cost ratio (see Part 2; for this to happen, the subtle bias needs to be sufficiently strong). When the bias is stronger, the discouragement effect is also stronger, implying a lower premium-cost ratio at which Red's investment declines with σ (see Part 3).

3.3.5 Comparison with the First Best

It is instructive to compare the equilibrium effort levels to their first-best counterparts. For $\beta > 0$, there is typically no contract that implements the first-best investment levels. If $H > k$, the first-best outcome is $e_b^{FB} = 1$ and $e_r^{FB} = 0$. This outcome is unachievable under subtle discrimination: From Proposition 2, to have $e_b^* = 1$ we need $\sigma \geq \bar{\sigma}$, in which case we have $e_r^* = \min\left\{\frac{\sigma(1-2\beta)}{2}, 1\right\} > 0$ (because $\beta < 0.5$ if $\bar{\sigma}$ is finite). If $H \leq k$, the first-best requires both agents to invest $\tilde{e} = \frac{H}{H+k}$. But agents' investments are the same if and only if $\sigma = 1$, in which case we have (from (8)) $e_r^* = e_b^* = 0.5 \geq \tilde{e}$. Thus, except for the case in which $H = k$, there is no σ that implements the first-best investment levels in the presence of subtle bias ($\beta > 0$).

Things are different if there is no subtle discrimination ($\beta = 0$). If $H \leq k$, the first-best can be achieved by choosing $\sigma^{FB} = \frac{2H}{H+k}$ (i.e., $W^{FB} = \frac{2kH}{H+k}$). If $H > k$, the first-best cannot be achieved.

To summarize: (i) if the principal is subtly biased, there is no contract that implements the first-best outcome, except for the (measure-zero) case in which $H = k$; (ii) if the principal is unbiased, the first-best outcome can be implemented by a suitably-designed promotion contest if and only if $H \leq k$. The comparison with the first-best shows that subtle discrimination is a friction. Without a subtle bias, the first-best can sometimes be achieved. If there are additional contractual frictions, subtle discrimination can nevertheless be welfare-enhancing in some cases, as we show in the Internet Appendix.

3.3.6 The Promotion Gap

We now define the promotion gap between blue and red agents:

Definition 1. Let p_i denote agent i 's promotion probability, $i \in \{b, r\}$. The promotion gap is

$$\Delta p \equiv p_b - p_r = e_b - e_r + [e_b e_r + (1 - e_b)(1 - e_r)] 2\beta. \quad (12)$$

That is, the promotion gap is the difference between the promotion probabilities of Blue and Red.

Note that the promotion gap in (12) has two terms. The first term, $e_b - e_r$, is the difference in the probabilities of skill acquisition. Given our broad interpretation of what skills are, we call this difference the *achievement gap*. All else constant, a larger achievement gap increases the promotion gap. The second term is the difference in promotion probabilities between Blue and Red that arises as a direct consequence of the subtle bias. It is the promotion gap conditional on a tie times the probability of a tie. We call this term the *favoritism gap*. Note that the subtle bias affects the equilibrium investment levels, thus β affects both the achievement gap and the favoritism gap.

The next proposition shows how the equilibrium promotion gap varies with the premium-cost ratio, σ .

Proposition 3. For each $\beta \in (0, 0.5]$, there exists a unique premium-cost ratio $\tilde{\sigma}(\beta)$ such that the promotion gap decreases in σ for $\sigma < \tilde{\sigma}(\beta)$ and increases in σ for $\sigma \in (\tilde{\sigma}(\beta), \bar{\sigma}(\beta))$.

Figure 2 illustrates how the promotion gap changes with the premium-cost ratio, σ . The promotion gap initially decreases with σ and then increases with σ . Note that, for large values of the premium-cost ratio, even a small subtle bias can be significantly amplified through the strategic interactions between the agents.

Note also that, in high-stakes careers, the contribution of the achievement gap to the promotion gap is greater than that of the favoritism gap. That is, differences in “observable” achievements (human capital, performance, experience, effort, etc.) explain most

of the promotion gap. In other words, because ties occur less often as the promotion premium increases, the principal is less likely to make biased promotion decisions as the stakes increase. In such scenarios, we would expect to find little *direct* evidence of discrimination. Nevertheless, promotion gaps are large in high-stakes situations.

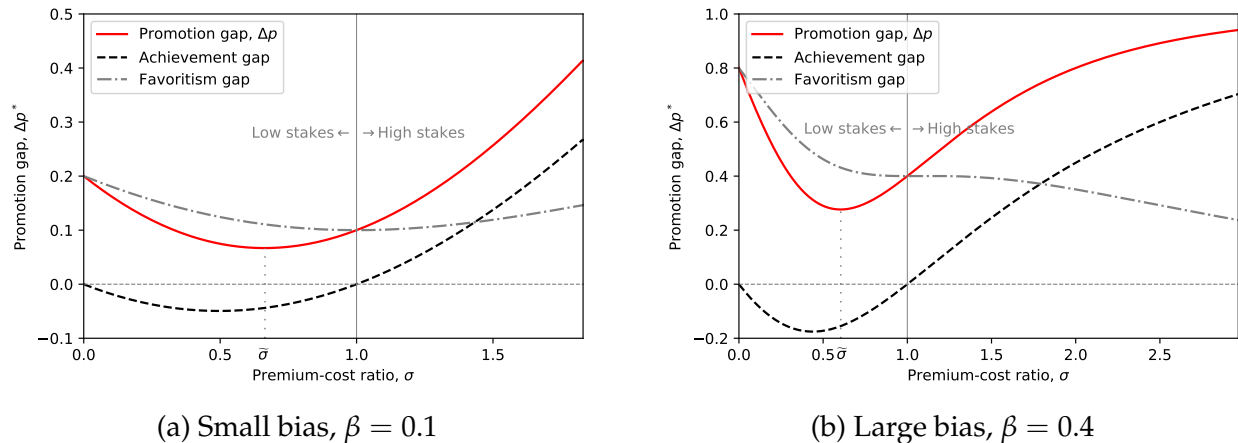


Figure 2: Equilibrium promotion gap, Δp^* , as a function of the premium-cost ratio, σ , for two levels of subtle bias, $\beta_1 = 0.1$ and $\beta_2 = 0.4$.

3.4 Optimal Compensation Contracts

We now allow the principal to design the compensation contract. As standard in principal-agent problems, we assume that the firm is a monopsonist in the labor market; that is, the firm has all the bargaining power. The principal is not allowed to explicitly discriminate through contracts; he must offer the same contract $w = (w_1, W)$ to both agents. Agents are assumed to be penniless; wages must be non-negative: $w_1 \geq 0$ and $w_1 + W \geq 0$.

To remain in a fully rational world, we assume that the principal knows that the agents behave as if promotions are subject to subtle bias β . One interpretation is that the principal is aware of his own bias, but finds it impossible to commit to flipping an unbiased mental coin, i.e., to behave as if $\beta = 0$. Under this interpretation, the subtle bias may create a dynamic inconsistency problem: the principal could be (in some cases) better off by committing not to discriminate, but there is no commitment technology available. In the language of O'Donoghue and Rabin (1999), the principal is a “sophisticate,” i.e., someone

who understands that they will subtly discriminate and, therefore, can correctly predict their future behavior. A second – and perhaps more empirically relevant – interpretation is that promotion decisions are made by a biased third party (e.g., a direct supervisor) and the principal designs the contract taking into account the supervisor’s bias (see Prendergast and Topel (1996) for a model along these lines).

Agents’ outside utilities are normalized to zero. To avoid corner solutions, we assume that the firm pays a fixed entry cost to operate; to save on notation, we assume that this cost is $l + \varepsilon$, with ε arbitrarily small. This assumption implies that the firm chooses to operate if and only if the expected profit after entry is strictly greater than l . The principal is risk-neutral and derives no utility from discrimination. His profit-maximization problem (after entry, i.e., gross of entry costs) is as follows:

$$\max_{w_1 \geq 0, w_1 + W \geq 0} l + H(e_b + e_r - e_b e_r) - 2w_1 - W, \quad (13)$$

subject to

$$e_b = \arg \max_{e \in [0,1]} eW \left[\left(\frac{1}{2} - \beta \right) + 2\beta e_r \right] - \frac{ke^2}{2}, \quad (14)$$

$$e_r = \arg \max_{e \in [0,1]} eW \left[\left(\frac{1}{2} + \beta \right) - 2\beta e_b \right] - \frac{ke^2}{2}. \quad (15)$$

The principal faces two incentive compatibility (IC) constraints (14 and 15). Agents’ participation constraints are not binding because $w_1 \geq 0$ and $w_1 + W \geq 0$ imply that agent i can guarantee a non-negative payoff by choosing $e_i = 0$. Because w_1 does not affect the IC constraints, the principal optimally sets $w_1^* = 0$. If the principal chooses $w_1 = W = 0$, the agents exert no effort, and the post-entry profit is l . In such a case, the firm’s profit from entering the market is $l - l - \varepsilon < 0$. Thus, we use $w_1 = W = 0$ to denote the case in which the firm does not operate.

For any given k , choosing the wage upon promotion, W , is equivalent to choosing the

“stake,” i.e., σ . Parameter σ denotes different contracts with different stakes involved; a high-stakes career path is a contract in which the prize from winning the promotion is high relative to the cost of investment. For both convenience and interpretation, from now on, we think of the principal’s problem as that of choosing σ . Proposition 2 implies that $e_b = 1$ if $\sigma > \bar{\sigma}(\beta)$, thus increasing σ beyond $\bar{\sigma}(\beta)$ has no impact on revenue. That is, in an optimal contract, $\sigma \leq \bar{\sigma}(\beta)$. With these observations, the principal’s problem can be written as follows:

$$\Pi(k, \beta, \theta) = \max_{\sigma \in [0, \bar{\sigma}(\beta)]} k\theta (e_b + e_r - e_b e_r) - k\sigma, \quad (16)$$

subject to

$$e_b = \frac{\sigma(0.5 - \beta) + 2\beta\sigma^2(0.5 + \beta)}{1 + 4\beta^2\sigma^2}, \quad (17)$$

$$e_r = \frac{\sigma(0.5 + \beta) - 2\beta\sigma^2(0.5 - \beta)}{1 + 4\beta^2\sigma^2}, \quad (18)$$

where $\theta \equiv \frac{H}{k}$ is the *productivity-cost ratio* and $\Pi(k, \beta, \theta)$ is the optimal expected profit net of entry costs (as $\varepsilon \rightarrow 0$).

We first solve a baseline case of the above problem with no subtle discrimination ($\beta = 0$). In this case, we can explicitly solve for the optimal contract.

Proposition 4. *If the principal is unbiased ($\beta = 0$), the firm operates if and only if $\theta > 1$ and the optimal stake, $\sigma^* = \frac{2(\theta-1)}{\theta}$, uniquely implements investment levels $e_b^* = e_r^* = \frac{\theta-1}{\theta}$.*

When there is no bias, both agents choose the same investment level in equilibrium. Note that the firm operates only when $\theta > 1$, i.e., the productivity gain for the principal is high relative to the marginal cost of investment for the agents. That is, firms with low productivity-cost ratios prefer to shut down. From a social welfare perspective, all firms should operate because the marginal cost from investing when $e_b = e_r = 0$ is zero (i.e., $c'(0) = 0$), while the marginal social benefit from investing when $e_b = e_r = 0$ is

positive and equal to $H > 0$. Thus, when $\theta \leq 1$, the firm inefficiently stays out of the sector. Such inefficiency occurs because the non-negative wage constraint prevents the firm from extracting all the surplus from the agents.

Consider now the general case in which $\beta \geq 0$. Although it is not possible to solve analytically for the optimal contract in all cases, the existence and uniqueness of the optimal contract are easily established:

Proposition 5. *For every set of parameters ($k > 0, \beta \in [0, 0.5], \theta > 0$), there exists a (generically) unique solution $\sigma(k, \beta, \theta)$ to the principal's problem (if the firm chooses not to operate, we set $\sigma = 0$).*

To save on notation, without loss of generality, from now on, we set $k = 1$. Let $\sigma(\beta, \theta)$ denote the optimal stake. The next result describes the properties of the optimal contract and how it changes with the productivity-cost ratio, θ . Parameter θ can also be interpreted as a measure of the relative importance of human capital at higher hierarchical levels. Thus, for interpretation, we call firms with high θ *human-capital-intensive firms*.

Proposition 6. *For every $\beta \in [0, 0.5]$, there exist values $\underline{\theta}(\beta) < \bar{\theta}(\beta)$ such that:*

1. *If $\theta \leq \underline{\theta}(\beta)$, the optimal stake is $\sigma(\beta, \theta) = 0$ (i.e., the firm does not operate). If $\theta \geq \bar{\theta}(\beta)$, the optimal stake is $\sigma(\beta, \theta) = \bar{\sigma}(\beta)$.*
2. *The optimal stake, $\sigma(\beta, \theta)$, is strictly increasing in $\theta \in [\underline{\theta}(\beta), \bar{\theta}(\beta)]$.*
3. *The firm's profit is strictly increasing in $\theta \geq \underline{\theta}(\beta)$.*

Part 2 of Proposition 6 implies that human-capital-intensive firms (high- θ firms) offer career paths involving higher stakes. Because the optimal stake is increasing in θ , all the comparative statics in the previous subsection are unchanged once we replace σ with θ . In particular, if we define $\tilde{\theta}(\beta)$ as the value of θ such that the optimal stake is $\sigma(\beta, \tilde{\theta}(\beta)) = 1$, we again have that Red invests more than Blue when stakes are low ($\theta < \tilde{\theta}(\beta)$) and Blue

invests more than Red when stakes are high ($\theta > \tilde{\theta}(\beta)$). Finally, Part 3 implies that high- θ firms are more profitable. Thus, we can also use θ as a proxy for firm profitability or productivity. Panels (a) and (b) of Figure 3 illustrate Proposition 6 (for $\beta = 0.4$), while panel (c) shows that similar to Figure 2, the equilibrium promotion gap is U-shaped in the productivity-cost ratio, θ .

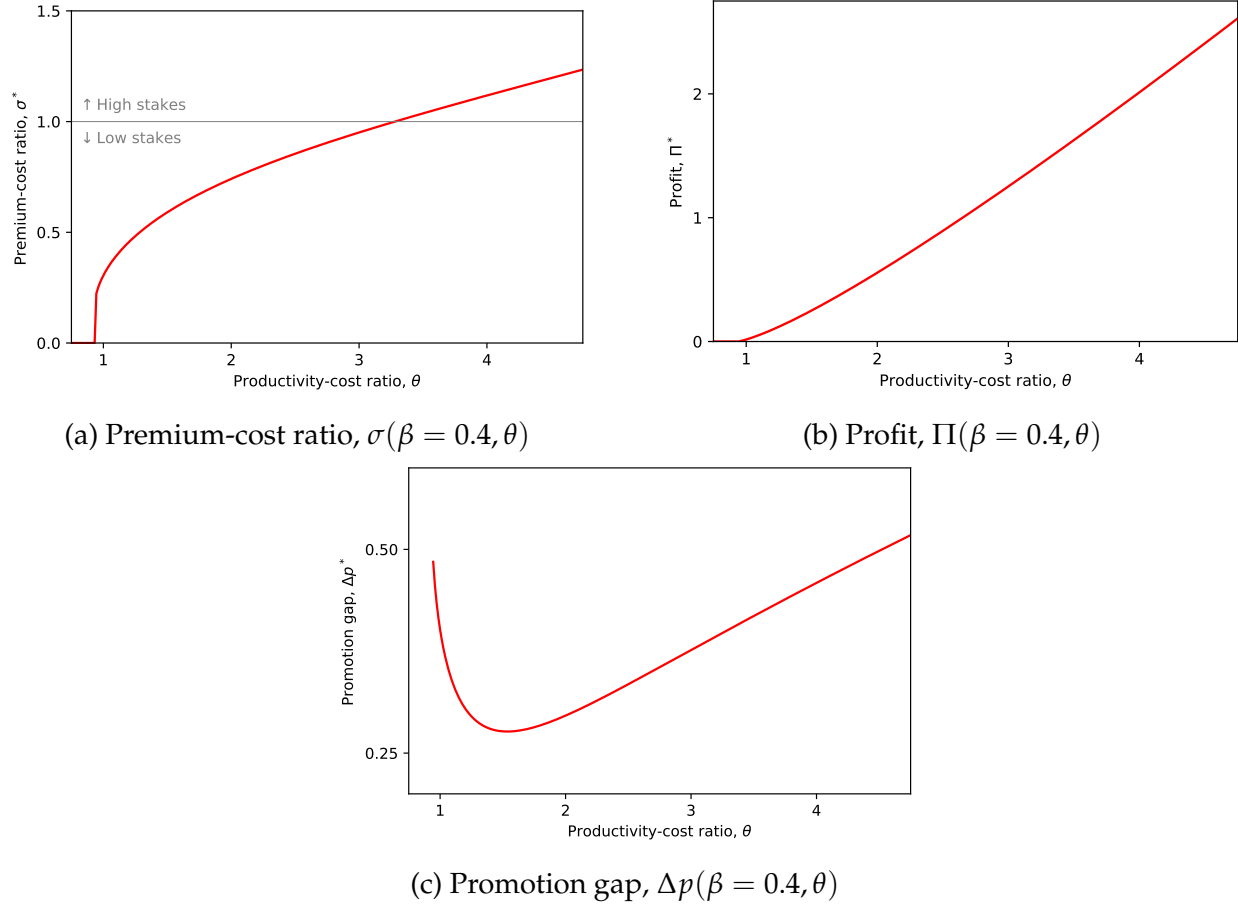


Figure 3: Optimal premium-cost ratio, equilibrium profit, and promotion gap, as functions of the productivity-cost ratio, θ for a given level of subtle bias ($\beta = 0.4$).

3.5 Optimal Anti-Discrimination Policies

Does subtle discrimination benefit or harm firms? To see how subtle discrimination affects profits, we now consider the problem of a principal who can choose both the compensation contract and the firm's own subtle bias. We have in mind a situation in which

the firm chooses an optimal anti-discrimination policy. For example, the firm can set up processes that lead to the selection of supervisors with high or low subtle biases. Similarly, the firm can invest in a corporate culture that is either friendly or hostile to diversity goals. Firms can become conservative by adopting policies associated with high β . Similarly, firms can become progressive by adopting policies associated with low β .

Suppose the principal can select policies associated with bias β at no cost (another interpretation is that market forces may drive firms with suboptimal biases out of the market). Which β would the principal choose? The principal's problem is

$$\Pi(\theta) = \max_{(\sigma, \beta) \in [0, \bar{\sigma}(\beta)] \times [0, 0.5]} \theta (e_b + e_r - e_b e_r) - \sigma, \quad (19)$$

subject to

$$e_b = \frac{\sigma(0.5 - \beta) + 2\beta\sigma^2(0.5 + \beta)}{1 + 4\beta^2\sigma^2}, \quad (20)$$

$$e_r = \frac{\sigma(0.5 + \beta) - 2\beta\sigma^2(0.5 - \beta)}{1 + 4\beta^2\sigma^2}. \quad (21)$$

Let $\beta(\theta)$ denote the profit-maximizing subtle bias and $\sigma(\theta)$ the corresponding optimal stake. Define $\bar{\theta}$ as the lowest value of θ such that $\sigma(\theta) = \bar{\sigma}(\beta(\theta))$. That is, the optimal stake is strictly interior if and only if $\theta \leq \bar{\theta}$. From now on, we focus on strictly interior solutions for σ .

Proposition 7. *There exists $\theta' < \bar{\theta}$ such that*

$$\beta(\theta) = \begin{cases} 0.5 & \text{if } \theta \in (0, \theta'] \\ 0 & \text{if } \theta \in [\theta', \bar{\theta}] \end{cases}. \quad (22)$$

Furthermore, $\sigma(\theta) < 1$ if $\theta \in (0, \theta']$ and $\sigma(\theta) > 1$ if $\theta \in [\theta', \bar{\theta}]$.

This proposition shows that, if the principal could optimally choose his subtle bias (or, equivalently, a supervisor with a given bias) at no cost, he would always choose a corner

solution for the bias: either no bias or the maximum bias. This choice is determined by the productivity-cost ratio. Figure 4 illustrates the optimal stake, profit and the resulting promotion gap as functions of θ . For less productive firms, i.e., firms with low θ , profits increase with subtle discrimination. Thus, firm profit is maximized at $\beta^* = 0.5$. Such firms also choose to offer low-stake careers (i.e., $\sigma(\theta) < 1$). Intuitively, subtle discrimination is profitable for firms that offer low-stakes careers because the overcompensation effect improves the performance of discriminated agents. Thus, in less productive (or less human-capital-intensive) sectors, firms perform better when they discriminate.

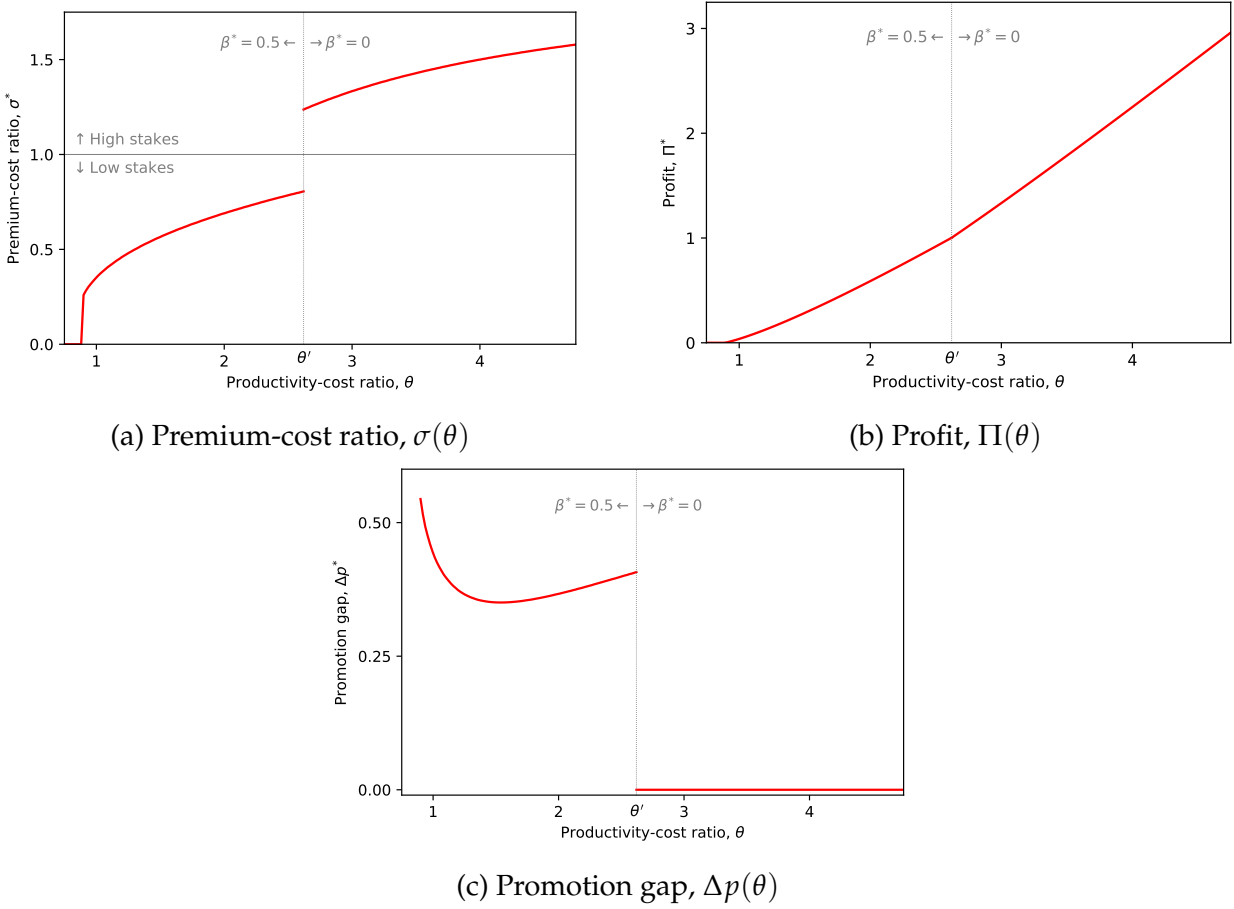


Figure 4: Optimal premium-cost ratio, equilibrium profit, and promotion gap, as functions of the productivity-cost ratio, θ .

By contrast, for high- θ firms, the profit is maximized when the subtle bias is zero. That is, firms that offer high-stakes careers are incentivized to choose robust anti-discrimination

policies. Intuitively, in firms with high-stake careers, the discouragement effect is strong, hindering the performance of discriminated agents. In such sectors, discriminating firms are less profitable than non-discriminating firms, and thus more likely to be driven out by competition.

3.5.1 Discussion: Firm Polarization

When firms optimally choose their biases, high- θ firms do not have a promotion gap (see Figure 4c). That is, such firms have more diversity at the top. By contrast, low- θ firms have positive promotion gaps. Thus, our model predicts a particular type of firm polarization, in which high- and low-productivity firms choose different policies with respect to discrimination and diversity. Note that our model predicts a large promotion gap for high-productivity firms *conditional on a given bias*. When firms can choose their biases, high-productivity firms have small promotion gaps.

High-productivity firms prefer to promote a work environment free of discrimination. These firms strive to be perceived as “progressive” and “activist.” If successful, they would have more diversity at the top (i.e., a smaller promotion gap). These firms also offer careers with higher stakes and are likely to be large, profitable, and human capital-intensive. By contrast, low-productivity firms do not take actions to counter subtle discrimination. These firms do not mind being perceived as “conservative” and will be less diverse at the top. They offer careers with low stakes and are smaller, less profitable, and less human capital-intensive than “progressive” firms. Note also that polarization implies that two firms with productivity-cost ratios $\theta' + \eta$ and $\theta' - \eta$ adopt very different anti-discrimination policies even if their differences in productivity are negligible (i.e., as $\eta \rightarrow 0$).

Firms may be able to achieve their diversity goals through voluntary actions, such as the adoption of a *soft quota* (or “soft affirmative action,” as in Fershtman and Pavan (2021)). Rather than setting a strict numeric target, we can think of soft quotas as a recom-

mendation to promote more red agents whenever possible. Suppose that, to implement a soft quota, the firm adopts a policy in which a supervisor pays a (vanishingly) small cost κ every time they promote a blue agent. For example, the supervisor needs to write a report explaining why the blue agent was more qualified than the red agent. As long as κ is sufficiently small and supervisors have strong incentives to maximize firm profit, the soft quota would only affect supervisors' behavior in tie-breaking situations.

What types of firms adopt soft quotas?¹³ The answer follows from Proposition 7:

Corollary 4. *The firm adopts a soft quota that incentivizes the promotion of red agents if and only if $\theta \geq \theta'$.*

4 Related Literature

Lazear and Rosen (1990) present a model to explain male-female differences in promotion rates. In their model, men have higher promotion rates than equally-qualified women because women have better non-work opportunities and are thus more likely to quit after being promoted. They use their model to explain why gender wage gaps are mainly caused by gender promotion gaps. Lazear and Rosen (1990) claim that the empirical findings of equal pay within jobs and large promotion gaps are “*difficult to incorporate into the main economic theory of discrimination based on taste factors alone*” (p. S107). By contrast, our model shows that subtle discrimination can cause large promotion gaps even when there are no differences in pay across groups within a job. In terms of empirical predictions, their model implies small promotion gaps in high-stake situations (i.e., when workers are more productive upon promotion), while our model predicts the opposite: for a given bias, promotion gaps are larger in high-level jobs.

Our model setup is similar (in spirit, but not in its details) to that of Prendergast (1993). Prendergast (1993) proposes a model of promotions in which the firm cannot contractu-

¹³See the Internet Appendix for an analysis of hard quotas.

ally commit to compensating workers for acquiring firm-specific human capital. Our model differs from his in two significant aspects. First, in our model, promotions are competitive, i.e., candidates compete for a limited number of positions. Second, in our model, the principal is subtly biased in favor of candidates from a particular group.

Our model relates to the literature on favoritism and other biases in subjective performance evaluations and their consequences for selection and promotion decisions (Prenndergast and Topel (1996); MacLeod (2003); Friebe and Raith (2004); Hoffman et al. (2018); Frederiksen et al. (2020); Letina et al. (2020); Frankel (2021); Pagano and Picariello (2022)). In these models, favoritism and other biases have ex-post payoff consequences for the decision-maker. By contrast, in our model, favoritism matters only because it affects ex-ante incentives.

More broadly, our study is related to the theoretical literature on discrimination (see Arrow (1998), Fang and Moro (2011), and Lang and Lehmann (2012) for reviews). In their seminal work on affirmative action, Coate and Loury (1993) show that negative stereotypes can be self-fulfilling because discriminated agents may not undertake investments that make them more productive. Similarly, in our model, discrimination may discourage some agents from investing. However, because workers compete for the same position, their investment decisions are interdependent. We show that such strategic considerations may further discourage investment or, instead, provide discriminated agents with stronger incentives to invest. Thus, differently from Coate and Loury (1993), in our model, the unfavored group may invest more than the favored group. We show that this result obtains only when discrimination is subtle. In addition, the strategic interactions between agents imply a unique equilibrium, which is rare in the literature on self-fulfilling discrimination.¹⁴

In our model, agents impose externalities on each other. In this sense, our model is similar to those by Mailath et al. (2000) and Moro and Norman (2004), who study in-

¹⁴Glover et al. (2017) provide evidence of a different form of self-fulfilling discrimination: workers perform worse when under the supervision of a biased manager.

tegrated labor markets where workers from one group impose externalities on another group. In both models, asymmetric equilibria exist in which agents with identical qualifications receive different wages. That is, discrimination is ex-post observable. By contrast, in our model, wages are not conditional on agents' labels, and therefore discrimination cannot be verified ex-post.

Unlike theories of discrimination based on differential screening abilities (Cornell and Welch (1996); Fershtman and Pavan (2021)), our model assumes that the principal knows each candidate's type. While we can still interpret subtle discrimination as a form of incorrect or exaggerated beliefs, as in Bordalo et al. (2016), it can also be seen as a limiting case of taste-based discrimination when the taste parameter is arbitrarily small.

Our paper is also related to a strand of the discrimination literature that focuses on bias amplification. Lang et al. (2005) show that in markets where firms post wages, weak discriminatory preferences can cause large wage differentials. Bartoš et al. (2016) show how "attention discrimination" can amplify animus and prior beliefs about group quality. Davies et al. (2021) demonstrate that an arbitrary small bias towards one candidate can have large consequences when the principal exerts effort to learn about candidates' abilities. Siniscalchi and Veronesi (2021) present a model in which mild population heterogeneity and self-image bias can lead to persistent differences between groups. Differently from these models, in our model the source of bias amplification is the competitive nature of promotion tournaments. While agents can take actions that amplify the consequences of small biases, we show that these actions can also lead to the attenuation of such biases.¹⁵

Additionally, our paper is related to a small theoretical literature on biased contests (Kawamura and de Barreda (2014); Pérez-Castrillo and Wettstein (2016); Drugov and Ryvkin (2017)). Drugov and Ryvkin (2017) show that under certain conditions, biased contests can be optimal from the organizer's point of view (e.g., total effort maximization)

¹⁵Recently, Kline et al. (2022) show that only a small fraction of U.S. employers discriminate. However, these models suggest that even a small bias can have substantial consequences for the economy.

even when contestants are symmetric. In that vein, Nava and Prummer (2022) present a model in which the principal can directly affect the contestants' valuations of the prize (promotion) through work culture. Our paper differs from this literature in many respects, particularly in our focus on subtle discrimination, its empirical implications, and its consequences for different types of firms.

Although we do not model the preferences and beliefs at the root of subtle discrimination, we note that our notion of subtle discrimination is compatible with models of lexicographic preferences. In particular, our decision-making heuristic can be mapped into Tversky's (1969) notion of lexicographic semiorder; see also Manzini and Mariotti (2012) for a generalization.

Our notion of subtle discrimination is closely related to (but also different from) Cunningham and de Quidt's (2022) concept of implicit preferences. They consider a setup in which a decision maker selects a woman over a man whenever both have identical qualifications but selects a man over a woman when their qualifications are mixed, i.e., they cannot be objectively ranked. Cunningham and de Quidt (2022) equate such choices to an explicit preference for women and an implicit preference for men. Applying our terminology to their example, we say that the decision-maker has a subtle bias towards women in the first case and a subtle bias towards men in the second. Thus, in their model, subtle biases depend on the nature of the tie (i.e., unambiguous versus ambiguous ties).

5 Empirical Implications and Conclusions

Most cases of discrimination we witness in day-to-day life are subtle. Subtle discrimination leaves no trace and is subject to plausible deniability. Although subtle discrimination may harm those at the receiving end of discriminatory actions, it may not have many immediate consequences for the perpetrating parties.

Our leading example of subtle discrimination is the use of biased tie-breaking rules

in promotion contests. When candidates are indistinguishable in terms of their future productivity, the firm is indifferent between biased and unbiased tie-breaking rules. This indifference opens the door for decision rules that favor characteristics unrelated to future productivity. Thus, subtle discrimination may result from a small bias, which manifests itself only when the ex-post consequences are of little significance. However, despite the small size of subtle biases, our model shows that they may have significant ex-ante implications. First, subtle biases in promotion decisions distort candidates' incentives to take actions that increase their promotability. Second, the competitive nature of promotion contests can amplify the ex-ante effects of subtle biases.

Our model generates several novel predictions. The model shows that unfavored agents are discouraged to invest in human capital when promotion stakes are high. While it is not always clear how to measure "promotion stakes," the gain from promotion is likely related to the importance of human capital for performing a task. For example, promotion benefits are widely perceived to be high in professional services careers, such as consulting, law, and finance. Azmat et al. (2020) find that the differences in promotion rates between men and women in law firms are explained by men working more hours (i.e., exerting more effort) than women in entry-level positions. Such evidence is consistent with a discouragement effect in high-stakes careers. In contrast, Benson et al. (2021) find that women on management-track careers in retail have better pre-promotion performance than men. This finding is consistent with an overcompensation effect that dominates in low-stakes situations.

The fact that the promotion gap eventually increases with the promotion premium can explain the "leaky pipe" phenomenon, i.e., increasing promotion gaps at higher hierarchical levels. In hierarchies with convex wage profiles, the net benefit from promotion increases with rank. In such hierarchies, we would expect the promotion gap to increase with rank.

Our model also explains why some firms invest in building a "progressive" corporate

culture while others are content to maintain a “conservative” image. Subtle discrimination is detrimental to high-productivity firms because discriminated workers are discouraged from investing in valuable skills. Thus, such firms prefer to foster equality as a means to incentivize a diverse workforce. By contrast, low-productivity firms benefit from holding discriminated workers to higher standards, as these employees overcompensate by working harder.

Consistent with our predictions, Edmans et al. (2023) find that employees’ perception of diversity, equity and inclusion is stronger in growing, high-valuation, and financially strong firms. Similarly, a robust empirical finding is that, in the cross-section, large and high-performing firms have more women on their boards (see, e.g., Adams and Ferreira (2009)). We are unaware of theoretical work that explains these cross-sectional correlations. Subtle discrimination can explain these findings. High-productivity (i.e., large and profitable) firms may choose to take actions that incentivize the recruitment of women to their boards. These actions would reduce their promotion gaps and thus increase the proportion of women in top jobs.

There are several ways in which one can test for subtle discrimination in competitive situations. One is to identify a direct shock to the bias, i.e., a shock to β . According to our model, an ex-post, unanticipated small shock to β would change the observed promotion gaps between the groups but would have no impact on firm performance in the short run. As an example of this approach, Ronchi and Smith (2021) find evidence that an exogenous shock to male managers’ gender attitudes – the birth of a daughter as opposed to a son – increases managers’ propensity to hire female workers. That is, the shock to gender preferences changes the observed hiring gaps. However, they also find that the shock has no effect on firm performance, which is explained by managers replacing men with women with comparable qualifications, experience, and earnings. Overall, the evidence is consistent with subtle discrimination affecting gender gaps but not profits (in the short run).

Another approach to testing for subtle discrimination is to consider the impact of discrimination on those who are subjected to discrimination (for an example of this approach, see Hengel (2022)). In our context, this requires testing the predictions of our model for the investment choices made by the agents, in particular, how they relate to the stakes faced by the agents. An instructive example – although in a somewhat different context – is the work of Filippin and Guala (2013), who ran a lab experiment where individuals assigned to different groups submit bids in an all-pay auction. The winner is selected by an auctioneer who has strong incentives to reward the highest bidder. Nevertheless, the auctioneers more frequently assign the prize to a member of their own group when bids of two or more players are tied. In response, out-group bidders reduce their bids, leading to a decrease in their earnings and a substantial gap in outcomes between groups.

Finally, direct tests of subtle discrimination can also be designed in the lab. Foschi et al. (1994) designed an experiment where subjects must promote at most one of two candidates. They can also choose to promote no one. When subjects choose between a pair of candidates of the same sex, sometimes no one is promoted. Thus, the authors can infer the minimum threshold of qualifications for promotions for each sex. They find that subjects use similar thresholds for pairs of male and female candidates. That is, they show that men and women are held to the same standards when competing against someone of the same sex. By contrast, when men and women with identical qualifications compete for the same position, they find that subjects are more likely to promote men. Thus, the evidence is consistent with our definition of subtle discrimination but inconsistent with statistical discrimination.

References

- ADAMS, R. B. AND D. FERREIRA (2009): “Women in the boardroom and their impact on governance and performance,” *Journal of Financial Economics*, 94, 291–309.
- ARROW, K. J. (1973): “The theory of discrimination,” in *Discrimination in Labor Markets*, ed. by O. Ashenfelter and A. Rees, Cambridge MA: Princeton University Press, 3–33.
- (1998): “What has economics to say about racial discrimination?” *Journal of Economic Perspectives*, 12, 91–100.
- AZMAT, G., V. CUÑAT, AND E. HENRY (2020): “Gender promotion gaps: Career aspirations and workplace discrimination,” Working paper, Sciences Po.
- BARTOŠ, V., M. BAUER, J. CHYTILOVÁ, AND F. MATĚJKA (2016): “Attention discrimination: Theory and field experiments with monitoring information acquisition,” *American Economic Review*, 106, 1437–75.
- BECKER, G. (1957): *The economics of discrimination*, Chicago: University of Chicago Press.
- BENSON, A., D. LI, AND K. SHUE (2021): ““Potential” and the gender promotion gap,” Working paper, University of Minnesota.
- BIRCAN, Ç., G. FRIEBEL, AND T. STAHL (2022): “Knowledge Teams, Careers, and Gender,” Working paper, Goethe-University Frankfurt.
- BOHREN, J. A., K. HAGGAG, A. IMAS, AND D. G. POPE (2019): “Inaccurate statistical discrimination: An identification problem,” Working paper, National Bureau of Economic Research.
- BORDALO, P., K. COFFMAN, N. GENNAIOLI, AND A. SHLEIFER (2016): “Stereotypes,” *The Quarterly Journal of Economics*, 131, 1753–1794.
- BOSQUET, C., P.-P. COMBES, AND C. GARCÍA-PEÑALOSA (2019): “Gender and promotions: Evidence from academic economists in France,” *The Scandinavian Journal of Economics*, 121, 1020–1053.
- CHEREPANOV, V., T. FEDDERSEN, AND A. SANDRONI (2013): “Rationalization,” *Theoretical Economics*, 8, 775–800.

- COATE, S. AND G. C. LOURY (1993): "Will affirmative-action policies eliminate negative stereotypes?" *The American Economic Review*, 1220–1240.
- CORNELL, B. AND I. WELCH (1996): "Culture, information, and screening discrimination," *Journal of Political Economy*, 104, 542–571.
- COTTERMAN, J. D. (2004): *Compensation plans for law firms*, American Bar Association.
- CUNNINGHAM, T. AND J. DE QUIDT (2022): "Implicit preferences inferred from choice," Working paper, Stockholm University.
- DAVIES, S., E. VAN WESEP, AND B. WATERS (2021): "On the magnification of small biases in decision-making," Working paper, University of Colorado at Boulder.
- DEITCH, E. A., A. BARSKY, R. M. BUTZ, S. CHAN, A. P. BRIEF, AND J. C. BRADLEY (2003): "Subtle yet significant: The existence and impact of everyday racial discrimination in the workplace," *Human Relations*, 56, 1299–1324.
- DHANANI, L. Y., J. M. BEUS, AND D. L. JOSEPH (2018): "Workplace discrimination: A meta-analytic extension, critique, and future research agenda," *Personnel Psychology*, 71, 147–179.
- DIPBOYE, R. L. AND S. K. HALVERSON (2004): *Subtle (and not so subtle) discrimination in organizations*, vol. 16 of *The dark side of organizational behavior*, 131–158.
- DOBBIE, W., A. LIBERMAN, D. PARAVISINI, AND V. PATHANIA (2021): "Measuring bias in consumer lending," *The Review of Economic Studies*, 88, 2799–2832.
- DOVIDIO, J. F. AND S. L. GAERTNER (1986): *Prejudice, discrimination, and racism.*, Academic Press.
- (2000): "Aversive racism and selection decisions: 1989 and 1999," *Psychological science*, 11, 315–319.
- DRUGOV, M. AND D. RYVKIN (2017): "Biased contests for symmetric players," *Games and Economic Behavior*, 103, 116–144.
- EDGECLIFFE-JOHNSON, A. (2022): "The war on 'woke capitalism'," *Financial Times*.

- EDMANS, A. (2020): *Grow the pie: How great companies deliver both purpose and profit—updated and revised*, Cambridge University Press.
- EDMANS, A., C. FLAMMER, AND S. GLOSSNER (2023): “Diversity, Equity, and Inclusion,” Working paper, National Bureau of Economic Research.
- ELIAZ, K. AND R. SPIEGLER (2020): “A model of competing narratives,” *American Economic Review*, 110, 3786–3816.
- ESSED, P. (1991): *Understanding everyday racism: An interdisciplinary theory*, vol. 2, Sage.
- FANG, H. AND A. MORO (2011): “Chapter 5 - Theories of Statistical Discrimination and Affirmative Action: A Survey,” in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. O. Jackson, North-Holland, vol. 1, 133–200.
- FERSHTMAN, D. AND A. PAVAN (2021): ““Soft” Affirmative Action and Minority Recruitment,” *American Economic Review: Insights*, 3, 1–18.
- FILIPPIN, A. AND F. GUALA (2013): “Costless discrimination and unequal achievements in an experimental tournament,” *Experimental Economics*, 16, 285–305.
- FOSCHI, M., L. LAI, AND K. SIGERSON (1994): “Gender and double standards in the assessment of job applicants,” *Social Psychology Quarterly*, 326–339.
- FRANKEL, A. (2021): “Selecting applicants,” *Econometrica*, 89, 615–645.
- FREDERIKSEN, A., L. B. KAHN, AND F. LANGE (2020): “Supervisors and performance management systems,” *Journal of Political Economy*, 128, 2123–2187.
- FRIEBEL, G. AND M. RAITH (2004): “Abuse of authority and hierarchical communication,” *RAND Journal of Economics*, 224–244.
- GLOVER, D., A. PALLAIS, AND W. PARIENTE (2017): “Discrimination as a self-fulfilling prophecy: Evidence from French grocery stores,” *The Quarterly Journal of Economics*, 132, 1219–1260.
- HEBL, M., S. K. CHENG, AND L. C. NG (2020): “Modern discrimination in organizations,” *Annual Review of Organizational Psychology and Organizational Behavior*, 7, 257–282.

- HEBL, M., J. M. MADERA, AND E. KING (2008): *Exclusion, Avoidance, and Social Distancing*, Psychology Press, 157–180, Diversity Resistance in Organizations.
- HEINZ, J. P., R. L. NELSON, R. L. SANDEFUR, AND E. O. LAUMANN (2005): *Urban lawyers: The new social structure of the bar*, University of Chicago Press.
- HENGEL, E. (2022): “Are Women Held to Higher Standards? Evidence from Peer Review,” *The Economic Journal*.
- HOFFMAN, M., L. B. KAHN, AND D. LI (2018): “Discretion in hiring,” *The Quarterly Journal of Economics*, 133, 765–800.
- HOSPIDO, L., L. LAEVEN, AND A. LAMO (2019): “The gender promotion gap: Evidence from central banking,” *Review of Economics and Statistics*, 1–45.
- JONES, K. P., D. F. ARENA, C. L. NITTROUER, N. M. ALONSO, AND A. P. LINDSEY (2017): “Subtle discrimination in the workplace: A vicious cycle,” *Industrial and Organizational Psychology*, 10, 51–76.
- KAHNEMAN, D. (2011): *Thinking, fast and slow*, Macmillan.
- KAWAMURA, K. AND I. M. DE BARREDA (2014): “Biasing selection contests with ex-ante identical agents,” *Economics Letters*, 123, 240–243.
- KLINE, P., E. K. ROSE, AND C. R. WALTERS (2022): “Systemic discrimination among large US employers,” *The Quarterly Journal of Economics*, 137, 1963–2036.
- LANG, K. AND J.-Y. K. LEHMANN (2012): “Racial discrimination in the labor market: Theory and empirics,” *Journal of Economic Literature*, 50, 959–1006.
- LANG, K., M. MANOVE, AND W. T. DICKENS (2005): “Racial discrimination in labor markets with posted wage offers,” *American Economic Review*, 95, 1327–1340.
- LAZEAR, E. P. AND S. ROSEN (1990): “Male-female wage differentials in job ladders,” *Journal of Labor Economics*, 8, S106–S123.
- LETINA, I., S. LIU, AND N. NETZER (2020): “Delegating performance evaluation,” *Theoretical Economics*, 15, 477–509.

- LUNDBERG, S. AND J. STEARNS (2019): "Women in economics: Stalled progress," *Journal of Economic Perspectives*, 33, 3–22.
- MACLEOD, W. B. (2003): "Optimal contracting with subjective evaluation," *American Economic Review*, 93, 216–240.
- MAILATH, G. J., L. SAMUELSON, AND A. SHAKED (2000): "Endogenous inequality in integrated labor markets with two-sided search," *American Economic Review*, 90, 46–72.
- MANZINI, P. AND M. MARIOTTI (2012): "Choice by lexicographic semiorders," *Theoretical Economics*, 7, 1–23.
- MORO, A. AND P. NORMAN (2004): "A general equilibrium model of statistical discrimination," *Journal of Economic Theory*, 114, 1–30.
- NAVA, F. AND A. PRUMMER (2022): "Profitable Inequality," Working paper, Queen Mary University of London.
- NOH, S., V. KASPAR, AND K. A. WICKRAMA (2007): "Overt and subtle racial discrimination and mental health: Preliminary findings for Korean immigrants," *American Journal of Public Health*, 97, 1269–1274.
- O'DONOGHUE, T. AND M. RABIN (1999): "Doing it now or later," *American Economic Review*, 89, 103–124.
- PAGANO, M. AND L. PICARIELLO (2022): "Corporate Governance, Favoritism and Careers," Working paper, University of Naples Federico II.
- PAO, E. K. (2017): *Reset: My fight for inclusion and lasting change*, Random House.
- PÉREZ-CASTRILLO, D. AND D. WETTSTEIN (2016): "Discrimination in a model of contests with incomplete information about ability," *International Economic Review*, 57, 881–914.
- PHELPS, E. S. (1972): "The statistical theory of racism and sexism," *The American Economic Review*, 62, 659–661.
- PRENDERGAST, C. (1993): "The role of promotion in inducing specific human capital acquisition," *The Quarterly Journal of Economics*, 108, 523–534.

- PRENDERGAST, C. AND R. H. TOPEL (1996): "Favoritism in organizations," *Journal of Political Economy*, 104, 958–978.
- RONCHI, M. AND N. SMITH (2021): "Daddy's girl: Daughters, managerial decisions, and gender inequality," Working paper, Bocconi University.
- SHERMAN, M. G. AND H. E. TOOKES (2022): "Female representation in the academic finance profession," *The Journal of Finance*, 77, 317–365.
- SINISCALCHI, M. AND P. VERONESI (2021): "Self-image bias and lost talent," Working paper, NBER.
- TVERSKY, A. (1969): "Intransitivity of preferences," *Psychological Review*, 76, 31.
- VAN LAER, K. AND M. JANSSENS (2011): "Ethnic minority professionals' experiences with subtle discrimination in the workplace," *Human Relations*, 64, 1203–1227.

Appendix

A Proofs

Proof of Proposition 1. Suppose first that both agents undertake strictly positive investments in skill acquisition in the first-best solution. The first-order conditions for an interior solution are

$$H(1 - e_j) - c'(e_i) = 0, \quad (23)$$

for $i \neq j \in \{b, r\}$. Under $c''(e_i) > 0$, an interior solution must be unique, which implies that the solution is symmetric and given by \tilde{e} , where

$$\tilde{e} = 1 - \frac{c'(\tilde{e})}{H}. \quad (24)$$

Note that \tilde{e} is well defined as long as $H > c'(0)$. We extend the definition of \tilde{e} so that $\tilde{e} = 0$ if $H \leq c'(0)$. We can then calculate the surplus associated with \tilde{e} : $\tilde{S} \equiv H\tilde{e}(2 - \tilde{e}) - 2c(\tilde{e})$.

Consider now the case in which only one agent, say b , is requested to exert effort. If $H > c'(0)$, the optimal investment is given by $\hat{e}_b = \min\{c'^{-1}(H), 1\}$. If $H \leq c'(0)$, we set $\hat{e}_b = 0$. The surplus associated with \hat{e}_b is $\hat{S} \equiv H\hat{e}_b - c(\hat{e}_b)$.

The first-best investment levels can take one of two forms. If $\tilde{S} \geq \hat{S}$, the gains from sharing effort are greater than the losses from effort duplication, in which case we have $e_b^{FB} = e_r^{FB} = \tilde{e}$. If, instead, $\tilde{S} < \hat{S}$, effort duplication is too costly, thus the first-best solution is $e_b^{FB} = \hat{e}_b$ and $e_r^{FB} = 0$. □

Proof of Proposition 2. Equations (8) and (9) represent the unique solution to the system of equations given by (6) and (7). From (8), we find that $e_b^* \leq 1$ requires

$$f_b(\sigma) = \beta(2\beta - 1)\sigma^2 - (0.5 - \beta)\sigma + 1 \geq 0.$$

Function f_b is strictly concave and has a unique positive root,

$$\bar{\sigma}(\beta) \equiv \frac{\beta - 0.5 + \sqrt{\frac{1}{4} + 3\beta - 7\beta^2}}{2\beta(1 - 2\beta)} \geq 0,$$

for all $\beta \in (0, 0.5)$. Thus, $e_b^* \leq 1$ if and only if $\sigma \leq \bar{\sigma}(\beta)$. To show that $\bar{\sigma}(\beta) > 1$, note that

$$\beta - 0.5 + \sqrt{\frac{1}{4} + 3\beta - 7\beta^2} = \beta - 0.5 + \sqrt{(\beta - 0.5)^2 + 4\beta(1 - 2\beta)} = -a + \sqrt{a^2 + 2b},$$

where $a = 0.5 - \beta > 0$ and $b = 2\beta(1 - 2\beta) > 0$. It suffices to show that $-a + \sqrt{a^2 + 2b} >$

b. Indeed, $-a + \sqrt{a^2 + 2b} > b \iff a^2 + 2b > a^2 + 2ab + b^2 \iff 2 > b + 2a \iff 2 > 2\beta - 4\beta^2 + 1 - 2\beta \iff 1 > -4\beta^2$.

Similarly, $e_r^* \leq 1$ requires

$$f_r(\sigma) = \beta(2\beta + 1)\sigma^2 - (0.5 + \beta)\sigma + 1 \geq 0.$$

Function f_r is strictly convex. If f_r has no real root, then trivially $e_r^* < 1$ for any value of σ . A pair of real roots exists when $\beta \in \left(0, \frac{1}{14}\right]$. In this case, the smallest real root is:

$$\sigma'(\beta) = \frac{0.5 + \beta - \sqrt{\frac{1}{4} - 3\beta - 7\beta^2}}{2\beta(2\beta + 1)} > 0.$$

Note that $f_r(1) > 0$, and its derivative at $\sigma = 1$ is

$$\frac{\partial f}{\partial \sigma}(\sigma = 1) = 2\beta(2\beta + 1) - (0.5 + \beta),$$

which is strictly negative for $\beta \in \left(0, \frac{1}{14}\right]$. Thus, it must be that $\sigma'(\beta) > 1$. Note also that $f_r(\sigma) - f_b(\sigma) = 2\beta\sigma(\sigma - 1)$, which is positive if and only if $\sigma > 1$. Thus, at $\sigma'(\beta)$ we have $f_r(\sigma'(\beta)) > f_b(\sigma'(\beta))$, which implies $\bar{\sigma}(\beta) < \sigma'(\beta)$.

If $\sigma \leq \bar{\sigma}(\beta)$, then both e_b^* and e_r^* are interior. If $\sigma > \bar{\sigma}(\beta)$, then we must have $e_b^* = 1$, which implies $e_r^* = \min\left\{\frac{\sigma(1-2\beta)}{2}, 1\right\}$. Notice that if $\beta = 0.5$, then $\bar{\sigma} \rightarrow \infty$, and the solution is interior for any σ . \square

Proof of Corollary 1. Note that, from (8) and (9), $e_r^* = e_b^* \frac{(0.5+\beta)-2\beta\sigma(0.5-\beta)}{(0.5-\beta)+2\beta\sigma(0.5+\beta)}$. Straightforward manipulation of this equality implies that $e_r^* \geq e_b^*$ if and only if $\sigma \leq 1$. \square

Proof of Corollary 2. Solving (10) and (11) yields

$$e_b^* = \frac{\sigma(0.5 - \beta) + (2\beta - \delta)\sigma^2(0.5 + \beta - \delta)}{1 + (2\beta - \delta)^2\sigma^2}; \quad (25)$$

$$e_r^* = \frac{\sigma(0.5 + \beta - \delta) - (2\beta - \delta)\sigma^2(0.5 - \beta)}{1 + (2\beta - \delta)^2\sigma^2}, \quad (26)$$

which implies $e_r^* = e_b^* \frac{(0.5+\beta-\delta)-(2\beta-\delta)\sigma(0.5-\beta)}{(0.5-\beta)+(2\beta-\delta)\sigma(0.5+\beta-\delta)}$. Straightforward manipulation of this equality implies that $e_r^* \geq e_b^*$ if and only if $\sigma(1 - \delta) \leq 1$. \square

Proof of Corollary 3. 1. Differentiating (8) with respect to σ yields

$$\frac{\partial e_b^*}{\partial \sigma} = \frac{0.5 - \beta + 4\beta\sigma [(0.5 + \beta) - \beta\sigma(0.5 - \beta)]}{(1 + 4\beta^2\sigma^2)^2},$$

which is strictly positive because $e_r^* \geq 0$ implies $0.5 + \beta - 2\beta\sigma(0.5 - \beta) \geq 0 \Rightarrow 0.5 + \beta - \beta\sigma(0.5 - \beta) > 0$.

2. Differentiating (9) with respect to σ yields

$$\frac{\partial e_r^*}{\partial \sigma} = \frac{(0.5 + \beta) - 4\beta\sigma [(0.5 - \beta) + \beta\sigma(0.5 + \beta)]}{(1 + 4\beta^2\sigma^2)^2}.$$

Note that $\frac{\partial e_r^*}{\partial \sigma} > 0$ for $\sigma = 0$ and the numerator is strictly decreasing in σ (with limit at $-\infty$). Solving for the unique positive root for the numerator yields

$$\sigma_{root}(\beta) \equiv \frac{\sqrt{(0.5 - \beta)^2 + (0.5 + \beta)^2} - (0.5 - \beta)}{2\beta(0.5 + \beta)} > 0.$$

We then define $\hat{\sigma}(\beta) \equiv \min\{\sigma_{root}(\beta), \bar{\sigma}(\beta)\}$.

3.

$$\frac{\partial \sigma_{root}(\beta)}{\partial \beta} = \frac{\left(1 + 2\beta \left(\frac{1}{2} + 2\beta^2\right)^{-\frac{1}{2}}\right) (\beta + 2\beta^2) - (1 + 4\beta) \left[\left(\frac{1}{2} + 2\beta^2\right)^{\frac{1}{2}} - (0.5 - \beta)\right]}{(\beta + 2\beta^2)^2}.$$

The numerator is negative for $\beta = 0$ and decreasing in β for $\beta \in (0, 0.5]$. Thus, we have $\frac{\partial \sigma_{root}}{\partial \beta} < 0$, that is, the region in which e_r^* declines starts earlier for larger values of β . □

Proof of Proposition 3. The equilibrium promotion gap is

$$\Delta p(\sigma) = 2\beta \frac{1 + 2\beta^2(1 + 4\beta^2)\sigma^4 + \left(\frac{3}{2} + 2\beta^2\right)\sigma^2 - 2\sigma}{(1 + 4\beta^2\sigma^2)^2}.$$

Its derivative with respect to σ is

$$\frac{\partial \Delta p}{\partial \sigma} = 2\beta \frac{-2 + 3(1 - 4\beta^2)\sigma + 24\beta^2\sigma^2 + 4\beta^2(4\beta^2 - 1)\sigma^3}{(1 + 4\beta^2\sigma^2)^3}.$$

Define the function $A(\sigma)$ as the the numerator of the expression above:

$$A(\sigma) = -2 + 3(1 - 4\beta^2)\sigma + 24\beta^2\sigma^2 + 4\beta^2(4\beta^2 - 1)\sigma^3.$$

$A(\sigma)$ is a third-degree polynomial of σ , thus, for $\sigma \in \mathbb{R}$, it has three (real or complex) roots (r_1, r_2, r_3) , a local minimum, and a local maximum. Consider its first derivative:

$$A'(\sigma) = 3(1 - 4\beta^2) + 48\beta^2\sigma + 12\beta^2(4\beta^2 - 1)\sigma^2.$$

To find the roots for $A'(\sigma) = 0$, apply the quadratic root formula to obtain

$$\sigma^m = \frac{4\beta - \sqrt{16\beta^2 + (1 - 4\beta^2)^2}}{2\beta(1 - 4\beta^2)}, \sigma^M = \frac{4\beta + \sqrt{16\beta^2 + (1 - 4\beta^2)^2}}{2\beta(1 - 4\beta^2)}.$$

Notice that $\sigma^m < 0$ and $\sigma^M > 0$. At $\sigma = 0$, we have $A(0) = -2 < 0$ and $A'(0) = 3(1 - 4\beta^2) > 0$. Thus, $A(\sigma^m)$ must be a local minimum and $A(\sigma^M)$ a local maximum. Thus, $A(\sigma)$ has one negative real root ($r_1 < \sigma^m$), while $r_2 \leq r_3$ must be positive if they are real numbers.

Notice that at $\sigma = 1$, the condition for an interior solution is trivially satisfied:

$$f_b(1) = \beta(2\beta - 1) - (0.5 - \beta) + 1 = 2\beta^2 + 0.5 > 0.$$

At $\sigma = 1$, we have

$$A(1) = 1 + 8\beta^2 + 16\beta^4 > 0.$$

That is, $\frac{\partial \Delta p}{\partial \sigma}$ is strictly positive at $\sigma = 1$. Thus, a real root $r_2 \in (0, 1)$ must exist; $r_3 > r_2$ must also be a real number. We then have that $\frac{\partial \Delta p}{\partial \sigma} < 0$ for $\sigma \in (0, r_2)$, $\frac{\partial \Delta p}{\partial \sigma} > 0$ for $\sigma \in (r_2, r_3)$, and $\frac{\partial \Delta p}{\partial \sigma} < 0$ for $\sigma > r_3$. Because σ^M is a local maximum, $\sigma^M < r_3$. Brute force comparison reveals that $\sigma^M > \bar{\sigma}$ for all $\beta \in (0, 0.5]$. Thus, σ^M cannot be an interior solution $\Rightarrow r_3 > \sigma^M$. Thus, in an interior solution, $\frac{\partial \Delta p}{\partial \sigma} < 0$ for $\sigma < r_1$ and $\frac{\partial \Delta p}{\partial \sigma} > 0$ for $\sigma > r_2$. We thus have that $\Delta p(\sigma)$ reaches a minimum at $\min\{r_2, \bar{\sigma}\} \equiv \tilde{\sigma}$. \square

Proof of Proposition 4. If $\beta = 0$, we have that, in an interior solution, $e_r = e_b = \frac{\sigma}{2}$. The principal's problem is thus

$$\max_{\sigma \in [0, \bar{\sigma}(0)]} \theta \left[1 - \left(1 - \frac{\sigma}{2} \right)^2 \right] - \sigma.$$

The first order condition for an interior solution is

$$\theta \left(1 - \frac{\sigma}{2} \right) - 1 = 0.$$

The second-order condition holds (the problem is globally concave): $-\frac{\theta}{2} < 0$. Thus, we have

$$\sigma^* = 2\frac{\theta - 1}{\theta}, e^* = \frac{\theta - 1}{\theta}.$$

Notice that for all $\theta \geq 1$, the solution is interior, and for all $\theta < 1$ the principal does not operate the firm. \square

Proof of Proposition 5. Notice that the firm can guarantee a non-negative profit by choosing $\sigma = 0$. Because the objective function is continuous in σ and $[0, \bar{\sigma}(\beta)]$ is a compact set, a maximum always exists. An optimal σ^* is generically unique because the objective function is a function of polynomials and thus has no flat regions in the interior of $[0, \bar{\sigma}(\beta)]$. \square

Proof of Proposition 6. Define

$$\sigma(\beta, \theta) \equiv \arg \max_{\sigma \in [0, \bar{\sigma}(\beta)]} \theta f(\sigma, \beta) - \sigma,$$

where

$$f(\sigma, \beta) = e_b(\sigma, \beta) + e_r(\sigma, \beta) - e_b(\sigma, \beta)e_r(\sigma, \beta),$$

where $e_b(\sigma, \beta)$ and $e_r(\sigma, \beta)$ are given by (8) and (9), respectively. From Proposition 5, the optimal σ is generically unique, thus $\sigma(\beta, \theta)$ is well-defined (except perhaps for a measure-zero combination of parameters (β, θ)). The maximum profit is thus defined as

$$\Pi(\beta, \theta) \equiv \theta f(\sigma(\beta, \theta), \beta) - \sigma(\beta, \theta).$$

First notice that, for $\sigma(\beta, \theta) > 0$, we have that the optimal profit strictly increases with θ (by the Envelope Theorem):

$$\frac{\partial \Pi}{\partial \theta} = f(\sigma(\beta, \theta), \beta) > 0. \quad (27)$$

To prove Part 1, notice first that at $\theta = 0$, trivially, $\sigma(\beta, 0) = 0$ and the profit is zero. For $\theta = \bar{\sigma}(\beta) + \varepsilon$, where $\varepsilon > 0$, if the principal chooses $\sigma = \bar{\sigma}(\beta)$ we have $f(\bar{\sigma}(\beta), \beta) = 1$ and the profit is strictly positive. Thus, we know that there exists $\underline{\theta}(\beta)$ such that $\sigma(\beta, \theta) > 0$ (and the profit is strictly positive) if and only if $\theta > \underline{\theta}(\beta)$.

Now define $\bar{\theta}(\beta)$ as

$$\bar{\theta}(\beta) \equiv \frac{1}{f_\sigma(\bar{\sigma}(\beta), \beta)}, \quad (28)$$

where f_σ denotes the derivative with respect to σ (note that $f(\sigma, \beta)$ is differentiable in σ in the interior of $[0, \bar{\sigma}(\beta)]$). We then have $\sigma(\beta, \bar{\theta}(\beta) + \varepsilon) = \bar{\sigma}(\beta)$ for all $\varepsilon \geq 0$. This proves Part 1.

To prove Part 2, consider $\theta \in (\underline{\theta}(\beta), \bar{\theta}(\beta))$, that is, the values for θ such that $\sigma(\beta, \theta)$ is interior, i.e., $\sigma(\beta, \theta) \in (0, \bar{\sigma}(\beta))$. Thus, the first-order condition at $\sigma^* = \sigma(\beta, \theta)$ must hold:

$$\frac{\partial \Pi}{\partial \sigma} = \theta f_{\sigma}(\sigma^*, \beta) - 1 = 0,$$

as well as the second order condition:

$$\frac{\partial^2 \Pi}{\partial \sigma^2} = \theta f_{\sigma\sigma}(\sigma^*, \beta) < 0.$$

We have that (by implicit differentiation of the first order condition):

$$\frac{\partial \sigma}{\partial \theta} = -\frac{f_{\sigma}(\sigma^*, \beta)}{\theta f_{\sigma\sigma}(\sigma^*, \beta)} = -\frac{1}{\theta^2 f_{\sigma\sigma}(\sigma^*, \beta)} > 0,$$

proving Part 2. Part 3 follows from (27). □

Proof of Proposition 7. For $e_b^* < 1$ (i.e., a strictly interior solution for effort levels), define $f(\sigma, \beta)$ as

$$f(\sigma, \beta) = e_b^* + e_r^* - e_b^* e_r^* = \frac{\sigma + 4\beta^2 \sigma^2}{1 + 4\beta^2 \sigma^2} - \frac{4\beta^2 \sigma^3 + (1 - 4\beta^2 \sigma^2)(\frac{1}{4} - \beta^2) \sigma^2}{(1 + 4\beta^2 \sigma^2)^2}.$$

If $e_b^* < 1$ we can write the profit as

$$\Pi(\sigma, \beta, \theta) = \theta f(\sigma, \beta) - \sigma.$$

We then have

$$\frac{\partial \Pi}{\partial \beta} = -\frac{2\beta\theta\sigma^2(\sigma - 1)\{4\sigma^2(\sigma + 1)\beta^2 - 3\sigma + 5\}}{(4\sigma^2\beta^2 + 1)^3},$$

which has non-negative roots at $\beta = 0$ and

$$\beta_{root}(\sigma) = \frac{1}{2\sigma} \sqrt{\frac{3\sigma - 5}{\sigma + 1}}.$$

Note that for $\sigma < 1$, $\frac{\partial \Pi}{\partial \beta}$ is strictly positive for $\beta > 0$, implying that the optimal bias is $\beta = 0.5$. At $\sigma \in (1, \frac{5}{3})$, the derivative is strictly negative for $\beta > 0$, implying that the optimal bias is $\beta = 0$. For $\sigma > 5/3$, $\frac{\partial \Pi}{\partial \beta}$ is positive for $\beta < \beta_{root}(\sigma)$ and negative for $\beta > \beta_{root}(\sigma)$, implying that the optimal bias is $\beta_{root}(\sigma)$.

Define the following:

$$\begin{aligned}\sigma(\beta, \theta) &= \arg \max_{\sigma \in [0, \bar{\sigma}(\beta)]} \Pi(\sigma, \beta, \theta), \\ \beta(\sigma, \theta) &= \arg \max_{\beta \in [0, 0.5]} \Pi(\sigma, \beta, \theta), \\ \sigma(\theta) &= \arg \max_{\sigma \in [0, \bar{\sigma}(\beta(\sigma, \theta))]} \Pi(\sigma, \beta(\sigma, \theta), \theta).\end{aligned}$$

For now we assume that θ is such that $\sigma(\theta) < 5/3$, so that the optimal profit is

$$\Pi(\theta) = \max\{\Pi(\sigma(0, \theta), 0, \theta), \Pi(\sigma(0.5, \theta), 0.5, \theta)\}.$$

Define

$$\Delta(\theta) = \Pi(\sigma(0, \theta), 0, \theta) - \Pi(\sigma(0.5, \theta), 0.5, \theta),$$

and let θ' denote an element of $\{\theta : \Delta(\theta) = 0\}$. We know that at least one such θ' exists because: (i) $\Pi(\sigma(0.5, 1), 0.5, 1) \geq \Pi(0.5, 0.5, 1) = 0.02 > \Pi(\sigma(0, 1), 0, 1) = 0$ (see Proposition 4) and (ii) $\Pi(\sigma(0.5, 4), 0.5, 4) = 2 < \Pi(\sigma(0, 4), 0, 4) = 2.25$. By continuity there must be a $\theta' \in (1, 4)$ ¹⁶ such that $\Pi(\sigma(0.5, \theta'), 0.5, \theta') = \Pi(\sigma(0, \theta'), 0, \theta')$.

We need to show that θ' is unique. By the Envelope Theorem,

$$\frac{\partial \Delta(\theta)}{\partial \theta} = f(\sigma(0, \theta), 0) - f(\sigma(0.5, \theta), 0.5).$$

If $\frac{\partial \Delta(\theta')}{\partial \theta} > 0$ for all $\theta' \in \{\theta : \Delta(\theta) = 0\}$, then θ' is unique. To show that this is indeed the case, note first that at θ' , it must be that $\sigma(0.5, \theta') \leq 1$, otherwise $\frac{\partial \Pi}{\partial \beta} < 0$ and thus $\Pi(\sigma(0, \theta'), 0, \theta') - \Pi(\sigma(0.5, \theta'), 0.5, \theta') > 0$. Similar reasoning implies that $\sigma(0, \theta') \geq 1$. We then have that $\Delta(\theta') = 0$ implies

$$f(\sigma(0, \theta'), 0) - f(\sigma(0.5, \theta'), 0.5) = \frac{\sigma(0, \theta') - \sigma(0.5, \theta')}{\theta'} > 0.$$

Thus, θ' is unique. Notice that $\theta' < \bar{\theta}$. If not, at $\bar{\theta}$ we have $\Delta(\bar{\theta}) < 0$, i.e., the optimal bias is $\beta = 0.5$. From Proposition 2, $\sigma(0.5, \bar{\theta}) > 1$. But then $\frac{\partial \Pi}{\partial \beta}$ is strictly negative for $\beta > 0$, thus the optimal bias cannot be $\beta = 0.5$.

Consider now values for θ such that $\sigma(\theta) \geq 5/3$. In any strictly interior solution for e_b^* , we have (after simplification)

$$f(\sigma, \beta_{root}(\sigma)) = \frac{\sigma^2 + 2\sigma + 25}{32},$$

¹⁶Numerically, we obtain that $\theta' \approx 2.62054$.

and thus

$$\Pi(\sigma, \beta_{root}(\sigma), \theta) = \theta \frac{\sigma^2 + 2\sigma + 25}{32} - \sigma$$

and

$$\frac{\partial \Pi(\sigma, \beta_{root}(\sigma), \theta)}{\partial \sigma} = \theta \frac{2\sigma + 2}{32} - 1,$$

implying that $\Pi(\sigma, \beta_{root}(\sigma), \theta)$ has a global minimum at $\sigma = 16 - \theta$. Thus, at any $\sigma \geq \frac{5}{3}$ with $e_b^* < 1$, the principal prefers either to increase or decrease σ . Thus, there is no strictly interior solution in which $\sigma(\theta) \geq \frac{5}{3}$. That is, $\sigma(\bar{\theta}) < \frac{5}{3}$.

□

Subtle Discrimination

Internet Appendix

Elena S. Pikulina

Daniel Ferreira

July 2023

1 Continuous Skill Levels

Our model may offer an impression that subtle discrimination matters only when skill levels are discrete and, thus, workers can be equally qualified with positive probability. This impression is false. While the case of discrete skill levels is compelling and realistic, it is not a requirement for any of the key results in the model. According to our definition, subtle discrimination can happen whenever the decision-maker can credibly claim that he used his private information to choose between two candidates. The case in which ties are exact is an obvious application of this idea, but subtle biases can operate even when observable differences in skills are large, provided that the principal can plausibly deny being biased.

Here we consider a variation of our model in which agents can choose to acquire a skill level in the unit interval, i.e., $s_i \in [0, 1]$. For simplicity, we assume that $e_i = s_i$ (i.e., effort deterministically translates into skill). Thus, the difference in observable skill is $\Delta s = e_r - e_b$. As in Section 2, let $F(\cdot)$ denote the c.d.f. of Δx . To simplify the analysis, suppose that $\underline{x} = -1$, $\bar{x} = 1$, $F(\cdot)$ is uniform, and $\omega = 1$. That is, for *any* differences in skill Δs , the principal can plausibly justify promoting Blue. Thus, subtle discrimination may occur for any Δs .¹

¹This is obviously not necessary for the results, but helps making the point that there is nothing special about the “ties” in our main model.

Let us first consider how an unbiased decision-maker would choose between the two candidates. For given e_b and e_r , the probability that Blue is chosen is given by

$$p(e_b, e_r) = \frac{1}{2}(1 + e_b - e_r). \quad (1)$$

In the contest literature, $p(e_b, e_r)$ is known as the *contest success function*, or CSF. Below we show that the key to our results is how subtle biases affect the CSF.

Under the unbiased CSF, the agents' expected utilities are

$$\sigma \frac{1}{2}(1 + e_b - e_r) - e_b^2 \quad (2)$$

$$\sigma \left[1 - \frac{1}{2}(1 + e_b - e_r) \right] - e_r^2 \quad (3)$$

leading to first-order conditions:

$$\frac{\sigma}{2} - e_b^* = 0 \quad (4)$$

$$\frac{\sigma}{2} - e_r^* = 0, \quad (5)$$

i.e., $e_b^* = e_r^* = \frac{\sigma}{2}$, as in our main model.

Suppose now the principal is subtly biased. Then, we write the CSF as

$$p(e_b, e_r) = \max \left\{ \frac{1}{2}(1 + e_b - e_r) + b(e_b, e_r), 1 \right\}, \quad (6)$$

where $b(e_b, e_r)$ is the bias in favor of Blue. The bias $b(e_b, e_r)$ may depend on the observed skill levels because the cost to the principal from making a biased decision may depend on the skill levels. How subtle discrimination distorts effort levels thus depends on the particular functional form of $b(e_b, e_r)$.

Consider first the case in which $b(e_b, e_r) = \beta e_b$, for $\beta \in (0, 0.5]$. Intuitively, this function illustrates a situation where the principal finds it easier to justify choosing Blue when Blue

is more objectively qualified. Assuming an interior solution, the first-order conditions are:

$$\sigma \left(\frac{1}{2} + \beta \right) - e_b^* = 0 \quad (7)$$

$$\frac{\sigma}{2} - e_r^* = 0, \quad (8)$$

which implies $e_b^* > e_r^*$. (An interior solution always obtains for $\sigma < \frac{2}{1+2\beta}$). Thus, subtle discrimination produces an *encouragement effect* for Blue. If $\sigma > \frac{2}{1+2\beta}$, the equilibrium is $e_b^* = 1$ and $e_r^* = 0$. That is, in high-stakes situations, subtle discrimination leads to a discouragement effect for Red.

Suppose now that $b(e_b, e_r) = \beta(1 - e_r)$, for $\beta \in (0, 0.5]$. Intuitively, this represents a case in which the principal finds it easier to justify choosing Blue when Red is less objectively qualified. Assuming an interior solution, the first-order conditions are:

$$\frac{\sigma}{2} - e_b^* = 0 \quad (9)$$

$$\sigma \left(\frac{1}{2} + \beta \right) - e_r^* = 0, \quad (10)$$

which implies $e_b^* < e_r^*$. Thus, in this case, subtle discrimination produces an *overcompensation effect* for Red. (Again, an interior solution always obtains for $\sigma < \frac{2}{1+2\beta}$).

These two examples illustrate that, under continuous skill levels and subtle discrimination, it is possible to generate either the discouragement effect or the overcompensation effect by suitably choosing different bias functions $b(e_b, e_r)$. The simple functions we present here can only generate one of such effects at a time. But by choosing more complex bias functions one can produce examples in which either effect may dominate, depending on the parameters. It is, however, difficult to justify more complex bias functions based only on intuition or introspection. Our main model with discrete skills and arbitrarily small subtle bias ($\omega \rightarrow 0$) provides a microfoundation for the bias function

$$b(e_b, e_r) = \beta(1 - e_b - e_r + 2e_b e_r). \quad (11)$$

Thus, we can replicate all the results in the paper using the continuous approach devel-

oped here if we use the bias function in (11).

To conclude, we show that the equilibrium effort levels depend on how the subtle bias affects the contest success function. Our main model shows a natural example where the subtle bias may lead to either discouragement or overcompensation. Importantly, our model has predictions for when each of these effects is likely to dominate. The simple examples in this Appendix subsection show that models with continuous skills and non-infinitesimal subtle biases can also generate overcompensation and discouragement.

2 General Functional Form of Investment Cost

In this section we show that our main results hold when we use a more general cost function $c(e)$, such that $c(0) = 0, c(1) \rightarrow \infty, c'(\cdot) > 0, c''(\cdot) > 0$. In particular, we (numerically) solve the agents' and the principal's problems for the following cost function:

$$c(e) = \frac{k}{\alpha} \frac{e^\alpha}{1 - e^\gamma}, \quad (12)$$

The above form has several advantages. First, for $\alpha = 2$ and $\gamma \rightarrow \infty$, it converges to the quadratic cost function $\frac{ke^2}{2}$ used in the main text. Second, for the agent's problem, it guarantees an interior solution for any value of the premium-cost ratio, $\sigma \equiv \frac{W}{k}$. Finally, we confirm that for a sizable interval of parameters values α and γ and for any value of the productivity-cost ratio θ , the social welfare is maximized when both agents are treated symmetrically, that is, invest in their human capital. In particular, we define social surplus under the asymmetric treatment (only one agent invests in her human capital) as:

$$s_{1a}(\theta; \alpha, \gamma) = \max_e \theta e - \frac{k}{\alpha} \frac{e^\alpha}{1 - e^\gamma}, \quad (13)$$

and under the symmetric treatment (both agents invest) as:

$$s_{2a}(\theta; \alpha, \gamma) = \max_e \theta e(2 - e) - 2 \times \frac{k}{\alpha} \frac{e^\alpha}{1 - e^\gamma}. \quad (14)$$

Then, for $\alpha \in [2.0, 10.0]$, we compute $\bar{\gamma}$ such that for all $\gamma < \bar{\gamma}$ (see Figure IA.1), the

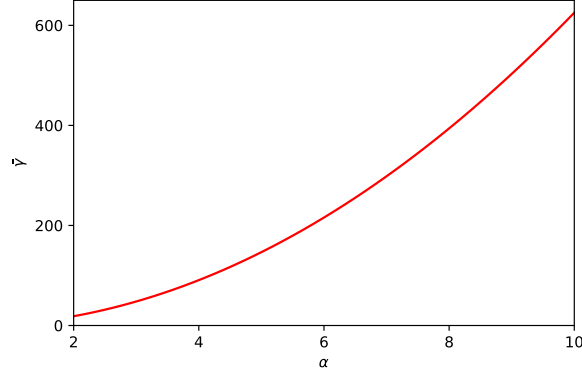


Figure IA.1: $\bar{\gamma}$ as a function of the cost function parameter α

social welfare is maximized when both agents invest in their human capital for all values of the firm productivity parameter θ . That is, for all values of θ , $s_{2a}(\theta; \alpha, \gamma) > s_{1a}(\theta; \alpha, \gamma)$ as long as $\alpha \in [2.0, 10.0]$ and $\gamma < \bar{\gamma}$. This way we confirm that our results in the main text are not driven by the fact that under a quadratic cost function, for $\theta > 1$ it is more socially efficient to treat agents asymmetrically, i.e., ask only one of them to invest. In the remainder of this appendix we assume $\alpha = 2.0$ ($\bar{\gamma} \approx 18.5013$ for $\alpha = 2.0$).

Figure IA.2 shows the equilibrium investment levels as a function of the premium-cost ratio, σ , for two levels of the subtle bias, $\beta \in \{0.1, 0.4\}$ and the following levels of the cost function parameters, $\gamma \in \{0.5, 9.0, 18.0\}$. Note that for $\alpha = 2.0$, all these values of γ are below $\bar{\gamma}$. Figure IA.2 confirms the results in Proposition 2 and Figure 1 of the main text. In particular, it shows that for low values of σ , Red invests more than Blue, while for high values of σ , it is Blue who invests more. Therefore, the existence of the overcompensation and the discouragement effects is not driven by our choice of the quadratic cost function.

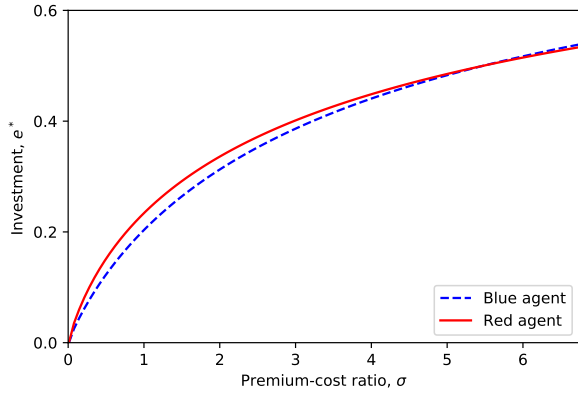
Figure IA.3 shows the equilibrium promotion gaps for the same values of β and γ as functions of the premium-cost ratio σ (it replicates the results in Proposition 3 and Figure 2 of the main text). Again, we confirm that for a wide range of parameters, the promotion gap has U -shape and that under high values of the premium-cost ratio σ , the contribution of the achievement gap to the promotion gap is lower than under low values of σ .

Figure IA.4 shows the solution for the principal's problem, the optimal premium-cost ratio, $\sigma^*(\theta; \beta, \gamma, \alpha)$, the optimal profit, $\Pi^*(\theta; \beta, \gamma, \alpha)$, and the resulting promotion gap, $\Delta p^*(\theta; \beta, \gamma, \alpha)$ as functions of the productivity-cost ratio θ and for given values of the

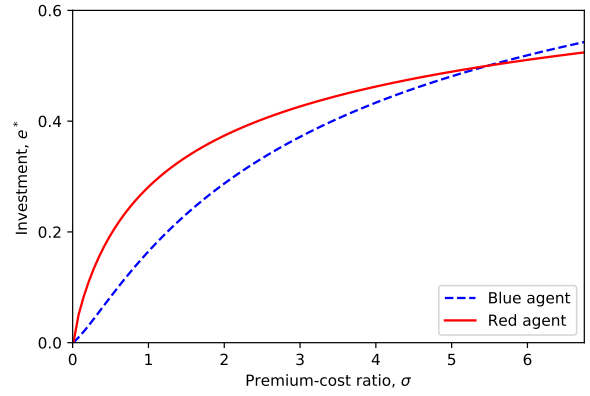
subtle bias β and parameters γ and α . In particular, $\beta \in \{0.1, 0.4\}$, $\gamma \in \{0.5, 9.0, 18.0\}$ and $\alpha = 2.0$.

Next, we consider a situation when a firm could optimally choose its own subtle bias. In particular, we confirm that there exists θ' such that for all $\theta < \theta'$, the firm prefers the maximum level of subtle bias, $\beta^* = 0.5$ and for all $\theta > \theta'$, the firm prefers the minimum level of subtle bias, $\beta^* = 0$. Figure IA.5 shows θ' as a function of γ , where $\gamma \in [0.5, 18.0]$ and $\alpha = 2$. Therefore, we confirm that firm polarization occurs even under a more general cost function.

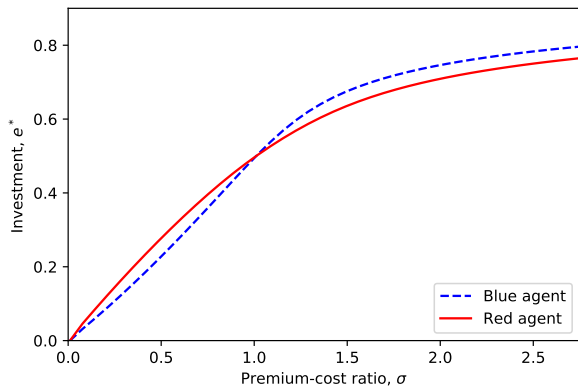
Finally, in Figure IA.6 we replicate the results in Figure 4 from the main text for $\gamma = 9.0$ and $\alpha = 2.0$. For other values of the cost function parameters, γ and α , the optimal stake, σ^* , the resulting profit, Π^* , and the promotion gap, Δp^* have similar shapes. These figures are available upon request.



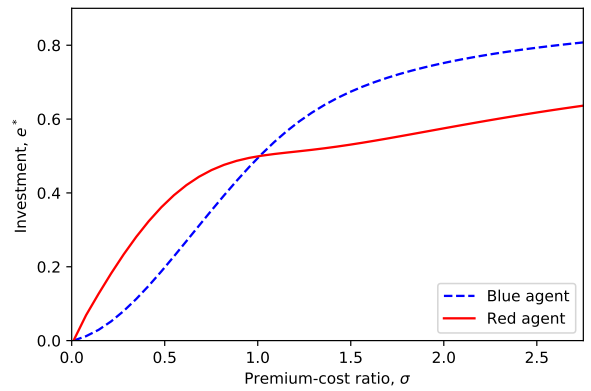
(a) $\beta = 0.1, \gamma = 0.5$



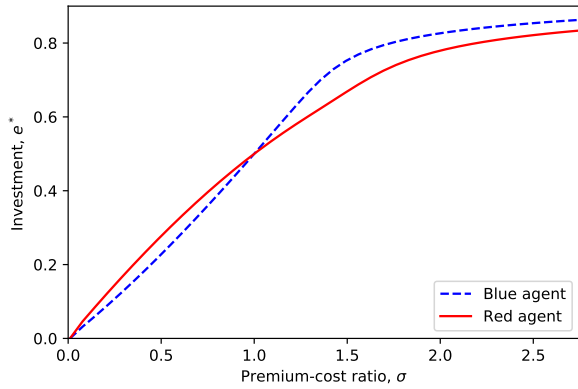
(b) $\beta = 0.4, \gamma = 0.5$



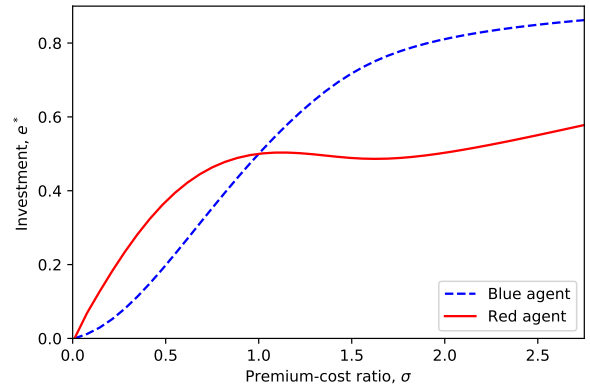
(c) $\beta = 0.1, \gamma = 9.0$



(d) $\beta = 0.4, \gamma = 9.0$

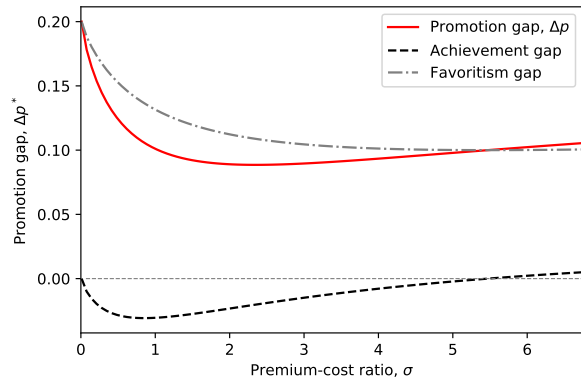


(e) $\beta = 0.1, \gamma = 18.0$

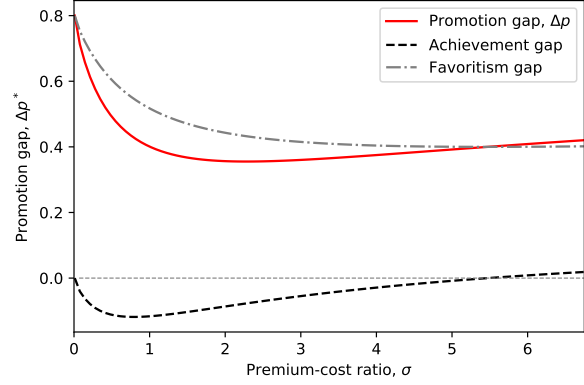


(f) $\beta = 0.4, \gamma = 18.0$

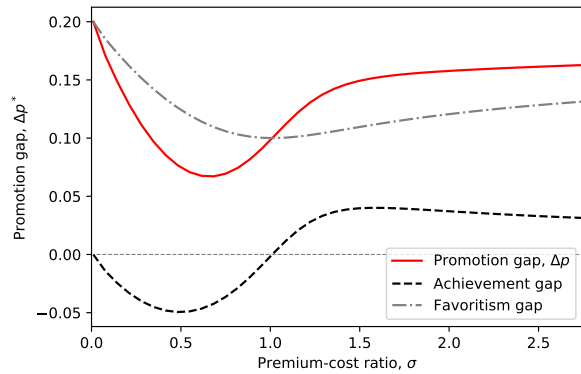
Figure IA.2: Equilibrium investments of blue and red agents, e_b^* and e_r^* , as functions of the premium-cost ratio, σ , for different values of the subtle bias β and the cost function parameter γ .



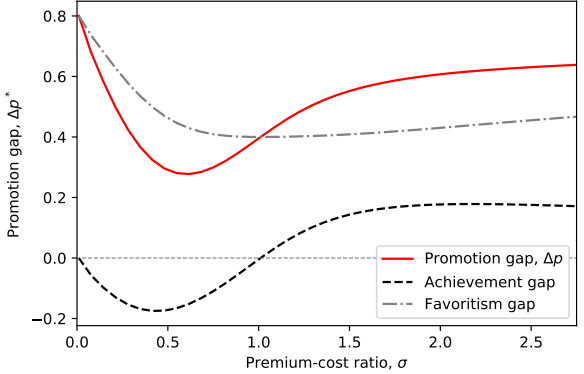
(a) $\beta = 0.1, \gamma = 0.5$



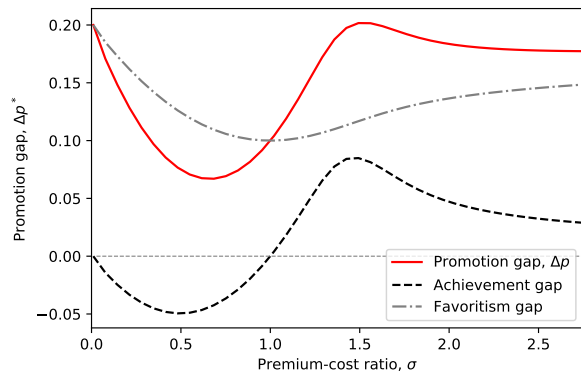
(b) $\beta = 0.4, \gamma = 0.5$



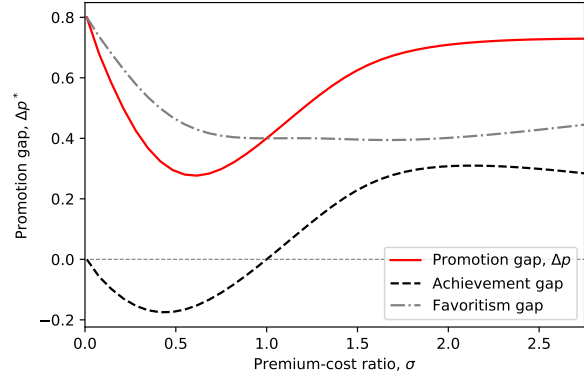
(c) $\beta = 0.1, \gamma = 9.0$



(d) $\beta = 0.4, \gamma = 9.0$

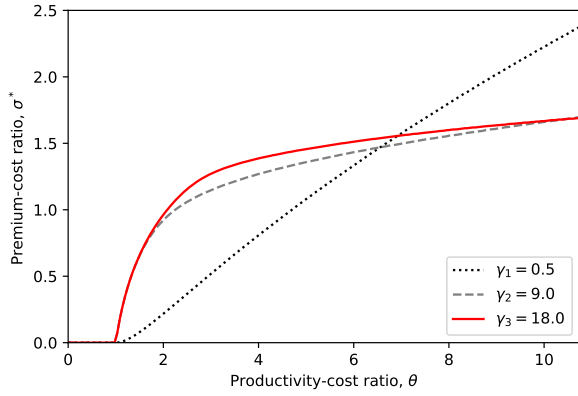


(e) $\beta = 0.1, \gamma = 18.0$

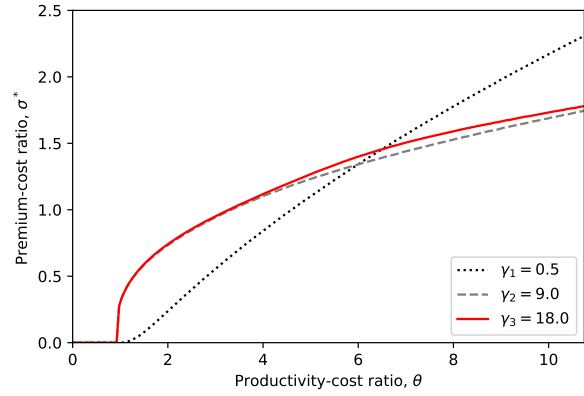


(f) $\beta = 0.4, \gamma = 18.0$

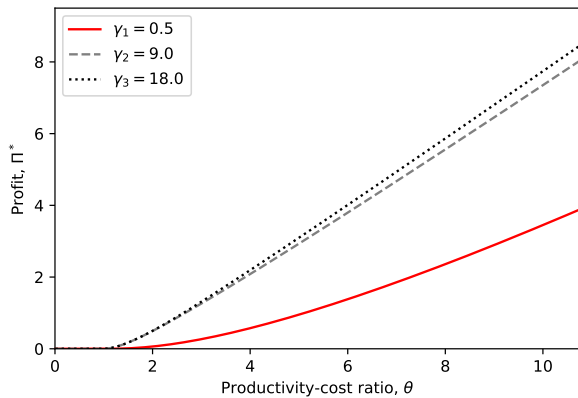
Figure IA.3: Equilibrium promotion gap, Δp^* , as a function of the premium-cost ratio, σ , for different values of the subtle bias β and the cost function parameter γ .



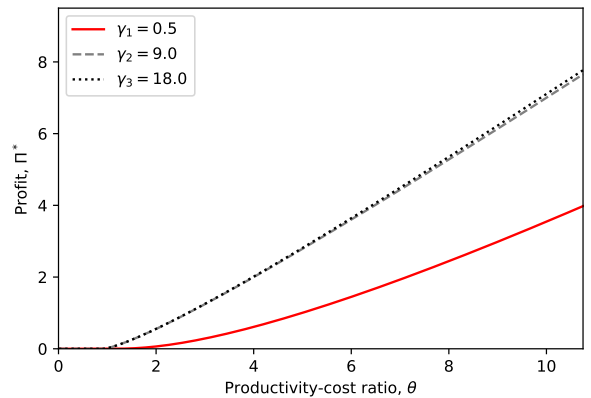
(a) Premium-cost ratio, σ^* , for $\beta = 0.1$



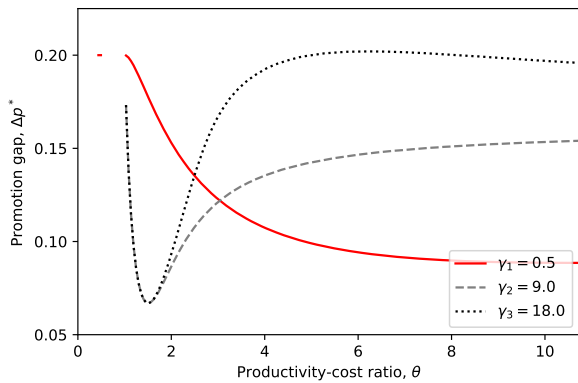
(b) Premium-cost ratio, σ^* , for $\beta = 0.4$



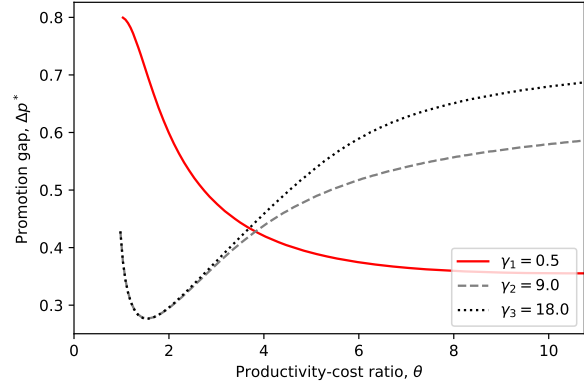
(c) Firm profit, Π^* , for $\beta = 0.1$



(d) Firm profit, Π^* , for $\beta = 0.4$



(e) Promotion gap, Δp^* , for $\beta = 0.1$



(f) Promotion gap, Δp^* , for $\beta = 0.4$

Figure IA.4: Optimal premium-cost ratio, σ^* , firm profit, Π^* , and promotion gap, Δp^* , as a function of the productivity-cost ratio, θ , for different values of the subtle bias β and the cost function parameter γ .

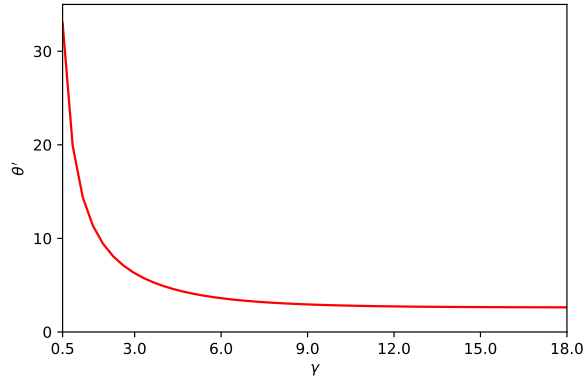


Figure IA.5: θ' as a function of the cost function parameter γ

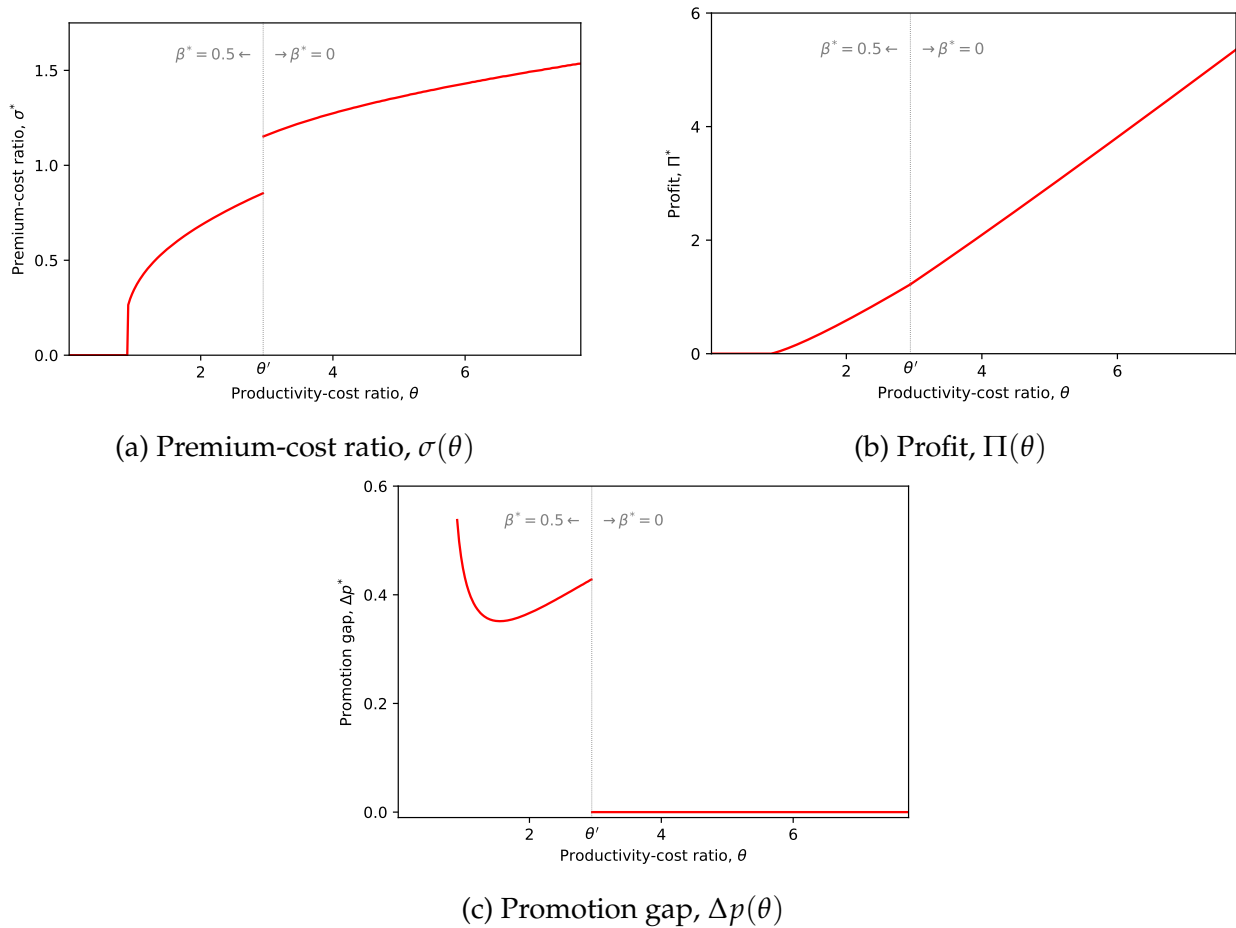


Figure IA.6: Optimal premium-cost ratio, equilibrium profit, and promotion gap, as functions of the productivity-cost ratio, θ , for $\gamma = 9.0$ and $\alpha = 2.0$.

3 Non-Binary Skill

Here we relax the assumption that the acquired skill is a binary variable and extend the model to three levels of skill, $s_i = \{0, 0.5, 1\}$. We show that our results are robust to skill being a non-binary variable.

3.1 Agent's Problem

As in the main model, at Date 1, the agents simultaneously undertake a nonverifiable investment (or effort), $e_i \in [0, 1]$, $i \in b, r$, in firm-specific human capital. Both agents are risk-neutral and have the same skill-acquisition cost function, $c(e_i)$, strictly increasing and convex and such that $c(0) = 0$. Agent i 's probability of acquiring the lower skill level, $s_i = 0.5$, is e_i and the higher skill level, $s_i = 1$, is αe_i , where $\alpha \in [0, 1)$. For example, in project management, project planning, scheduling, and budgeting are considered parts of the foundational skill level ($s_i = 0.5$ in our interpretation), while vision and goal-setting for projects as well as ability to align project objectives with organizational strategy are considered parts of more advanced skill level ($s_i = 1$ in our model). For completeness, we assume that the probability of an agent remaining unskilled is $1 - e_i - \alpha e_i \geq 0$, which implies that the following condition is necessary for an interior solution: $e_i \leq \frac{1}{1+\alpha}$.

Before writing down the agent's maximization problem, we compute probabilities of three disjoint outcomes:

1. Both agents are equally skilled with probability $p_{tie} = e_b e_r + \alpha^2 e_b e_r + (1 - \alpha e_b - e_b)(1 - \alpha e_r - e_r) = 1 + 2e_b e_r(1 + \alpha + \alpha^2) - (e_b + e_r)(1 + \alpha)$;
2. Blue is more skilled with probability $p_{btop} = \alpha e_b(1 - \alpha e_r) + e_b(1 - e_r - \alpha e_r)$;
3. Red is more skilled with probability $p_{rtop} = \alpha e_r(1 - \alpha e_b) + e_r(1 - e_b - \alpha e_b)$.

Now we can write down the maximization problem for Blue and Red under the as-

sumption that the effort cost function is quadratic, $c(e_i) = \frac{ke_i^2}{2}$:

$$\max_{e_b} \sigma \left(p_{btop} + \left(\frac{1}{2} + \beta \right) p_{tie} \right) - \frac{e_b^2}{2}; \quad (15)$$

$$\max_{e_r} \sigma \left(p_{rtop} + \left(\frac{1}{2} - \beta \right) p_{tie} \right) - \frac{e_r^2}{2}. \quad (16)$$

The first order conditions imply the following reaction functions for Blue and Red:

$$e_b = \sigma \left((\alpha + 1) \left(\frac{1}{2} - \beta \right) + 2\beta e_r (\alpha^2 + \alpha + 1) \right) \quad (17)$$

$$e_r = \sigma \left((\alpha + 1) \left(\frac{1}{2} + \beta \right) - 2\beta e_b (\alpha^2 + \alpha + 1) \right) \quad (18)$$

Note that in the absence of subtle bias $\beta = 0$, the reaction functions are flat, just as in the main text:

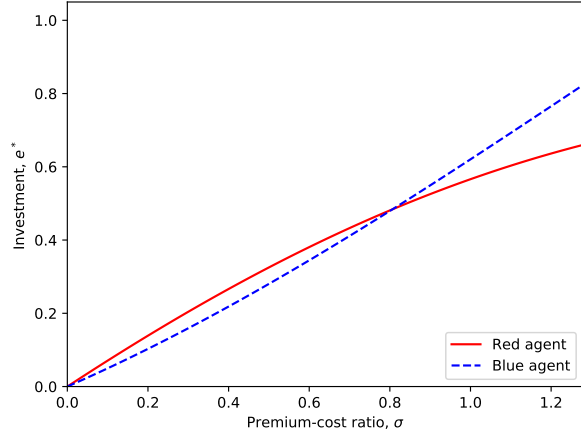
$$e_b^* = e_r^* = \frac{\sigma}{2} (\alpha + 1) \quad (19)$$

If $\alpha = 0$, the optimal investment levels are the same as in the main text (see Eq. (8) and (9) in the main text). For $\alpha \in (0, 1)$, we solve for e_b^* and e_r^* and obtain the solution in the closed form:

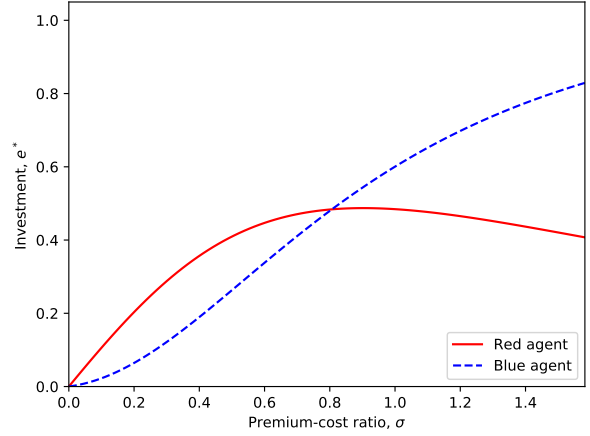
$$e_b^* = \sigma(1 + \alpha) \frac{0.5 - \beta + 2\beta\sigma(0.5 + \beta)(1 + \alpha + \alpha^2)}{1 + 4\sigma^2\beta^2(\alpha^2 + \alpha + 1)^2} \quad (20)$$

$$e_r^* = \sigma(1 + \alpha) \frac{0.5 + \beta - 2\beta\sigma(0.5 - \beta)(1 + \alpha + \alpha^2)}{1 + 4\sigma^2\beta^2(\alpha^2 + \alpha + 1)^2} \quad (21)$$

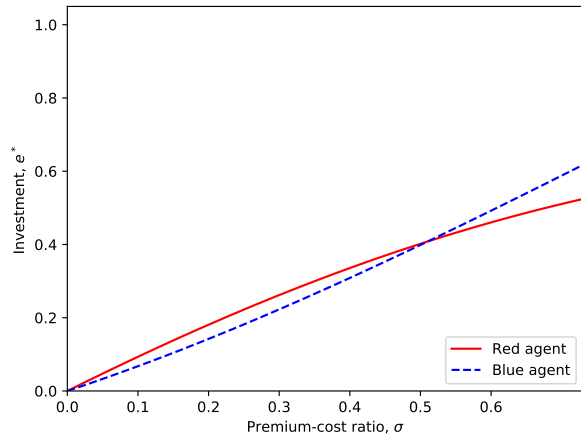
The equilibrium investment levels of Blue and Red for $\alpha \in (0, 1)$ have the same functional form as in the main version of the model. Figure IA.7 shows the equilibrium investment levels of Blue and Red as functions of the premium-cost ratio, σ , for two levels of the subtle bias β and parameter α . As in the main text, in low-stakes situations, the overcompensation effect dominates and in high-stakes situations, the discouragement effect dominates. Therefore, firms that offer low-stakes careers (or less human-capital-intensive) find it profitable to engage in subtle discrimination in order to benefit from the overcompensation effect. In contrast, firms that offer high-stakes careers (and are more human-capital-intensive) prefer not to discriminate.



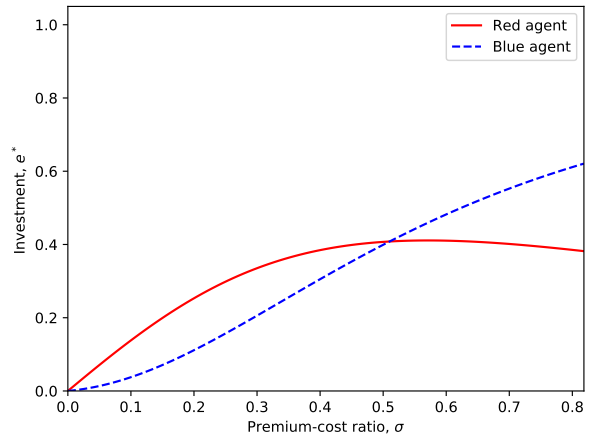
(a) $\beta = 0.1, \alpha = 0.2$



(b) $\beta = 0.4, \alpha = 0.2$



(c) $\beta = 0.1, \alpha = 0.6$



(d) $\beta = 0.4, \alpha = 0.6$

Figure IA.7: Equilibrium investments of blue and red agents, e_b^* and e_r^* , as functions of the premium-cost ratio, σ , for different values of the subtle bias β and the parameter α .

3.2 Principal's Problem

In this subsection, we show that even under a finer skill assessment, involving three levels instead of two, the principal prefers high levels of subtle discrimination in cases of low productivity and low levels of subtle discrimination in cases of high productivity. We assume that upon promoting a moderately skilled agent with $s_i = 0.5$, the principal's profit increases by H , as stated in the main text. Similarly, when promoting a highly skilled agent with $s_i = 1$, the profit increases by μH , where $\mu \geq 1$. Then, the principal's

profit is:

$$\max_{w_l \geq 0, w_l + W \geq 0} l + \mu H \cdot P(s_{b \vee r} = 1) + H \cdot P(s_{b \vee r} = 0.5) - W - 2w_l \quad (22)$$

When $w_l = 0$ and the solution is interior, that is $\sigma \leq \bar{\sigma}(\alpha, \beta)$, we can rewrite the principal's profit maximization problem as:

$$\pi(\alpha, \mu, \theta, k) = \max_{\sigma \in [0, \bar{\sigma}(\alpha, \beta)]} k \left[\mu \theta P(s_{b \vee r} = 1) + \theta P(s_{b \vee r} = 0.5) - \sigma \right], \quad (23)$$

where

$$P(s_{b \vee r} = 1) = \alpha e_b + \alpha e_r - \alpha^2 e_b e_r; \quad (24)$$

$$P(s_{b \vee r} = 0.5) = e_b + e_r - e_b e_r (1 + 2\alpha), \quad (25)$$

subject to

$$e_b(\sigma, \alpha, \beta) = \sigma(1 + \alpha) \frac{0.5 - \beta + 2\beta\sigma(0.5 + \beta)(1 + \alpha + \alpha^2)}{1 + 4\sigma^2\beta^2(1 + \alpha + \alpha^2)^2}; \quad (26)$$

$$e_r(\sigma, \alpha, \beta) = \sigma(1 + \alpha) \frac{0.5 + \beta - 2\beta\sigma(0.5 - \beta)(1 + \alpha + \alpha^2)}{1 + 4\sigma^2\beta^2(1 + \alpha + \alpha^2)^2}. \quad (27)$$

Figure IA.8 shows the optimal principal's profit as a function of the productivity-cost ratio θ , for different levels of subtle bias ($\beta_1 = 0$ and $\beta_2 = 0.5$) and different levels of the parameters α and μ . As we have shown in the main text, for low productivity (low- θ firms), the principal is better off under high levels of subtle discrimination ($\beta = 0.5$), while for high levels of productivity (high- θ firms), she is better off under no discrimination ($\beta = 0$). Note that the region where high subtle discrimination is optimal from the firm's point of view is decreasing in α . In other words, subtle discrimination becomes less profitable for firms as their workers obtain high skill with a greater chance, conditional on an effort level. Figure IA.9 illustrates this insight. It depicts the threshold productivity-cost ratio θ' such that for $\theta < \theta'$, the principal's profit is maximized under the highest level of subtle discrimination, $\beta^{pm} = 0.5$ and for $\theta > \theta'$ it is maximized under no sub-

the discrimination, $\beta^{pm} = 0$ as a function of parameter α and for two different values of parameter μ . Note that θ' decreases in both α and μ .

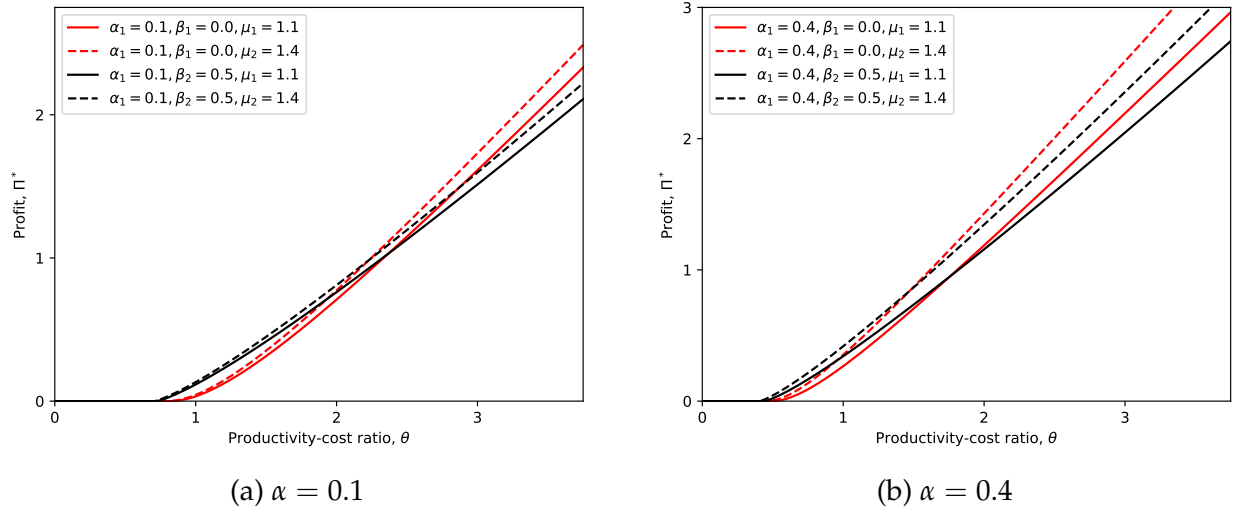


Figure IA.8: Equilibrium profit as a function of the premium-cost ratio, σ , for different values of the subtle bias β and the parameters α and μ .

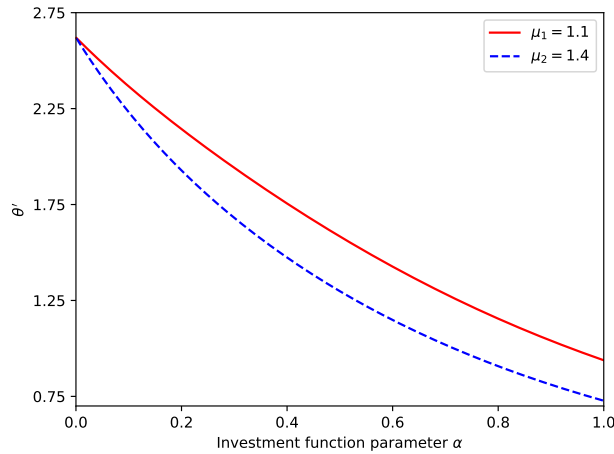


Figure IA.9: Threshold θ' as a function of α for different values of the parameter μ .

3.3 Principal's Choice of Skill Assessment Precision

In this subsection, we investigate under what conditions the principal prefers a more granular skill partition. For example, if an agent is highly skilled, $s_i = 1$, the principal might need to pay an extra cost to distinguish such agent from a skilled agent $s_i = 0.5$.

As before, we assume that promoting a moderately skilled agent increases the principal's profit by H , while promoting a highly skilled agent increases his profit by μH , where $\mu > 1$. Below we compare the principal's profit when agents' skill can take three levels: $s_i = \{0, 0.5, 1\}$ and when the principal chooses to distinguish between all the three levels and when he chooses to separate only skilled ($s_i = 0.5$ and $s_i = 1$) and non-skilled agents ($s_i = 0$).

We solve the principal's problem without any subtle discrimination. In this case, the principal's profit when he recognizes three levels of skill is given by Eq. (23 - 27), where $\beta = 0$.

Now, let us consider a scenario wherein the principal lacks the ability (or intentionally chooses not) to differentiate between individuals possessing high skill levels ($s_i = 1$) and those with medium skill levels ($s_i = 0.5$). Although the principal is aware of the existence of highly productive agents with $s_i = 1$, their screening technology is incapable of distinguishing them from individuals with $s_i = 0.5$. In this case, agent i faces a problem similar to the one in Eq. (4) of the main text:

$$\max_{e_i} \sigma \left[(1 + \alpha)e_i(1 - (1 + \alpha)e_{-i}) + \frac{1}{2} \left(1 - (1 + \alpha)e_i - (1 + \alpha)e_{-i} + 2(1 + \alpha)^2 e_i e_{-i} \right) \right] - \frac{e_i^2}{2}, \quad (28)$$

because when an agent chooses effort level e_i , she becomes skilled (either moderately skilled or highly skilled) with probability $e_i + \alpha e_i$. With no subtle discrimination, the reaction functions are the same as when the principal can distinguish between all three skill levels:

$$e_b^* = e_r^* = \frac{\sigma}{2}(1 + \alpha). \quad (29)$$

The principal's problem in this case is

$$\pi(k, \theta, \mu, \alpha) = \max_{\sigma \in [0, \sigma(\alpha)]} k \left[\theta (e_b + e_r - e_b e_r) \left(\mu \frac{\alpha}{1 + \alpha} + \frac{1}{1 + \alpha} \right) - \sigma \right], \quad (30)$$

subject to

$$e_i = \frac{\sigma}{2}(1 + \alpha).$$

Figure IA.10 shows the principal's optimal profit for cases when she can distinguish between three skill levels (red lines) and for cases when she can only distinguish between two skill levels (black lines). We plot the principal's profit for several values of parameters α and μ .

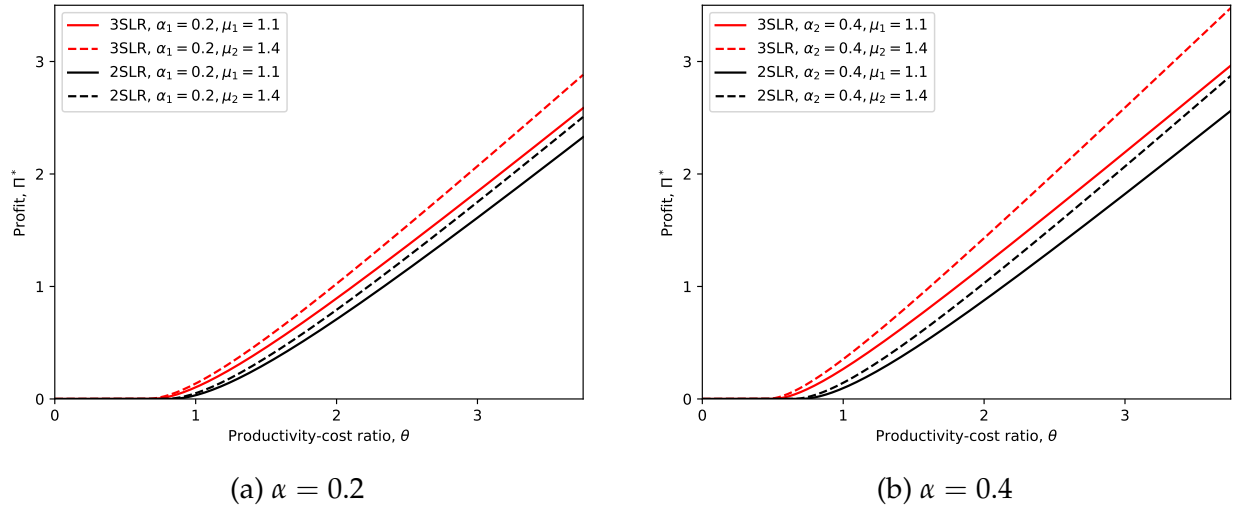


Figure IA.10: Equilibrium profit of the principal as a function of the productivity-cost ratio, θ , for different values of μ and α . Red lines correspond to cases when the principal distinguishes between three skill levels, while black lines correspond to cases when the principal cannot distinguish between medium and highly skilled agents at the promotion stage.

The results in Figure IA.10 show that the principal's profit is higher when he possesses an ability to distinguish between all three skill levels. As a result, the principal would be willing to pay a fixed cost in order to obtain this ability. Importantly, the principal's willingness to pay is increasing in the productivity-cost ratio θ . That is, in low-productivity firms, the principal may decide not to separate between moderately and highly skilled agents if such separation is costly. For example, one might need to conduct additional rounds of interviews to identify highly skilled candidates. If promotion of highly skilled candidates is not profitable enough, the firm may choose to conduct fewer rounds of interviews and to "pool" moderately and highly skilled agents in promotions.

4 Welfare Analysis

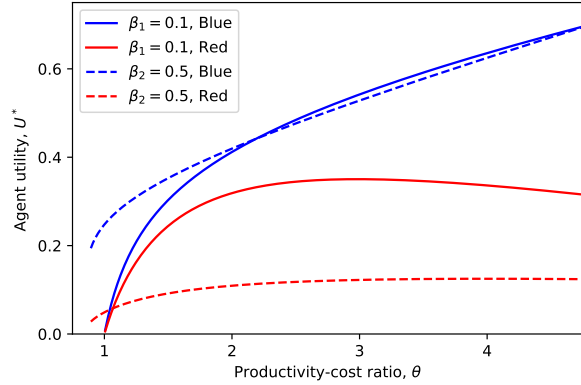
Unlike traditional models of taste-based discrimination, in our model, changing the bias does not mechanically affect utilities. This property allows our model to produce sharper welfare and policy implications. In this section, we address a number of normative questions: What are the welfare implications of subtle discrimination? When is subtle discrimination inefficient?

4.1 When Does Discrimination Harm Workers?

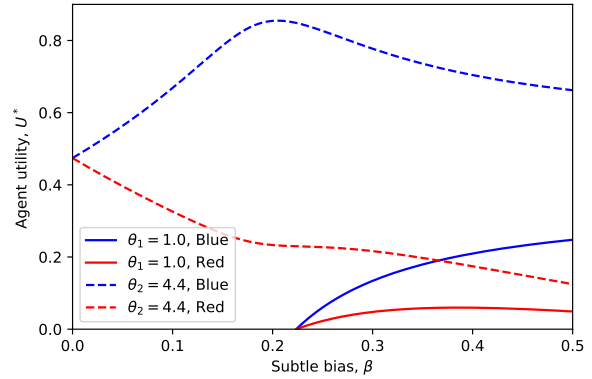
Figures IA.11a and IA.11b show the utilities of blue and red agents as functions of the productivity-cost ratio, θ , and the subtle bias, β , when the contract $\sigma(\theta, \beta)$ is optimally chosen by the principal. Two features are worth highlighting.

First, a stronger bias is not always beneficial to Blue. For high θ , increasing the bias may decrease Blue's utility. How could a bias in favor of blue agents harm these exact agents? A more biased principal offers lower stakes, reducing the benefits of promotion. As the figure shows, this dampening of incentives can offset Blue's gains from a higher bias. Thus, since profits may decrease with the subtle bias, there exist regions in which reducing the bias is a strict Pareto improvement, even in the absence of side transfers.

Second, there exists a region (for small values of θ) where the red agent prefers more discrimination to less. Therefore, for low levels of the productivity-cost ratio, all (the principal and both agents) prefer more discrimination to less. This result highlights that players at different layers of the corporate hierarchy, as well as in different industries, are heterogeneous in their preferences with respect to anti-discriminatory policies. While in positions or industries where productivity gains upon promotion are high everyone may benefit from decreased discrimination, this is not always the case in positions or industries with low productivity gains.



(a) As a function of θ for small and large values of the subtle bias, $\beta_1 = 0.1$ and $\beta_2 = 0.5$.



(b) As a function of β for low and high levels of the productivity-cost ratio, $\theta_1 = 1.0$ and $\theta_2 = 4.4$.

Figure IA.11: Agents' utilities, U^* , under optimal contract $\sigma(\theta, \beta)$.

4.2 Social Surplus

Figure IA.12 presents the level of subtle bias that maximizes the total social surplus, S , as a function of the productivity-cost ratio, θ . The relationship between subtle bias and social surplus is complex. There are three regions. In the first region, low- θ firms benefit from high subtle biases because the overcompensation effect helps to incentivize red agents. As we see from Figure IA.11b, for sufficiently low values of θ both Blue and Red benefit from increasing the bias.² In the second region, Red no longer benefits from the bias and, eventually, the discouragement effect becomes dominant, thus the firm also prefers a lower bias. Thus, for firms with intermediate levels of θ , the social-surplus-maximizing bias is $\beta = 0$. In the third region, Blue's utility is hump-shaped in the subtle bias (see Figure IA.11b), while the firm's profit is relatively flat in β . The optimal bias trades off the gains and losses to the agents. The socially-optimal bias is increasing in the productivity-cost ratio because discouraging Red is efficient when Blue is more likely to win, as it reduces the deadweight costs of effort duplication.

²Note that the bias itself does not directly affect utilities. Thus, our welfare results fundamentally differ from those of models with non-subtle biases. For example, in Prendergast and Topel (1996), an increase in bias directly benefits supervisors.

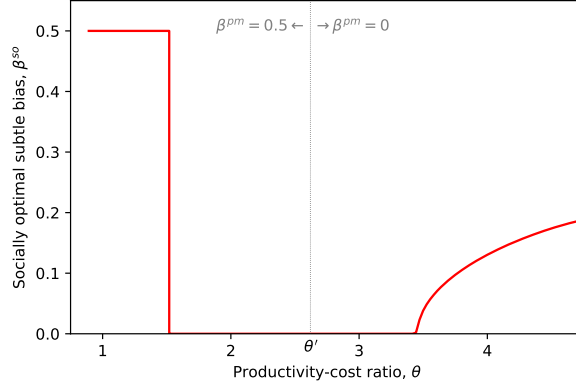


Figure IA.12: Socially optimal level of subtle bias as a function of the productivity-cost ratio, $\beta^{so}(\theta)$.

5 Hard Quotas

The analysis in Section 3.5 of the main text reveals that not all firms would voluntarily take steps towards reducing subtle biases. At the same time, the welfare analysis shows that reducing subtle biases is sometimes socially desirable. Thus, it is instructive to consider possible interventions aimed at reducing or eliminating subtle discrimination.

Setting a (hard) quota is a popular policy tool to tackle a lack of diversity at top positions. Quotas are unlikely to deliver efficiency gains in our model, for two reasons: they constrain the principal’s maximization problem and directly interfere with the agents’ incentives to invest. Nevertheless, quotas may be a policy option for reasons other than efficiency, such as equity and fairness.

To consider quotas at the firm level, we extend the model as follows. At Date 0, the firm has a continuum of vacancies for job 1, with mass 2μ , and for job 2, with mass μ . Each worker in job 1 competes with exactly one worker for promotion. In line with the basic model, all pairs of workers are mixed (one red and one blue). In equilibrium, the probability that an agent of type i is promoted, p_i , is also the proportion of agents of type i found in job 2 at the end of the game. A quota is a target for p_i or, equivalently, a target for the promotion gap, Δp . For convenience we use the latter, thus a quota is fully described by a number $q \in [-1, 1]$.

Without loss of generality, we assume that the quota’s goal is to reduce the promotion

gap, that is, to promote more red agents: $q < \Delta p_0$ (the *pre-quota promotion gap*). Here we adopt the interpretation that the principal designs a firm-wide promotion policy, which is then implemented by a mass μ of supervisors, one for each pair of workers in job 1. We assume that supervisors have incentives aligned with the firm but are subtly biased. Here, unlike in Subsection 3.5 we assume that the firm cannot choose the bias of its supervisors. Because only supervisors observe the skill s_i of their pairs of subordinates, any rule that allows supervisors some discretion can be abused. Thus, the only way to comply with the quota is for the principal to force some supervisors to promote red agents regardless of skill. To do so, the principal offers a proportion δ of supervisors discretion over promotion decisions and forces a proportion $1 - \delta$ of supervisors to promote only red agents.

The principal chooses δ to maximize profit subject to the quota constraint, $\Delta p = q$. The principal has two options: he can reveal the identities of the “constrained” and “unconstrained” supervisors to their subordinates, or he can keep them secret. For brevity, we only consider the full disclosure case.³ The principal’s problem is

$$\Pi(\beta, \theta, q) = \max_{\sigma \in [0, \bar{\sigma}(\beta)], \delta \in [0, 1]} \delta \theta (e_b + e_r - e_b e_r) - \sigma, \quad (31)$$

subject to

$$e_b = \frac{\sigma(0.5 - \beta) + 2\beta\sigma^2(0.5 + \beta)}{1 + 4\beta^2\sigma^2}, \quad (32)$$

$$e_r = \frac{\sigma(0.5 + \beta) - 2\beta\sigma^2(0.5 - \beta)}{1 + 4\beta^2\sigma^2}, \quad (33)$$

$$\Delta p \equiv \delta \{e_b - e_r + [e_b e_r + (1 - e_b)(1 - e_r)] 2\beta\} - (1 - \delta) = q, \quad (34)$$

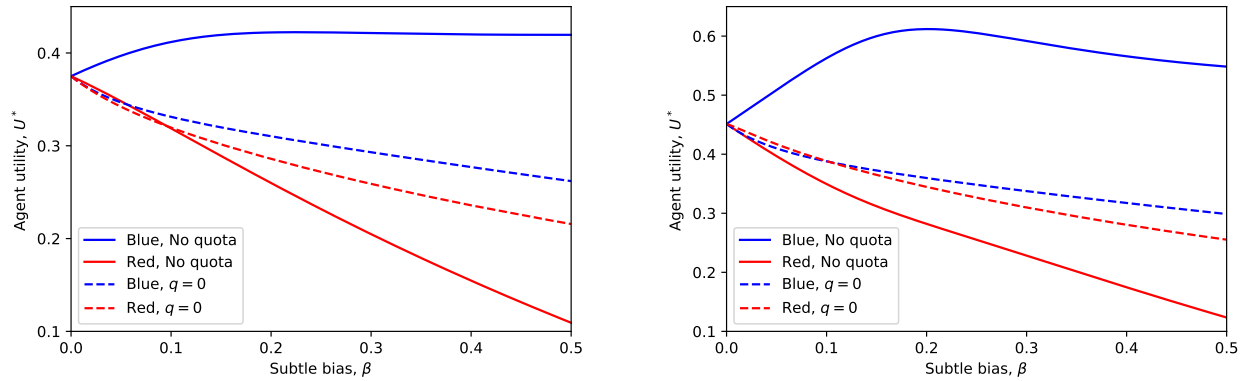
where the last equation is the quota constraint: the promotion gap must be q .

Firm profit is always higher when there is no quota or if the quota is not binding (i.e., $q = \Delta p_0$). This reduction in profit is expected; the quota constrains the principal’s maximization problem. Still, there might be reasons to support quotas on grounds of redistributive equity. The key question is then: when do discriminated agents benefit from quotas?

³The no disclosure case yields similar results.

Figure IA.13 shows the utilities of Blue and Red under a 50% quota (i.e., $q = 0$). As expected, the quota typically reduces Blue's utility and increases Red's utility. However, for low biases, the quota may reduce Red's utility. This counterintuitive result occurs because, under a quota, the firm chooses to offer a smaller promotion bonus. This negative effect dominates when the bias is low because, in such a case, Red's probability of promotion increases by only a small amount after the quota.

We also see that the favored agent is typically better off than the discriminated agent, even when the quota imposes full parity. For Red to do better than Blue under a quota, the bias must be small and productivity must be high.



(a) For $\theta_1 = 2.0$, $\beta = 0$ is socially optimal and $\beta = 0.5$ is profit maximizing

(b) For $\theta_2 = 3.2$, $\beta = 0$ is both socially optimal and profit maximizing

Figure IA.13: Agents' utilities as functions of subtle bias under no quota and under a fully disclosed quota, $\Delta p = q = 0$, for $\theta_1 = 2.0$ and $\theta_2 = 3.2$.

References

PRENDERGAST, C. AND R. H. TOPEL (1996): "Favoritism in organizations," *Journal of Political Economy*, 104, 958–978.