

Gürtler, Marc; Gürtler, Oliver

**Article — Published Version**

## Promotion signaling, discrimination, and positive discrimination policies

The RAND Journal of Economics

**Provided in Cooperation with:**

John Wiley & Sons

*Suggested Citation:* Gürtler, Marc; Gürtler, Oliver (2019) : Promotion signaling, discrimination, and positive discrimination policies, The RAND Journal of Economics, ISSN 1756-2171, Wiley, Hoboken, NJ, Vol. 50, Iss. 4, pp. 1004-1027,  
<https://doi.org/10.1111/1756-2171.12303>

This Version is available at:

<https://hdl.handle.net/10419/230269>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>

# Promotion signaling, discrimination, and positive discrimination policies

Marc Gürtler\*

and

Oliver Gürtler\*\*

*This article studies discrimination in a model in which promotions are used as signals of worker ability. The model can account for statistical and taste-based discrimination. In the short run, a positive discrimination policy is beneficial for workers in the middle of the ability distribution, because these workers are promoted if and only if the policy is in place. Instead, workers of either high or low ability suffer from the policy. In the long run, the policy benefits all targeted workers. The model can explain empirical findings about the effects of a gender quota on the boards of Norwegian companies.*

## 1. Introduction

■ In many countries, certain worker groups are underrepresented in high-level jobs, and their underrepresentation in these typically very well-paid jobs is one of the main reasons for the wage gaps that are observed in practice. As an illustration, consider *The Economist Group*, which, in response to the newly enacted United Kingdom (UK) gender pay gap reporting regulation, reported a median pay gap of 29.5% and declared: “The fact that there are more women in junior roles and more men in senior roles [...] is a major reason for the pay gap.”<sup>1</sup> There is ample evidence that certain groups of workers are discriminated against in the workplace.<sup>2</sup> It is thus conceivable that discrimination is an important factor explaining the underrepresentation of

---

\* Braunschweig Institute of Technology; marc.guertler@tu-bs.de.

\*\* University of Cologne; oliver.guertler@uni-koeln.de.

We thank two anonymous reviewers, David Martimort (the Editor), as well as Jed DeVaro, Christian Ewerhart, Ori Haimanko, Matthias Kräkel, Patrick Legros, Johannes Münster, Hideo Owan, Martin Ruckes, Patrick Schmitz, Anders Stenberg, David Wettstein, and participants at the 41st Annual Conference of the Eastern Economic Association, the 18th Colloquium on Personnel Economics, the Annual 2015 Conference of the Verein für Socialpolitik, the 85th Annual Meetings of the Southern Economic Association, the Econ and Finance Workshop in Karlsruhe, the research seminar ULME in Ulm, the DICE research seminar in Düsseldorf, the DART workshop in Graz, the research seminar at Linnaeus University, the research seminar at Ben-Gurion University of the Negev, and the research seminar at Erasmus University Rotterdam for helpful comments.

The copyright line for this article was changed on February 06, 2020 after original online publication.

<sup>1</sup> [www.economist.com/paygap](http://www.economist.com/paygap) (last retrieved on February 28, 2019).

<sup>2</sup> Fang and Moro (2011) as well as Lang and Lehmann (2012) provide surveys of the economic literature on discrimination.

© 2019 The Authors. *The RAND Journal of Economics* published by Wiley Periodicals, Inc. on behalf of The RAND Corporation.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

certain groups in high-level jobs. The objective of the current article is twofold. First, we aim to develop a promotion-based theory of discrimination. Second, we wish to study the effects of a quota policy, which enforces more equal representation of groups in high-level jobs and which has recently been adopted in several countries (e.g., in Norway and Germany).<sup>3</sup>

Our model builds upon the idea, first illustrated by Waldman (1984), that promotions are signals about worker ability. Waldman (1984) assumes that employers learn their workers' abilities and he shows that they find it optimal to use a threshold promotion rule, promoting all workers whose ability is sufficiently high. As a consequence, when external firms observe a worker's promotion, they infer that the worker's ability exceeds the cutoff value and, thus, they upgrade their assessment of the worker's ability. In turn, they have a greater interest in hiring this worker, with the result that the worker receives more generous wage offers. Recent empirical studies find results that are consistent with predictions derived from the promotion-signaling model.<sup>4</sup>

We consider a promotion-signaling model with two periods and two different groups of workers. In the first period, firms hire workers and assign them to a low-level job. Each firm learns the employed workers' abilities and then decides which workers to promote to a high-level job for the second period. We say that a worker group is discriminated against if the workers from this group face a more stringent promotion threshold than the workers from the other group, implying that workers with given abilities have a smaller chance of being promoted. We show that the model can capture both (endogenous and exogenous) statistical discrimination and taste-based discrimination against a certain group of workers. Endogenous statistical discrimination occurs because of a potential multiplicity of equilibria, meaning that workers who are discriminated against are trapped in an inefficient equilibrium with a high promotion threshold. Exogenous statistical discrimination exists if the ability distributions are different for the two worker groups, whereas taste-based discrimination is caused by employer preferences regarding the workers emanating from the two different groups.

We then introduce a quota policy into the model, requiring workers from an originally disadvantaged group to fill a certain share of the positions in the high-level job. We show that introducing the quota lowers the promotion standard for workers who were originally disadvantaged and are thus favored by the policy, and increases the promotion standard for workers disadvantaged by the policy. We distinguish between the effects of the policy on the employed workers at the time of the policy introduction, and which we term the policy's short-run effects, and the policy's long-run effects on workers who begin their career after the policy has been introduced.

For an employed worker, the policy only affects the worker's second-period payoff. There are two effects. First, the worker is more likely to be promoted and to obtain a wage increase, which obviously benefits him or her. Second, the positive signal of promotion to the high-level job becomes weaker, whereas the negative signal of being reassigned to the low-level job becomes stronger. This results in lower wages in each of the jobs, which clearly hurts workers. Summing up, the policy introduction leaves those workers worse off who have either very high ability, such that they would have been promoted even without the policy, or who have such low ability, that they are denied promotion even when the policy is in place. In contrast, workers in the middle of the ability distribution benefit from the policy because those workers are promoted if and only if it is in place. It is possible for the negative effects on the workers of either low or high ability to outweigh the positive effects on the workers of middle ability, so that a worker's expected payoff may actually decrease. For a worker who begins his career only after the policy has been introduced, the policy also affects the worker's period-1 wage. The reason is that the policy changes a firm's profit from hiring a marginal worker (i.e., the marginal revenue product of labor). When hiring a worker from the group targeted by the policy, a firm finds it easier to fulfill the quota constraint and will change the two groups' promotion standards as a result. This

<sup>3</sup> Fryer and Loury (2005) argue that such positive discrimination policies are highly controversial, and therefore emphasize the importance of economic reasoning in the evaluation of these programs.

<sup>4</sup> See DeVaro and Waldman (2012), Bognanno and Melero (2016), and Cassidy, DeVaro, and Kauhanen (2016).

in turn has an effect on the productivity of workers and, thus, the firm's profit. As a consequence of these effects, we find that a specific quota exists such that, in the long run, the targeted workers benefit from the policy regardless of their exact ability.

The article is organized as follows. In the next section, we present related literature. In Section 3, we turn to the basic model. The effect of positive discrimination policies on workers' wages is analyzed in Section 4. In Section 5, we derive several empirical implications from the model, and we discuss how these implications relate to existing empirical evidence. Section 6 presents the conclusions. The proofs of all lemmas and propositions are in the Appendix.

## 2. Related literature

■ Our article is most closely related to the literature that studies discrimination with respect to promotion decisions. In the model by Athey, Avery, and Zemsky (2000), workers are employed in a low-level job and need mentoring to enhance their skills and become productive in a high-level job. The skills of a considered worker in the low-level job are enhanced more effectively through mentoring if the share of workers in the high-level job who belong to the same group as the worker is high. As a consequence, if the high-level job is dominated by a group of workers, low-level workers of that group receive more effective mentoring, thereby becoming more productive in the high-level job. It is possible that dominance of a group of workers persists, and these workers are more likely to be promoted than the workers of some other group. Bjerk (2008) considers a three-level hierarchy, and new workers begin their career at the lowest level. A worker is promoted if and only if firms have a sufficiently high expectation of the worker's ability. If the average ability in a specific group of workers is low, if the workers belonging to this group cannot send very precise signals of their ability, or if these workers have few chances to signal their ability, even a very able worker from this group needs a long time to upgrade the expectation about his or her ability sufficiently in order to be promoted. Our model differs from the models by Athey, Avery, and Zemsky (2000) and Bjerk (2008) in that we assume that a worker's employer receives more comprehensive information about the worker's ability than external firms, implying that promotions serve as signals about worker ability.<sup>5</sup> In contrast, Athey, Avery, and Zemsky (2000) assume that abilities are observable, whereas Bjerk (2008) assumes that there is uncertainty regarding abilities, but learning is symmetric and all firms receive the same information about a worker's ability.<sup>6</sup>

Another related article is by Milgrom and Oster (1987). They consider a model with two types of workers, whom they call *Visibles* and *Invisibles*. *Visibles* are workers whose skills are publicly observable. On the contrary, the skill of an *Invisible* can only be observed by the current employer and becomes commonly known only if the employer decides to promote the *Invisible* to a high-level job. A consequence of this assumption is that firms have an incentive to "hide" their *Invisibles* in the low-level job, implying that these workers are discriminated against with

<sup>5</sup> The promotion-signaling model was developed by Waldman (1984) and extended by Bernhardt (1995), Zábojník and Bernhardt (2001), Owan (2004), Ghosh and Waldman (2010), DeVaro and Waldman (2012), DeVaro, Ghosh, and Zoghi (2012), Zábojník (2012), Waldman (2013), Gürtler and Gürtler (2015), Cassidy, DeVaro, and Kauhanen (2016), DeVaro and Kauhanen (2016), Shankar (2016), and Waldman (2016).

<sup>6</sup> A few recent articles investigate whether symmetric or asymmetric learning about worker ability is relatively more common in labor markets. Schönberg (2007) develops a formal model that allows for both types of learning. The main finding is that low-ability and high-ability workers are equally likely to switch firms when all firms learn at the same rate about worker ability (i.e., when learning is symmetric), whereas low-ability workers are relatively more likely to move to a different firm when there is asymmetric learning. Schönberg (2007) uses data from the National Longitudinal Survey of Youth 1979 (NLSY79) to test the model, and she finds evidence consistent with asymmetric learning only for college-educated workers. Pinkston (2009) considers a model with multiple periods. A worker is initially hired by some firm and in each period, another firm tries to hire the worker away from the incumbent employer. Each firm receives a private signal about the worker's ability. Importantly, it is assumed that firms engage in a bidding war for the worker's services. This means that firms are able to deduce other firms' signals through the observation of those firms' wage offers, implying that the incumbent employer becomes ever better informed about the worker's true ability. An empirical test of the model is constructed, and the empirical evidence highlights the importance of asymmetric employer learning.

respect to promotion decisions. Obviously, one commonality between our model and the model by Milgrom and Oster (1987) is that promotions serve as signals about worker ability. The important difference is that Milgrom and Oster (1987) assume that the signal is perfect, meaning that a worker's ability becomes common knowledge upon promotion. Instead, in our model, the signal is not perfect. Therefore, lowering the promotion standard may actually hurt workers by making the positive signal of promotion weaker.<sup>7</sup>

Our article is also more generally related to the economic literature on discrimination. This literature can be divided into two different strands. First, there is a body of work on taste-based discrimination that was originated by Becker (1957) and further developed by, for example, Coate and Loury (1993a) and Black (1995). According to this literature, certain workers are discriminated against because firms incur disutility when interacting with these workers. Second, there is a body of literature on statistical discrimination. It is assumed that a worker's ability is not fully known to potential employers. Firms therefore use all the available information to estimate abilities. When there are differences between groups, these differences influence the ability assessments so that it makes sense for the firms to treat two workers differently who belong to two different groups, but are otherwise identical. Differences between groups can either be imposed exogenously (e.g., Phelps, 1972; Aigner and Cain, 1977), or can emerge endogenously (e.g., Coate and Loury, 1993b; Moro and Norman, 2003; Fryer, 2007).

Our article contributes to this literature in two ways. First, we show how discrimination may arise in a promotion-signaling framework. Second, we examine the effects of a quota policy on the payoffs of the affected workers. In contrast to our study, many of the articles mentioned above assume that wages are exogenous. One article that endogenizes wages is by Moro and Norman (2003), who study statistical discrimination in a general equilibrium model with two different types of jobs. They show that positive discrimination policies can lower the wages of workers who are supposed to benefit from the policy. This happens exclusively in low-level jobs, however, and not in high-level jobs. In the model by Moro and Norman (2003), all firms obtain the same information about workers' abilities. Therefore, assignment to a specific job does not serve as a signal about worker ability, meaning that the signaling effects that drive results in the current model are not present in their model.

### 3. The basic model

□ **Description of the model and notation.** We consider a model of a competitive labor market with two periods,  $\tau = 1, 2$ . There are  $N$  identical firms and a continuum of workers of (Lebesgue-) measure  $n$ ; all parties are risk-neutral. There are two different types of jobs, a low-level job 1 and a high-level job 2. Jobs are indexed by  $k = 1, 2$ . Each firm has a continuum of jobs of either type that the firm wishes to fill in period  $\tau$ , and the respective measure for job  $k$  is denoted by  $M_{k\tau}$ . Workers emanate from two different groups,  $l = A, B$ , whose measures are  $n_A$  and  $n_B$ , respectively, ( $n_A + n_B = n$ ). Throughout Section 3, we assume that jobs are “plentiful.” In particular, we assume that  $M_{k\tau} > n$  for all  $k$  and  $\tau$ , so that all workers could, in principle, be hired by the same firm and assigned to the same job.<sup>8</sup>

If worker  $j$  from group  $l$  is hired by firm  $i \in \{1, \dots, N\}$  in period  $\tau$  and assigned to job  $k$ , the worker's output is given by

$$y_{ijl\tau}^k = (1 + s_{ij\tau})(c_k + d_k a_{jl}). \quad (1)$$

The worker's ability is denoted by  $a_{jl} \in [\underline{a}_l, \bar{a}_l]$ , with  $\bar{a}_l > \underline{a}_l \geq 0$ . Within group  $l$ , we define function  $F_l : [\underline{a}_l, \bar{a}_l] \rightarrow [0, 1]$  with  $F_l(x) := \lambda(\{j | a_{jl} \in [\underline{a}_l, x]\})/n_l$ , where  $\lambda$  denotes the Lebesgue measure, that is,  $F_l(x)$  denotes the share of workers from group  $l$  with an ability not greater

<sup>7</sup> DeVaro, Ghosh, and Zoghi (2012) extend the model by Milgrom and Oster (1987), and they derive some implications that they test with data obtained from the personnel records of a large US firm. Dato et al. (2016) conduct a laboratory experiment to test the models by Waldman (1984) and Milgrom and Oster (1987).

<sup>8</sup> We introduce slot constraints in Section 4.

than  $x$ .  $F_l$  is assumed to be continuously differentiable and the corresponding derivative  $f_l$  is assumed to have (full) support  $[a_l, \bar{a}_l]$ . We assume that each  $a_{jl}$  is initially unknown to all firms and all workers (as, e.g., in Holmström, 1982), meaning that even worker  $j$  does not know his or her position in the ability distribution. Consequently,  $F_l(x)$  corresponds to the probability that the ability of an arbitrary worker is not greater than  $x$ . In this sense,  $F_l$  can be interpreted as a cumulative distribution function (cdf) and  $f_l$  as the corresponding probability density function (pdf).

Furthermore,  $c_k \geq 0$  and  $d_k > 0$  are parameters characterizing worker productivity. Following Waldman (1984), we assume that  $d_2 > d_1$  (and  $c_2 < c_1$ ), so that output is more responsive to ability in the high-level job. We define  $a^e := (c_1 - c_2)/(d_2 - d_1)$  as the ability level at which output is equalized across jobs, and we assume  $a^e \in (a_l, \bar{a}_l)$ . Finally,  $s_{ij\tau} \in \{0, S\}$  is an indicator variable capturing firm-specific human capital acquired in the first period of employment. Its realization is equal to zero ( $s_{ij\tau} = 0$ ) if the first period is considered or if the second period is considered, and worker  $j$  has moved to a different firm after the first period. The variable equals  $S > 0$  if the second period is considered and the worker continues to work for the same firm as in the first period.

The workers' mean ability,  $E[a_{jl}] := \int_{a_l}^{\bar{a}_l} a_{jl} f_l(a_{jl}) da_{jl}$ , is assumed to be lower than  $a^e$ , so that each firm finds it optimal to assign workers to the low-level job 1 in  $\tau = 1$ .<sup>9</sup> At the end of the first period, a firm observes the output of each of the own workers and then decides which job the workers are assigned to in  $\tau = 2$ . External firms (which are also referred to as the "labor market") cannot observe individual outputs, but can observe the job assignment of workers. They use this information to update their ability assessment for the workers. We assume that  $a_l$  and  $\bar{a}_l$  are such that firms always decide to assign positive shares of their workers to both jobs at the end of the first period.

In Section 4, we study a positive discrimination policy that is aimed at improving the career prospects of people who are discriminated against. The policy sets a quota  $\alpha \in [0, 1]$  for the workers of group  $A$  in the high-level job, meaning that, in period 2, workers belonging to group  $A$  must fill a share of at least  $\alpha$  of the positions in the high-level job. To simplify the comparison between the models without a quota and with one, we introduce the quota already at this point, but set  $\alpha = 0$  throughout Section 3.

At the beginning of the second period, external firms attempt to hire workers by making wage offers. It is assumed that all wage offers (including the one from the current employer) are made simultaneously. Each worker is hired by the firm making the highest offer. When a worker's current employer matches the highest offer by the external firms, the worker remains with the current employer. We assume  $S$  to be sufficiently high so that, in equilibrium, firms are never successful at hiring workers away from the first-period employer.<sup>10</sup> As in Greenwald (1986) and Waldman (2013), however, there is a (small) share  $\gamma$  of workers who will switch employers after the first period for exogenous reasons that are unrelated to ability and job assignment. The corresponding workers learn of these reasons only after the first period (i.e., they do not know whether they will leave their employer when being hired in  $\tau = 1$ ), and the decision to switch employers is taken only after job assignments have been made. As explained in the two articles mentioned above, and in the proof of Proposition 1, these assumptions eliminate the winner's-curse effect. When a worker switches employers for exogenous reasons, he or she accepts the highest wage offer by the external firms. If multiple external firms make the highest offer to a set of workers who switch employers, each of these firms hires the same share of workers.

Explicit incentive schemes that link pay to performance are not feasible; nor are long-term contracts that bind workers to firms for both periods. There is no discounting.

<sup>9</sup> The firm's profit would be the same if some group of workers of zero measure were assigned to the high-level job 2 in  $\tau = 1$ . Here and in some other instances, we disregard deviations for groups of workers of zero measure; see also Assumption 1 in the next subsection.

<sup>10</sup> See Proposition 1 in the subsequent section.



The time structure can be summarized as follows: at the beginning of period 1, employment relationships are initiated when firms employ workers at the prevailing market wages (all parties are assumed to be price-takers in  $\tau = 1$ ). Workers are assigned to the low-level job and produce period-1 outputs. At the end of the first period, their current employers observe these outputs and decide about which workers to promote to the high-level job. These decisions are followed by all firms making simultaneous wage offers to all workers for the second period.<sup>11</sup> All newly hired workers are then assigned to a job, whereas the firms' existing workers remain in their assigned jobs. Finally, the workers produce their period-2 outputs.

□ **Model solution.** Next, we present the solution to the model. We start with some preliminaries in the next subsection. In particular, we introduce and discuss one additional assumption that simplifies the derivation of the equilibrium. Afterward, we state the equilibrium and we provide intuitive explanations for the observed behavior.

*Preliminaries.* Each firm observes the own workers' first-period outputs and then decides which of the workers to promote, that is, the promotion decisions depend on the observed output levels. According to equation (1), there is a unique linear relationship between first-period output and a worker's ability level  $a_{jl}$ . This implies that a firm can perfectly infer each worker's ability from the output observation. In the following, we will thus write the promotion decisions as a function of the realized ability levels (instead of the output levels). A promotion rule for some worker is then a mapping from the set of ability levels into the set of jobs.

We simplify the analysis by imposing Assumption 1. The assumption is not very restrictive, as we explain in the paragraph following the assumption's statement.

*Assumption 1.* We only consider equilibria with the following properties:

- (a) Workers of a specific ability level are promoted if and only if the second-period profit<sup>12</sup> of promotion is nonnegative. If this condition does not lead to a unique promotion rule, each firm chooses the same promotion rule for all workers from group  $l$ .
- (b) An external firm offers the same period-2 wage to all workers from group  $l$  who were hired by the same firm  $i$  and assigned to job  $k$  at the end of the first period.
- (c) None of the firms hires a nonempty set of workers from group  $l$  of zero measure.

Parts (a) and (b) of Assumption 1 impose symmetry within groups in the sense that all workers from a group who are hired by the same firm  $i$  are subject to the same promotion rule and offered the same wage by an external firm, conditional on their job assignment. These two parts come with very little loss of generality. First, they preclude the possibility of firms treating almost all of the workers with identical observable characteristics in the same way, but displaying different behavior toward a set of workers of zero measure. Because, in the economic literature, it is typically assumed that strategies are identical if they entail a different treatment only for a set of workers of zero measure, this preclusion is not really restrictive.<sup>13</sup> Second, parts (a) and (b) of Assumption 1 do not allow firms to divide groups into subgroups of workers, each with positive measure, and to treat workers in the different subgroups differently. Again, this is not really restrictive. The reason is that we could simply extend the number of groups from two to any larger number without changing any of our results. Thus, if a firm wishes to divide the workers from a group into two subgroups, and to treat workers from the different subgroups differently, we could simply model this by adding another group with the same ability distribution as the

<sup>11</sup> As explained before, the workers' current employers gain market power after the first period, so that price-taking behavior no longer applies in  $\tau = 2$ .

<sup>12</sup> Strictly speaking, this is a profit "density" or a marginal contribution to profit, as the set of workers of a specific ability level has zero measure.

<sup>13</sup> See, for example, Ewerhart (2014).

originally considered group. Finally, part (c) of the assumption means that firms either hire none of the workers from a specific group or a set of workers of positive measure.

Denote by  $n_{li}$  the measure of the set of workers from group  $l$  that firm  $i$  manages to hire at the beginning of  $\tau = 1$ . In the following, we will show that firms find it optimal to promote all workers whose ability exceeds some threshold-value. The corresponding threshold-value can depend on the group  $l$  that workers emanate from, the quota  $\alpha$  (which we set equal to zero in this section), and on the measures of the hired workers  $n_{Ai}$  and  $n_{Bi}$ ; we therefore write it as  $a_{li}^\alpha(n_{Ai}, n_{Bi})$  and sometimes, to shorten the exposition, simply as  $a_{li}^\alpha$ .<sup>14</sup> For ease of notation, we define  $Y_{l1}$  and  $Y_{l2}$  to be the mean outputs of workers from group  $l$  in periods 1 and 2 (excluding firm-specific human capital):

$$\begin{aligned} Y_{l1} &:= c_1 + d_1 E[a_{jl}] \text{ and} \\ Y_{l2}(a_{li}^\alpha(n_{Ai}, n_{Bi})) &:= F_l(a_{li}^\alpha(n_{Ai}, n_{Bi}))(c_1 + d_1 E[a_{jl}|a_{jl} < a_{li}^\alpha(n_{Ai}, n_{Bi})]) \\ &\quad + (1 - F_l(a_{li}^\alpha(n_{Ai}, n_{Bi})))(c_2 + d_2 E[a_{jl}|a_{jl} \geq a_{li}^\alpha(n_{Ai}, n_{Bi})]), \end{aligned}$$

where  $E[\cdot|\cdot]$  denotes the conditional expectation operator.

*The equilibrium.* The next proposition characterizes the equilibrium.

*Proposition 1.* There exists a threshold-value  $S_1 > 0$  for the value of firm-specific human capital  $S$  such that, if  $S > S_1$ , the following holds in equilibrium:

- The first-period wage for the workers from group  $l$  is  $w_{l1} = Y_{l1} + (1 - \gamma)SY_{l2}(a_{li}^0(n_{Ai}, n_{Bi}))$ , where  $Y_{l2}(a_{li}^0(n_{Ai}, n_{Bi}))$  is independent of  $n_{Ai}$  and  $n_{Bi}$  and depends on the lowest promotion threshold  $a_{li}^0$  that some firm  $i$  chooses (see part (c)).
- Any allocation of workers to firms with  $\sum_{i \in I_l} n_{li} = n_l$  can result in  $\tau = 1$ , where  $I_l$  is the set of firms for which the threshold-value  $a_{li}^0$  from part (c) is lowest.
- There exists a threshold-value  $a_{li}^0$  such that firm  $i$  promotes a worker  $j$  from group  $l$  at the end of  $\tau = 1$  if and only if  $a_{jl} \geq a_{li}^0$ .
- The workers' second-period wages on the two job levels are given by

$$\begin{aligned} w_{l2}(1, a_{li}^0) &= c_1 + d_1 E[a_{jl}|a_{jl} < a_{li}^0] \text{ and} \\ w_{l2}(2, a_{li}^0) &= c_2 + d_2 E[a_{jl}|a_{jl} \geq a_{li}^0] > w_{l2}(1, a_{li}^0). \end{aligned}$$

These wages are offered by the first-period employer and at least two external firms. All other external firms offer a wage smaller than  $w_{l2}(k, a_{li}^0)$ .

- A firm's additional second-period profit when promoting worker  $j$  (as a function of  $a_{jl}$  and  $a_{li}^0$ ) corresponds to

$$\Delta\pi_l^p(a_{jl}, a_{li}^0) = (1 - \gamma)[(1 + S)(c_2 - c_1 + (d_2 - d_1)a_{jl}) - (w_{l2}(2, a_{li}^0) - w_{l2}(1, a_{li}^0))].$$

The threshold-value  $a_{li}^0$  is implicitly defined by  $\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$ .

- Any solution to the condition  $\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$  satisfies  $a_{li}^0 > a^e$ .

The intuition for the proposition is as follows. When  $S$  is high enough, workers are always sufficiently more valuable to the incumbent firms than to external firms that the incumbent firms manage to retain the workers unless they switch employers for exogenous reasons. Because of the job moves for exogenous reasons, external firms manage to hire some workers and the equilibrium wages are determined such that they equal the output an external firm would expect a worker to produce. The external firms' profit is therefore zero, implying that, if at least two external firms offer the equilibrium wage, the other firms are indifferent between offering the equilibrium wage and any lower wage.

<sup>14</sup> This means that  $a_{li}^0$  denotes the promotion threshold for group  $l$  when no quota policy is in place.



Because very able workers have a comparative advantage at performing the high-level job, each worker is promoted at the end of the first period if and only if his or her ability is sufficiently high. As a consequence, promotion serves as a (positive) signal of worker ability, and firms offer higher wages to promoted workers rather than to workers who are denied promotion. Because of the wage increase in response to promotion, firms promote inefficiently few workers, that is, the promotion standard  $a_{li}^0$  exceeds the efficient standard of  $a^e$ . This replicates the main finding in Waldman (1984). Note that it is possible that the optimal promotion standard  $a_{li}^0$  is not uniquely defined, that is, the condition  $\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$  may have more than one solution.<sup>15</sup>

In  $\tau = 1$ , labor demand is completely elastic, whereas labor supply is fully inelastic, implying that period-1 wages are determined such that the firms' profit over both periods is zero. Because of firm-specific human capital and asymmetric learning, the firms that manage to hire workers in the first period earn a strictly positive profit in the second period. This means that the marginal revenue product of labor, which equals the period-1 wage and is constant (i.e., it does not depend on  $n_{Ai}$  and  $n_{Bi}$ ), exceeds a worker's expected first-period output, so that firms incur a loss in  $\tau = 1$ .<sup>16</sup>

Finally, when no positive discrimination policy restricting the firms' decisions is in place (i.e., when  $\alpha = 0$ ), all the workers are hired and promoted independently of each other. Therefore, any allocation of workers to firms such that all workers are employed is feasible. We can explain this in a different way that facilitates the later comparison to the model in which a positive discrimination policy is present. Define

$$\begin{aligned} NTO(\alpha, n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN}) &:= \sum_{i=1}^N (n_{Ai} Y_{A1} + n_{Bi} Y_{B1}) \\ &+ (1 - \gamma) S \sum_{i=1}^N (n_{Ai} Y_{A2}(a_{Ai}^\alpha(n_{Ai}, n_{Bi})) + n_{Bi} Y_{B2}(a_{Bi}^\alpha(n_{Ai}, n_{Bi}))) \end{aligned}$$

as the total output that all workers produce in the two periods net of the total period-2 wages. We show in the proof of Proposition 1 that the allocation of workers to firms in  $\tau = 1$  maximizes  $NTO(\alpha, n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN})$ . If  $\alpha = 0$ ,  $NTO$  is linear, and the partial derivatives with respect to  $n_{Ai}$  and  $n_{Bi}$ , respectively, are the same for all firms  $i$  who choose the same promotion standards for the different groups of workers, and they are higher for firms who choose relatively lower standards (in case multiple solutions to the condition  $\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$  exist). Therefore, the value for  $NTO(0, n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN})$  is the same for all possible allocations of workers to firms as long as  $\sum_{i \in I_l} n_{li} = n_l$ .

□ **Discrimination.** We proceed by highlighting the different ways in which workers can be discriminated against in the equilibrium from Proposition 1. Our model is able to capture both endogenous and exogenous statistical discrimination. Furthermore, by slightly modifying the model, we could also address the situation in which firms discriminate against some workers because of distaste for these workers.

We begin with *statistical discrimination* that emerges *endogenously*, as in Coate and Loury (1993b). Consider two workers emanating from the different groups  $A$  and  $B$  and, for simplicity, denote the two workers as workers  $A$  and  $B$ . Suppose that the two groups have an identical ability

<sup>15</sup> A specific example illustrating this possibility is available from the authors upon request. Note that, in the original model by Waldman (1984), the ability distribution was assumed to be uniform, in which case, the optimal promotion standard is always unique.

<sup>16</sup> One feature of the current model (and the one by Waldman, 1984) is that wages can decrease over time, meaning that  $w_{l1}$  can be higher than  $w_{l2}$ . This result is at odds with real-world compensation schemes, which tend to be increasing over time. One way to reconcile the model results with the empirical facts is to assume that workers acquire general human capital (in addition to firm-specific human capital) in the model's first period, which makes them more productive at other firms in the second period, implying that the second-period wage increases.

distribution, that is,  $f_A = f_B$ . Moreover, let there be more than one solution for  $a_{li}^0$  to the condition  $\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$ . It is then conceivable that the two workers face different promotion standards  $a_{Ai}^0$  and  $a_{Bi}^0 (\neq a_{Ai}^0)$ , although the groups' ability distribution is the same.<sup>17</sup> The intuition for this result is the following: if external firms expect a worker's first-period employer to set a rather high promotion standard, they conclude that a promoted worker must be of exceptionally high ability. Therefore, they make very generous wage offers to promoted workers, which in turn makes it optimal for the first-period employer to reassign the worker to the low-level job unless his performance is exceptionally high. In other words, the external firms' expectation of a very high promotion standard may become self-fulfilling in the current model. The same is true for a relatively lower promotion standard, implying that the optimal promotion standard may not be uniquely defined. Assume that worker  $A$  faces a higher promotion standard than worker  $B$ ,  $a_{Ai}^0 > a_{Bi}^0$ . If both workers have the same ability  $\hat{a}$ , and if  $\hat{a} \in [a_{Bi}^0, a_{Ai}^0)$ , worker  $B$  is promoted to the high-level job, whereas worker  $A$  is reassigned to the low-level job. In addition, because the lower promotion standard  $a_{Bi}^0$  is already inefficiently high, worker  $A$  receives a lower total income than worker  $B$ . This means that two equally able workers are treated differently and worker  $A$  is discriminated against. The latter effect requires firms correctly to anticipate the equilibrium that is played at the end of the first period. Here, it is conceivable that firms use identifiable factors such as the race or sex of a worker to coordinate equilibrium, implying discrimination against workers who are "trapped" in the inefficient equilibrium because of these factors.

When the solution for  $a_{li}^0$  to the condition  $\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$  is unique, the model can still capture *exogenous statistical discrimination*. To see this, consider again two different workers  $A$  and  $B$ , but assume that the groups' ability distributions are different,  $f_A \neq f_B$ , in which case, the workers typically differ in their value of  $c_2 - c_1 + d_2 E[a_{jl}|a_{jl} \geq a_{li}^0] - d_1 E[a_{jl}|a_{jl} < a_{li}^0] = w_{l2}(2, a_{li}^0) - w_{l2}(1, a_{li}^0)$ . In turn, it is optimal for firms to set different promotion standards for the two workers ( $a_{Ai}^0 \neq a_{Bi}^0$ ), so that there are first-period output levels at which one of the workers is promoted, whereas the other worker is reassigned to the low-level job.

Finally, the model could also be modified to account for the firms' *taste-based discrimination* against workers. The easiest way to incorporate taste-based discrimination into the model is to assume that firm  $i$  suffers some disutility  $\Delta_{ik} \geq 0$  (per period) when the firm hires a member of some specific group of workers into job  $k$ . For the moment, let us continue to assume that all firms are identical, so that  $\Delta_{ik} := \Delta_k$  for all  $i$ . Furthermore, suppose that  $\Delta_1 \leq \Delta_2$ , meaning that firms care more about what type of workers they hire into the high-level rather than the low-level job. Because the second-period wage upon assignment to job  $k$  is determined by the external firms' wage offers and these offers decrease by  $\Delta_k$  compared to our original model, the second-period wage decreases by  $\Delta_k$  as well. A direct implication is that the wage spread becomes lower (unless  $\Delta_1 = \Delta_2$ ), whereas the optimal promotion rule stays the same. To understand the latter result, notice that firms now suffer a relatively greater disutility from promoting the worker of  $\Delta_2 - \Delta_1$ , while the wage increase upon promotion decreases by  $\Delta_2 - \Delta_1$ ; these effects cancel out, so that the promotion decision is the same as in our main model.<sup>18</sup> Finally, with a similar argument as for the second-period wage, the first-period wage is lowered by  $\Delta_1$ .

The argumentation is a bit more complicated when  $\Delta_{ik}$  is not constant across firms. To simplify the argumentation, we assume that  $\Delta_{i'k} = \Delta_{i''k}$  for all  $i', i'' \in \{2, \dots, N\}$  and  $k \in \{1, 2\}$ , but that  $\Delta_{11} < \Delta_{21}$  and  $\Delta_{12} - \Delta_{11} < \Delta_{22} - \Delta_{21}$ . If there are no slot constraints, as assumed so far, the assumption of price-taking behavior in  $\tau = 1$  no longer seems appropriate, and the least discriminatory firm 1 should possess some market power. The firm then manages to hire all

<sup>17</sup> This would, for example, happen in a symmetric equilibrium in which all firms behave in the same way by choosing identical standards, that is,  $a_{A1}^0 = \dots = a_{AN}^0$  and  $a_{B1}^0 = \dots = a_{BN}^0$ , which differ for the two groups of workers ( $a_{A1}^0 \neq a_{B1}^0$ ).

<sup>18</sup> Of course, a worker would always prefer not to work in  $\tau = 2$  rather than to work and to receive a negative wage. In other words, wages cannot be negative, so that the preceding argument only holds if  $\Delta_k$  does not exceed the wage in job  $k$  that would prevail in the absence of discrimination. Instead, if  $\Delta_k$  is higher than this wage, taste-based discrimination affects the promotion rule.

the workers and pays them a wage depending on the preferences of the other firms. The payoff from hiring the workers is higher for the firm for two reasons. First, there is the obvious effect that the firm suffers lower disutility from hiring the workers. Second, the firm sets a relatively lower promotion standard, that is, in contrast to the model with identical firms, taste-based discrimination can now affect the promotion rule. Interestingly, taste-based discrimination then leads to relatively more workers being promoted. The reason is simple. When promoting a worker, firm 1 suffers (additional) disutility of  $\Delta_{12} - \Delta_{11}$ . The wage increase upon promotion, however, is lowered by  $\Delta_{22} - \Delta_{21} > \Delta_{12} - \Delta_{11}$ . Netting out these effects, firm 1 has a higher incentive to promote the worker, thus leading to a relatively lower promotion standard.<sup>19</sup> Finally, if there exist slot constraints such that the least discriminatory firm is unable to hire all the workers, workers will obviously be hired by multiple firms.

It is to be noted that the three types of discrimination have different implications for workers' career prospects and their wages. Both types of statistical discrimination imply that some workers find it more difficult than others to become promoted, whereas taste-based discrimination does not always have an effect on the promotion rule or may even lead to more of the disliked workers being promoted. To understand the effect of discrimination on wages, consider a worker who faces a relatively higher promotion standard because of (some type of) statistical discrimination, but still manages to become promoted. In the case of endogenous statistical discrimination, it is assumed that the ability distribution is the same for all workers. Thus, by being promoted, the worker signals a high ability and is rewarded with a relatively high wage offer. In contrast, in the case of exogenous statistical discrimination, the ability distributions for different groups of workers are assumed to be different. As a result, a worker can be offered a lower wage even if he or she faces a higher promotion standard. Similarly, our simple model of taste-based discrimination predicts lower wages for workers who are discriminated against. In light of the frequent observation that people who are perceived to be discriminated against also receive lower wages upon promotion,<sup>20</sup> these results allow the conclusion that exogenous statistical discrimination and taste-based discrimination are potentially more important than endogenous statistical discrimination.

#### 4. Positive discrimination policy

■ Having demonstrated how discrimination against a certain group of workers occurs in our model, we assume that a policy maker introduces a quota policy to address discrimination, that is, we now assume  $\alpha > 0$ . If the promotion standard is not unique, even a small policy intervention may induce firms and workers to switch from one equilibrium to another, implying a substantial change in the promotion standard. We avoid such difficulties by focusing on the effects of the policy conditional on a specific equilibrium being played. The easiest way to justify this procedure is to come up with conditions that guarantee that the solution for  $a_{li}^0$  to the condition  $\Delta\pi_i^p(a_{li}^0, a_{li}^0) = 0$  is unique. This is assumed in the following.<sup>21</sup> We assume that the quota needs to be fulfilled both at the time when firms make their promotion decisions and also after workers switch firms at the beginning of  $\tau = 2$ . As we focus on symmetric equilibria in the following, in which firms manage to replace all their departing workers with newly hired ones, this distinction does not matter.

<sup>19</sup> Setting a lower promotion standard could be detrimental if the standard becomes inefficiently low. An inefficiently low standard could always be ruled out by assuming that the firms' preferences do not differ too strongly.

<sup>20</sup> For instance, the study by Bertrand et al. (2019), investigating the effects of a gender quota in Norway, finds that women appointed to the boards of publicly traded Norwegian companies earned about 36% less than their male counterparts before the quota policy was introduced.

<sup>21</sup> As mentioned before, if the ability distribution is uniform, then the solution for  $a_{li}^0$  to the condition  $\Delta\pi_i^p(a_{li}^0, a_{li}^0) = 0$  is always unique. For general distributions, we have a unique solution if  $S$  is sufficiently large.

□ **Workers who are already employed when the policy is introduced.** Our analysis is divided into two parts. We begin by studying the effects of the policy on workers who have already been hired and who are in the first half of their career (i.e., the first period) when the policy is introduced, and which we term the policy's short-run effects. Thereafter (in the next subsection), we investigate the policy's long-run effects on workers who begin their career after the policy was introduced. Studying the short-run effects of the policy makes sense for a number of reasons. First, it may take a considerable amount of time until the ultimate effects of the policy are borne out. Firms may react to the introduction of the policy by changing the composition of their workforce, potentially hiring new workers from specific worker groups, while separating from other workers. These types of decisions are lengthy, meaning that the transition from the old to the new equilibrium may take quite long, so that it is important to study the effects of the policy during the transitional period. Second and related to the first point, policy makers often face a lot of pressure to reverse policies, in particular when these policies fail to achieve their goals in the short run. To understand whether policy makers should succumb to the pressure or whether they should stick with their policies, it is important to know whether and how the policy's short-run effects differ from the effects in the long run. Third, several countries have lately introduced policies aimed at facilitating career advancement for certain groups of people, and researchers have started to investigate the effects of these policies. Given the recent introduction of the policies, it is likely that the considered countries are situated in the transitional period in which firms are still adapting to the policy introduction. Formally studying the short-run effects of the policy is therefore important in order for one to derive some implications that can be contrasted with the existing empirical evidence. Fourth, considering the short-run effects is also beneficial for pedagogical reasons. Understanding the short-run effects of the policy makes it easier to understand the policy's long-run effects, as the short-run effects operate in the long run as well.

In the short run, two variables are unaffected by the policy. First, the workers' period-1 wages are fixed because these wages have already been determined before the policy was put in place. Second, each firm's composition of the workforce is fixed as well, because period-1 hiring decisions have already been made. Regarding the latter, we consider a symmetric situation and impose the assumption that each firm  $i$  originally managed to hire a continuum of workers emanating from group  $l$  of measure  $n_{li} = \frac{n_l}{N}$ . As shown before, such a symmetric equilibrium always exists in the model from Section 3. The following proposition characterizes the short-run effects of the policy.

*Proposition 2.* Suppose that all firms hired a continuum of workers from groups  $A$  and  $B$  of measures  $n_{Ai} = \frac{n_A}{N}$  and  $n_{Bi} = \frac{n_B}{N}$ , respectively, before the positive discrimination policy was introduced. Further, suppose that  $S > S_2$ , where  $S_2 > 0$  is a threshold for the value of firm-specific human capital, and define  $\hat{\alpha} := \inf\{\alpha \in [0, 1] | E[a_{jA} | a_{jA} \geq a_{Ai}^\alpha] \leq a^e\}$ . Then,  $\hat{\alpha} \in (0, 1] \cup \{\infty\}$ , and for  $\alpha \in (0, 1)$  with  $\alpha \leq \hat{\alpha}$ , an equilibrium with the following properties exists:

- (a) Each firm promotes all workers from group  $l \in \{A, B\}$  with ability  $a_{jl} \geq a_l^\alpha$  to the high-level job at the end of  $\tau = 1$  and reassigns all other workers to the low-level job, where  $a_A^\alpha \leq a_A^0$  and  $a_B^\alpha \geq a_B^0$  are the new promotion thresholds.
- (b) At the beginning of  $\tau = 2$ , all firms offer the same wage  $w_{l2}(k, a_l^\alpha)$  to each worker from group  $l$  who was assigned to job  $k$ . The wages for the two jobs are given by

$$\begin{aligned} w_{l2}(1, a_l^\alpha) &= c_1 + d_1 E[a_{jl} | a_{jl} < a_l^\alpha] \text{ and} \\ w_{l2}(2, a_l^\alpha) &= c_2 + d_2 E[a_{jl} | a_{jl} \geq a_l^\alpha] > w_{l2}(1, a_l^\alpha). \end{aligned}$$

- (c) The period-1 employers always manage to retain their workers unless workers switch firms for exogenous reasons. Each firm loses a continuum of workers from group  $l$  who were assigned to job 1 of measure  $\gamma F_l(a_l^\alpha) \frac{n_l}{N}$ . At the same time, each firm manages to hire a continuum of workers from group  $l$  who were assigned to job 1 at other firms of the same

measure, and replaces the departing workers from group  $l$  in job 1 with the newly hired ones from the same group and job. The same holds for job 2, where the corresponding measure is  $\gamma(1 - F_l(a_l^\alpha)) \frac{n_l}{N}$ .

The most important implication of Proposition 2 is that the introduction of the policy leads to changes in the two groups' promotion standards; for group  $A$ , the standard is lowered from  $a_A^0$  to  $a_A^\alpha$ ; for group  $B$ , the standard is increased from  $a_B^0$  to  $a_B^\alpha$ . Intuitively, because of the quota, it is possible that the first-period employer is ordered to promote more workers from group  $A$  or fewer workers from group  $B$  than he or she voluntarily would. It is then optimal for the employer to promote the next-best workers from group  $A$  or to demote the worst workers from group  $B$ , implying the described changes in the promotion standards.<sup>22</sup>

Except for the change in the promotion standards, the results are similar to those of the basic model. Because workers are more valuable to the current employer than to external firms, the employer manages to retain workers unless they decide to switch firms for exogenous reasons. Also, as firms are initially of equal size, they hire the very same measures of workers that they lose to external firms, meaning that the newly hired workers can simply replace the leaving workers.

The condition  $\alpha \leq \hat{\alpha}$  is important for this latter finding. It implies  $E[a_{jA}|a_{jA} \geq a_A^\alpha] \geq a^e$ , which means that the promotion standard for group  $A$  never becomes so low in response to the policy introduction that firms expect promoted workers from this group to be more productive in the low-level rather than the high-level job. Notice that, in many instances, the constraint  $\alpha \leq \hat{\alpha}$  is not really restrictive. For instance, if either  $n_A$  is big relative to  $n_B$  or  $a^e - E[a_{jA}]$  is small, then  $\hat{\alpha}$  is either close to, equal to, or even bigger than 1. In these cases, the condition  $\alpha \leq \hat{\alpha}$  would hold even for rather large quotas and, in particular, for those that are considered in practice and which usually do not exceed 0.5.

In the remainder of this subsection, we will take a closer look at the effects of the policy on workers' equilibrium wages. We focus on the workers from group  $A$ , who are targeted by the policy.

*Proposition 3.* Let  $\alpha \leq \hat{\alpha}$ , where  $\hat{\alpha}$  is defined in Proposition 2.

- (a) We have  $w_{A2}(k, a_A^\alpha) \leq w_{A2}(k, a_A^0)$  with the inequality being strict when  $a_A^\alpha < a_A^0$ , that is, for  $\alpha \in (0, 1)$  with  $n_A(1 - F_A(a_A^0))/(n_A(1 - F_A(a_A^\alpha)) + n_B(1 - F_B(a_B^0))) < \alpha$ .
- (b) Whenever  $a_A^\alpha < a_A^0$ , workers of rather low or high ability are assigned to the same job as in the initial situation without the policy, and they receive a lower second-period wage when the policy is in place. Workers of intermediate ability are promoted because of the policy, and they receive a higher wage.

The policy has two effects on the second-period payoff of a targeted worker from group  $A$ . First, the worker is more likely to be promoted and obtain a wage increase, which obviously benefits the worker. Second, the positive signal of promotion to the high-level job becomes weaker, whereas the negative signal of being reassigned to the low-level job becomes stronger. If the worker is promoted, external firms may question the worker's ability and may believe that the worker was promoted only because of the policy. If the worker is not promoted, external firms may believe that the worker's ability must be extremely low, because he or she was not promoted in spite of the positive discrimination program. Because external firms interpret the job-assignment signal differently when a positive discrimination policy is in place, their wage offers also differ.

<sup>22</sup> Estevan et al. (2014) make similar arguments in the context of college admission. They study the Texas top ten percent policy, which guarantees Texas students who graduated in the top ten percent of their high school class automatic admission to all state-funded colleges. The rule makes it easier for students from high schools with many disadvantaged students to be admitted to college, which is equivalent to a decrease in the educational standard that these students must meet.

In particular, both the wage for a promoted worker and a worker reassigned to the low-level job decrease, as shown in part (a) of Proposition 3.<sup>23</sup> This obviously hurts workers.

The second part of the proposition demonstrates that the introduction of the policy leaves those workers worse off who either have a very high ability so that they would have been promoted even without the policy, or who have such low ability that they are denied promotion even when the policy is in place. In contrast, workers in the middle of the ability distribution benefit from the policy because those workers are promoted if and only if firms are bound to the quota policy. Importantly, these results hold independent of the exact value of  $\alpha$ , meaning that there is no quota such that all targeted workers benefit from the policy's introduction in the short run.

It is possible that the negative effects on workers of either low or high ability outweigh the positive effects on the workers of middle ability, so that a worker's expected payoff may actually decrease. This happens only if the promotion standard is sufficiently below the efficient level, so that the inefficiency due to a standard that is too low is larger than the promotion-signaling distortion, which causes an inefficiently high standard. The conditions in the first sentence of the following proposition ensure that a specific quota  $\alpha$  implementing such an inefficiently low standard exists.

*Proposition 4.* Define  $\hat{a} \in [a_A, a^e]$  implicitly by  $E[a_{jA} | a_{jA} \geq \hat{a}] = a^e$  and suppose  $\hat{a} < \infty$ , where  $\hat{a}$  is defined in Proposition 2, that is,  $a_A^{\hat{a}} = \hat{a}$ .<sup>24</sup>

- (a) The expected second-period wage of a worker from group  $A$  is given by  $Y_{A2}(a_A^\alpha)$  and has a global maximum at a quota  $\alpha^e$  such that  $a_A^{\alpha^e} = a^e$ .
- (b)  $Y_{A2}(a_A^{\hat{a}}) < Y_{A2}(a_A^0)$ .<sup>25</sup>

□ **Workers who begin their career after the policy is introduced.** In this subsection, we investigate the policy's long-run effects. In analogy to Proposition 2, we assume that external firms expect workers from group  $A$  who are promoted by their original employer to be relatively more productive in the high-level than in the low-level job. A sufficient condition for this is that both  $\alpha$  and  $a^e - E[a_{jA}]$  are sufficiently low.

*Proposition 5.* Suppose that the assumption from the preceding paragraph holds and that, for given  $\alpha$ ,  $NTO(\alpha, n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN})$  is concave in  $(n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN})$ .<sup>26</sup> If  $S > S_2$ , an equilibrium with the following properties exists.

- (a) The first-period wage for the workers from group  $l$  is given by

$$w_{l1}^* = Y_{l1} + (1 - \gamma)SY_{l2}(a_l^\alpha(n_A/N, n_B/N)) \\ + (1 - \gamma)S \left( (n_l/N) \frac{\partial Y_{l2}}{\partial n_{li}}(a_l^\alpha(n_A/N, n_B/N)) + (n_\ell/N) \frac{\partial Y_{l2}}{\partial n_{li}}(a_\ell^\alpha(n_A/N, n_B/N)) \right),$$

with  $\ell \in \{A, B\} \setminus \{l\}$ .

- (b) Each of the firms hires a continuum of workers of measure  $n_{li} = n_l/N$  from both groups  $l$  and assigns all these workers to the low-level job in the first period.
- (c) The remainder of the game proceeds in the same way as in Proposition 2.

<sup>23</sup> Notice that the promotion standard  $a_A^\alpha$  is not necessarily a monotone function of  $\alpha$ . However, for each  $\alpha$ , we observe  $a_A^\alpha \leq a_A^0$ , as shown in Proposition 2. This relation between the promotion standards when the policy is or is not in place drives the results of Proposition 3.

<sup>24</sup> The condition  $\hat{a} < \infty$  is, for example, fulfilled if abilities are uniformly distributed and both  $S$  and  $n_B$  are sufficiently large. A corresponding proof is available from the authors upon request.

<sup>25</sup> The inequality  $Y_{A2}(a_A^{\hat{a}}) < Y_{A2}(a_A^0)$  would also hold for each quota  $\alpha$  corresponding to some standard  $a_A^\alpha < \hat{a}$ , and the only difference is that external firms which manage to hire some workers who were promoted by their initial employer would now find it optimal to assign these workers to the low-level job.

<sup>26</sup> This is, for example, the case if the ability distribution is uniform and  $S$  is sufficiently large. A corresponding proof is again available from the authors upon request.



The main difference to Proposition 2 is that the policy, which is designed to affect promotion decisions in  $\tau = 2$ , now “trickles down” to the first period and affects the wages that workers earn in the low-level job in this period. An important result in this context is that a worker’s total expected wage no longer necessarily equals the worker’s expected output, and the reason is that the decisions to hire different workers are now interdependent because of the quota policy. More precisely, a firm’s return from hiring a marginal worker (i.e., the marginal revenue product of labor) comprises the worker’s expected output net of the period-2 wage,  $Y_{l1} + (1 - \gamma)SY_{l2}(a_l^\alpha(n_A/N, n_B/N))$ , and also the worker’s effect on the other hired workers’ outputs,  $(1 - \gamma)S((n_l/N)\frac{\partial Y_{l2}}{\partial n_{li}}(a_l^\alpha(n_A/N, n_B/N)) + (n_l/N)\frac{\partial Y_{l2}}{\partial n_{li}}(a_l^\alpha(n_A/N, n_B/N)))$ . The latter effect is not present in the basic model. It arises because the hiring of workers from one of the groups makes it either easier to satisfy the quota constraint (if the workers emanate from group  $A$ ) or more difficult (if they emanate from group  $B$ ). In either case, it is possible for the promotion thresholds to change, which affects the other workers’ period-2 outputs. For the targeted workers from group  $A$ , the term  $(n_A/N)\frac{\partial Y_{A2}}{\partial n_{Ai}}$  can be stated as  $(n_A/N)\frac{\partial Y_{A2}}{\partial a_{Ai}^\alpha}\frac{\partial a_{Ai}^\alpha}{\partial n_{Ai}}$ , where  $\frac{\partial Y_{A2}}{\partial a_{Ai}^\alpha}$  is negative if and only if  $a_{Ai}^\alpha \geq a^e$  and  $\frac{\partial a_{Ai}^\alpha}{\partial n_{Ai}}$  is always nonnegative. The intuition is that hiring more workers from group  $A$  makes it easier to satisfy the quota constraint, implying an increase in the promotion threshold  $a_{Ai}^\alpha$ . This leads to a reduction in output unless the promotion threshold was inefficiently low to begin with. With similar reasoning, the term  $(n_B/N)\frac{\partial Y_{B2}}{\partial n_{Bi}}$  can be stated as  $(n_B/N)\frac{\partial Y_{B2}}{\partial a_{Bi}^\alpha}\frac{\partial a_{Bi}^\alpha}{\partial n_{Bi}}$ . Here,  $\frac{\partial Y_{B2}}{\partial a_{Bi}^\alpha}$  is always negative (as  $a_{Bi}^\alpha > a^e$ ), and  $\frac{\partial a_{Bi}^\alpha}{\partial n_{Bi}}$  is nonpositive. As an increase in  $n_{Ai}$  makes the quota constraint easier to satisfy, firms are able to promote relatively more workers from group  $B$ , implying a lower, and hence more efficient, promotion standard  $a_{Bi}^\alpha$ .

In the next step, we take a closer look at the effects of the positive discrimination policy on workers’ equilibrium wages. We find that the long-run effects of the policy differ substantially from the corresponding short-run effects. In particular, under some additional conditions, there always exists a quota such that the targeted workers benefit from the policy regardless of their ability level.

**Proposition 6.** Let  $\hat{\alpha} < \infty$  (as in Proposition 4) and  $S > S_3$ , where  $S_3$  is a threshold-value. Then, there exists a quota  $\alpha$  such that any worker  $j$  of group  $A$  receives a higher total wage than in the situation without the policy (where total wage refers to the sum of first- and second-period wages when conditioning on the worker’s ability level). In consequence, workers’ expected total wage is higher as well.

With the same reasoning as in Proposition 3, for a worker of group  $A$ , the second-period wage, conditional on the job level, always decreases when the promotion standard is lowered, due to the quota policy. In contrast, the trickle-down effects described before may lead to a higher first-period wage for the targeted workers. Under the condition  $\hat{\alpha} < \infty$ , the quota can be chosen such that an efficient promotion standard is implemented for group  $A$ . This leads to  $\frac{\partial Y_{A2}}{\partial n_{Ai}} = 0$ , implying that the period-1 wage for the targeted workers is higher than in a situation without a quota. As Proposition 6 indicates, the increase in the period-1 wage can be sufficiently strong such that the effects of the policy on workers’ period-2 wages are outweighed, and workers receive higher total wages regardless of their ability. A sufficient condition for this to happen is that  $S$  is sufficiently high. The simple reason is that period-1 wages are increasing in  $S$ , whereas period-2 wages do not depend on  $S$ . Finally, when all workers from group  $A$  benefit from the policy regardless of their ability level, they also benefit from the policy in expectation, that is, before their ability is known.

□ **Effect on people disadvantaged by the policy.** In this subsection, we briefly address the effects of the positive discrimination policy on workers who are disadvantaged by the policy and who therefore face a relatively higher promotion standard, as explained before. Using our previous results, it is clear that some of the workers who are supposed to be disadvantaged by the

policy may actually be better off, at least in the short run.<sup>27</sup> Extremely able workers who manage to become promoted in spite of the policy receive a higher second-period wage because the positive signal of being promoted grows stronger. The same holds for rather unable workers who are not promoted, even if there is no positive discrimination policy that makes it more difficult for them to move up the corporate ladder. In contrast, workers who are promoted if and only if no positive discrimination policy is in place are typically worse off. Given that the first-period employer promotes inefficiently few workers, a further increase in the promotion standard reduces expected output. In expectation, workers therefore suffer and receive a lower total wage. This may explain why, in practice, members of an initially advantaged group rarely lobby to have positive discrimination policies enacted, although some of them may actually be better off when such policies are introduced.

□ **Slot constraints in the high-level job.** Up to this point, we imposed the assumption  $M_{k\tau} > n$  for all  $k$  and  $\tau$ , so that all workers could in principle be hired by the same firm and assigned to the same job. An implication of this assumption is that each of the firms is able to promote any number of its first-period employees. In practice, however, positions on higher levels of the firms' hierarchies are often scarce. As firms determine their organizational structure themselves, they seem to limit the number of high-level jobs voluntarily, and the specific choice is likely dependent on the firms' production technology. To account for this, we modify the model in the following way. We consider an additional stage at the beginning of the game, where each firm  $i$  chooses the measure of high-level jobs that it wishes to fill in period 2, and which we denote by  $M_{22i}$ . We assume that firm  $i$  can commit to  $M_{22i}$  and that the firm must fill all these jobs in  $\tau = 2$  (unless it does not manage to hire enough workers). We discuss the main findings of this model variant in this subsection, and the proofs of the statements are available upon request.

In the basic model without slot constraints, we demonstrated that any allocation of workers to firms is possible as long as all of the workers are hired by some firm (so that both the markets for the workers from group  $A$  and  $B$  clear). In particular, it is possible that firms "segregate" in the sense that some firms hire only workers from group  $A$ , whereas others hire only workers from group  $B$ , although it is equally possible that each firm hires workers from both of the groups. When we introduce slot constraints into the model, an equilibrium exists in which all of the firms choose efficient promotion standards and which necessarily leads to worker segregation.

The intuition is simple. If a firm hires workers from only one of the groups, the firm can commit to a specific promotion standard by choosing how many workers to promote, that is, by committing to a specific value of  $M_{22i}$ . Hence, the firm is able to implement an efficient promotion standard. On the contrary, if the firm hires workers from both of the groups and if the groups' ability distributions differ, then the firm generally has an incentive to choose a different promotion standard for each group, similar to what we showed in our main model. Thus, the firm would not choose an efficient promotion standard for each of the groups. Stated differently, the firm has one available instrument ( $M_{22i}$ ) to affect the promotion standards. If the firm hires workers from only one of the groups, the firm needs to determine just one promotion standard, and the instrument suffices to achieve an efficient solution. This is different from the case in which the firm hires workers from both groups and in which the one instrument is insufficient to determine both promotion standards efficiently.

In the described equilibrium, the same (efficient) promotion standard is chosen for each of the groups, implying that discrimination does not occur. Accordingly, it would not seem

<sup>27</sup> This finding is related to the observation by Fang and Norman (2006) that people who (are thought to) suffer from government-mandated discriminatory policies are at times economically more successful than people who benefit from these policies. Fang and Norman (2006) present the example of ethnic Chinese in Malaysia, who were discriminated against by governmental implementation of the "New Economic Policy," but who did surprisingly well economically afterward. Fang and Norman (2006) explain this observation using a model with a public and a private sector. If a group of people is discriminated against in the public sector, this group faces higher incentives to succeed in the private sector, implying that the group may perform better economically.

necessary to introduce a positive discrimination policy. That said, the segregation of workers is not a desirable outcome, either. Interestingly, one can show that some quota rule exists, which leads to an equilibrium in which each firm hires workers from both groups and efficient promotion standards are chosen nonetheless.

The quota rule differs somewhat from the one we considered before in that it exactly determines the shares of workers from the two groups in the high-level job rather than just a lower bound for the share of workers from group *A*. The rule could never be satisfied if firms hired workers from only one of the groups, so that segregation is prevented. Furthermore, if the quota is determined appropriately, firms cannot discriminate between worker groups, given that they choose an efficient standard for at least one of the groups. Thus, an equilibrium exists in which firms choose efficient standards for both groups. In a certain sense, the quota rule represents the second instrument that firms require to choose both promotion standards efficiently (even though it is determined by a policy maker rather than the firms themselves).

## 5. Empirical implications of the model

■ Several empirical implications can be derived from the model. We studied the short-run and long-run effects of the positive discrimination policy and we begin by discussing the implications that are common to both these studies. In general, the policy could have an effect on the firms' hiring decisions, their promotion decisions, and their wage offers to the workers. When we studied the policy's short-run effects, the initial hiring decisions had already been made, so we start by discussing the effects of the policy on promotions and period-2 wage offers. We demonstrated that the introduction of the policy leads to changes in the promotion standards for the different groups of workers (Proposition 2). A direct implication is that the policy's introduction changes the employment of people from an initially disadvantaged group and an initially advantaged group in high-level jobs. Although more people from the former group are promoted to the high-level job, there are fewer people from the latter group in the very same job. In line with this prediction, Holzer and Neumark (2000), surveying the literature on positive discrimination programs, conclude that these "programs redistribute employment [...] from white males to minorities and women".<sup>28</sup>

An important implication of the model is that the introduction of the policy lowers the wages that workers who are advantaged by the policy earn later in their careers, and increases the wages for workers who are disadvantaged by the policy when controlling for job level (Proposition 3). The starting wage of workers who began their working careers before the policy was implemented is unaffected. This, in turn, implies that the wage increase upon promotion is changed by the policy. For workers who are advantaged by the policy, the post-promotion wage decreases, while the starting wage is fixed so that the wage increase upon promotion also decreases. For workers who are disadvantaged by the policy, the effect is the opposite and the wage increase upon promotion increases. Bertrand et al. (2019) study the effects of a law that was passed in Norway in 2003, which mandates a 40% representation of each gender on the boards of publicly traded companies. They show that election to the board of a company entails a substantial financial reward for the elected worker. In line with the predictions of our model, this reward fell for women (from about 10% of annual earnings to 8%), whereas it increased substantially for men (from about 5% to 9%) after the board quota was introduced.

Another implication is that the expected wage a worker earns later in his or her career potentially decreases when a positive discrimination policy favoring the worker is introduced (Proposition 4). As a consequence, it is possible that such policies do not diminish average "wage gaps" between different groups of workers. Support for this implication is provided by several studies. Smith and Welch (1984) study wage gaps between white and black men in the United

<sup>28</sup> Recent studies such as Kurtulus (2012) underscore this observation. She finds that the share of minorities and women who are employed in high-paying skilled jobs in the United States grew more between 1973 and 2003 in firms that were subject to affirmative action regulations than in firms that were not.

States between 1967 and 1980. They show that the wage gap became narrower in the first years of their analysis (until 1972), but that it stabilized after effective monitoring systems for affirmative action were established. The study by Leonard (1996) focuses on the US labor market in the 1980s and comes to a similar conclusion: “Between 1980 and 1990 the average wage gap decreased dramatically for women but increased for every other group.” The most striking evidence is provided by Burger and Jafta (2012), who study the impact of affirmative action programs on employment and wages in South Africa. They show that the average wage gap between white and black workers increased substantially after the affirmative action policies were put in place. As argued before, these at first sight perverse effects are consistent with the results of our model.

When we investigated the policy’s long-term effects, we showed that the policy trickles down to the first period and also affects the wages that workers earn at the beginning of their careers (Proposition 5). Bertrand et al. (2019) do not observe such trickle-down effects in their study of Norwegian companies, but, as argued before, it is conceivable that they studied a transitional period in which firms are still adapting to the policy introduction. It would thus be interesting to repeat the analysis of Bertrand et al. (2019) in a few years to study the policy’s long-term effects.

A related finding of our model is that, after the introduction of the policy, a worker’s total expected wage could be different from the worker’s expected output (Proposition 5). Hiring decisions are interdependent with the consequence that the zero-profit condition does not necessarily hold on the individual level. Thus, an empirical implication of the model is that positive discrimination programs lead to systematic differences between wages and productivity for workers emanating from different groups. We are not aware of any study investigating such differences, but believe that this would be an interesting avenue for future research.

Finally, our model predicts that all external job moves are lateral in the sense that workers stay at the same job level when moving to an external firm. This prediction is independent of whether or not the quota policy is in place. When the policy is in place, however, the result hinges on the assumption of symmetric firms, which implies that firms manage to hire the very same measures of workers that they lose to external firms (conditional on group affiliation and job assignment), so that the newly hired workers can simply replace the leaving workers. If we extended the model (e.g., Proposition 2) to allow for asymmetric firms, the results may be different. If firms differ in size, for instance, a smaller firm may hire more workers than it loses to the external firms. In this case, it is conceivable that, due to the quota constraint, the firm cannot assign all workers to the high-level job who were promoted by their initial employers. Two reactions by the firm are possible. To meet the quota constraint, it may either promote some workers from group *A* to the high-level job who were assigned to the low-level job by their initial employer, or it may demote some workers from group *B* to the low-level job who were initially assigned to the high-level job. The model could thus be modified such that the policy would lead to some promotions across firms for workers favored by the policy and demotions across firms for the workers disadvantaged by the policy. Again, to the best of our knowledge, no study investigating this prediction exists.

## 6. Conclusion

■ In this article, we consider a model in which promotions are used as signals of worker ability, and we examine the impact of a positive discrimination policy. The policy lowers the promotion standard for the workers who are discriminated against. In the short run, this is beneficial for the workers in the middle of the ability distribution, because these workers are promoted if and only if the policy is in place. In contrast, workers of either high or low ability generally suffer from the policy. This is because the policy does not change the promotion decision for these workers, but rather weakens the positive signal of being promoted and strengthens the negative signal of not being promoted. These findings imply that policies aimed at “leveling the playing field” are not always as beneficial as they may appear. If workers succeed despite many obstacles, the labor

market learns a great deal about their characteristics, so it can reward the workers generously on this basis.<sup>29</sup>

In the long run, the policy can benefit all targeted workers. The policy increases the wage that these workers earn early in their career, and this wage increase more than outweighs the possible future wage reduction. We derive several empirical implications from our model and we discuss empirical studies that are supportive of some of our predictions. We also offer new implications that researchers could take to the data.

## Appendix

The Appendix contains the proofs of the propositions.

### *Proof of Proposition 1.*

- (i) Before we turn to the proposition's parts (a) to (f), we establish that there exists a threshold-value  $S_1 > 0$  for the value of firm-specific human capital  $S$  such that, in equilibrium, external firms are never successful at hiring workers away from the first-period employer if  $S > S_1$ . If the first-period employer managed to hire a worker again in the second period, the worker's second-period output would be at least  $(1 + S)(c_1 + d_1 \bar{a}_l)$ . If one of the external firms managed to hire a worker, the worker's second-period output would be at most  $c_2 + d_2 \bar{a}_l (= \max\{c_1 + d_1 \bar{a}_l, c_2 + d_2 \bar{a}_l\})$ . For both  $l$ , suppose that

$$(1 + S)(c_1 + d_1 \bar{a}_l) > c_2 + d_2 \bar{a}_l \Leftrightarrow S > \frac{c_2 + d_2 \bar{a}_l}{c_1 + d_1 \bar{a}_l} - 1 =: S_{1l},$$

and define  $S_1 := \max\{S_{1A}, S_{1B}\}$ . Now, denote the highest offer that worker  $j$  receives from one of the external firms by  $w_{j2}$ . According to Assumption 1(b), the corresponding external firm makes the same wage offer to all other workers from the same group, hired by the same firm, and assigned to the same job. By Assumption 1(c), this set of workers has a positive measure. Hence, we must have  $w_{j2} \leq c_2 + d_2 \bar{a}_l (< (1 + S)(c_1 + d_1 \bar{a}_l))$ , as otherwise, the external firm would suffer a loss (because a share of the workers switches firms for exogenous reasons). It follows that the first-period employer always finds it optimal to match  $w_{j2}$ .

- (ii) Promotion rules and period-2 wage offers are interdependent and thus need to be determined simultaneously. Accordingly, we now turn to the proposition's parts (c) to (f), imposing the assumption  $S > S_1$ . Denote by  $A_{li}^1 \subseteq [\underline{a}_l, \bar{a}_l]$  the set of ability levels for which firm  $i$  decides to reassign workers from group  $l$  to the low-level job in  $\tau = 2$ , and by  $A_{li}^2 \subseteq [\underline{a}_l, \bar{a}_l]$  the set of ability levels for which workers from group  $l$  are promoted to the high-level job. By Assumption 1(a), for each ability level, the firm either promotes all workers of that ability or it reassigns all these workers to the low-level job. Hence,  $\{A_{li}^1, A_{li}^2\}$  is a partition of the set  $[\underline{a}_l, \bar{a}_l]$ . We denote the external firms' belief regarding  $\{A_{li}^1, A_{li}^2\}$  by  $\{\tilde{A}_{li}^1, \tilde{A}_{li}^2\}$ .

Consider an arbitrary worker  $j$  of group  $l$  who is assigned to job  $k$ . According to part (b) of Assumption 1, the highest wage offer of an external firm for this worker is also made to all other workers of group  $l$  who are hired by the same firm and assigned to the same job. Also, because this set of workers has positive measure, the workers' first-period employer has an incentive to exactly match this highest wage offer from the external firms. Thus, it remains to determine this offer, which we denote by  $w_{il2}(k)$ . The mean ability of workers who are actually switching firms is equal to the overall mean ability of workers (conditional on job assignment at the end of  $\tau = 1$ ). This is because workers are never successfully hired away, but a small fraction  $\gamma$  of workers leaves the first-period employer for reasons that are unrelated to ability and job assignment. It follows that workers' expected outputs for the external firms are  $\max\{c_1 + d_1 E[a_{jl}|a_{jl} \in \tilde{A}_{li}^1], c_2 + d_2 E[a_{jl}|a_{jl} \in \tilde{A}_{li}^2]\} =: Z_{il}$ . As external firms manage to hire only those workers who switch firms for exogenous reasons, they compete against each other in a regular Bertrand-fashion. Thus, at least two of the external firms offer  $Z_{il}$  and the external firms that do not offer  $Z_{il}$  offer a lower wage, so that we obtain  $w_{il2}(k) = Z_{il}$ .

Recall that firm  $i$ 's promotion rule is the same for all workers from group  $l$ . Consider arbitrary  $\tilde{a}_1 \in A_{li}^1$  and  $\tilde{a}_2 \in A_{li}^2$ , meaning that the firm does not want to promote a worker from group  $l$  when the worker has ability  $\tilde{a}_1$ , but that it wishes to promote a worker with ability  $\tilde{a}_2$ . Against the background of Assumption 1(a), this implies that the second-period profit of promotion of the worker with ability  $\tilde{a}_2$  ( $\tilde{a}_1$ ) is nonnegative (negative):

$$\begin{aligned} (1 + S)(c_2 + d_2 \tilde{a}_2) - w_{il2}(2) &\geq (1 + S)(c_1 + d_1 \tilde{a}_2) - w_{il2}(1) \\ (1 + S)(c_2 + d_2 \tilde{a}_1) - w_{il2}(2) &< (1 + S)(c_1 + d_1 \tilde{a}_1) - w_{il2}(1). \end{aligned} \quad (A1)$$

<sup>29</sup> A related argument is advanced in Krishnamurthy and Edlin (2014). In a study of college admission rules, they find that stereotypes against a disadvantaged group of students can only be eliminated if these students face higher admission standards. Formally, they assume that the ability distributions of the disadvantaged and the advantaged students satisfy the monotone likelihood ratio property. This assumption implies that the expected ability of admitted students can only be equalized across groups when the students of the disadvantaged groups have to meet a higher standard in order to be admitted to college.

Rearranging the two conditions leads to

$$\begin{aligned} (1+S)(c_2 - c_1 + (d_2 - d_1)\tilde{a}_2) &\geq w_{il2}(2) - w_{il2}(1) \\ &> (1+S)(c_2 - c_1 + (d_2 - d_1)\tilde{a}_1), \end{aligned} \quad (\text{A2})$$

which immediately implies  $\tilde{a}_2 > \tilde{a}_1$ . Hence, because we assumed that  $\underline{a}_l$  and  $\bar{a}_l$  are such that firms always decide to assign positive shares of their workers to both jobs at the end of  $\tau = 1$ , the sets  $A_{li}^1$  and  $A_{li}^2$  must be of the form  $A_{li}^1 = [\underline{a}_l, a_{li}^0]$  and  $A_{li}^2 = [a_{li}^0, \bar{a}_l]$ , with  $a_{li}^0 \in [\underline{a}_l, \bar{a}_l]$ .

It remains to determine the equilibrium value of  $a_{li}^0$ . Having hired a continuum of group- $l$  workers of measure  $n_{li}$ , firm  $i$  chooses the promotion standard  $a_{li}^0$  to maximize second-period profit

$$\begin{aligned} &(1-\gamma)n_{li}F_l(a_{li}^0)[(1+S)(c_1 + d_1E[a_{jl}|a_{jl} < a_{li}^0]) - w_{il2}(1)] \\ &+ (1-\gamma)n_{li}(1 - F_l(a_{li}^0))[(1+S)(c_2 + d_2E[a_{jl}|a_{jl} \geq a_{li}^0]) - w_{il2}(2)]. \end{aligned}$$

The first-order condition to the maximization problem is given by

$$\begin{aligned} &(1-\gamma)n_{li}f_l(a_{li}^0)[(1+S)(c_1 + d_1E[a_{jl}|a_{jl} < a_{li}^0]) - w_{il2}(1)] \\ &+ (1-\gamma)n_{li}F_l(a_{li}^0)(1+S)d_1 \frac{\partial E[a_{jl}|a_{jl} < a_{li}^0]}{\partial a_{li}^0} \\ &- (1-\gamma)n_{li}f_l(a_{li}^0)[(1+S)(c_2 + d_2E[a_{jl}|a_{jl} \geq a_{li}^0]) - w_{il2}(2)] \\ &+ (1-\gamma)n_{li}(1 - F_l(a_{li}^0))(1+S)d_2 \frac{\partial E[a_{jl}|a_{jl} \geq a_{li}^0]}{\partial a_{li}^0} = 0. \end{aligned}$$

We further have

$$\frac{\partial E[a_{jl}|a_{jl} < a_{li}^0]}{\partial a_{li}^0} = \frac{f_l(a_{li}^0)}{F_l(a_{li}^0)}(a_{li}^0 - E[a_{jl}|a_{jl} < a_{li}^0])$$

and

$$\frac{\partial E[a_{jl}|a_{jl} \geq a_{li}^0]}{\partial a_{li}^0} = \frac{f_l(a_{li}^0)}{1 - F_l(a_{li}^0)}(E[a_{jl}|a_{jl} \geq a_{li}^0] - a_{li}^0),$$

implying that the preceding first-order condition can be stated as:<sup>30</sup>

$$\begin{aligned} &(1-\gamma)n_{li}f_l(a_{li}^0)[(1+S)(c_1 + d_1a_{li}^0) - w_{il2}(1)] \\ &- (1-\gamma)n_{li}f_l(a_{li}^0)[(1+S)(c_2 + d_2a_{li}^0) - w_{il2}(2)] = 0 \\ \Leftrightarrow &(1-\gamma)((1+S)(c_2 - c_1 + (d_2 - d_1)a_{li}^0) - (w_{il2}(2) - w_{il2}(1))) = 0. \end{aligned}$$

In equilibrium, the external firms correctly anticipate the firm's promotion rule, that is,  $\tilde{A}_{li}^k = A_{li}^k$  for  $k = 1, 2$ . Thus, the workers' second-period wages (conditional on job assignment) can be stated as

$$\begin{aligned} w_{il2}(1) &= w_{l2}(1, a_{li}^0) = \max \{c_1 + d_1E[a_{jl}|a_{jl} < a_{li}^0], c_2 + d_2E[a_{jl}|a_{jl} < a_{li}^0]\} \text{ and} \\ w_{il2}(2) &= w_{l2}(2, a_{li}^0) = \max \{c_1 + d_1E[a_{jl}|a_{jl} \geq a_{li}^0], c_2 + d_2E[a_{jl}|a_{jl} \geq a_{li}^0]\} > w_{il2}(1). \end{aligned}$$

As  $E[a_{jl}|a_{jl} < a_{li}^0] \leq E[a_{jl}] < a^e$ , it follows that  $w_{l2}(1, a_{li}^0) = c_1 + d_1E[a_{jl}|a_{jl} < a_{li}^0]$ . Furthermore, because  $a_{li}^0 \in A_{li}^2$  and  $w_{l2}(2, a_{li}^0) - w_{l2}(1, a_{li}^0) > 0$ , inequality (A2) leads to

$$(1+S)(c_2 - c_1 + (d_2 - d_1)a_{li}^0) > 0 \Rightarrow c_2 + d_2a_{li}^0 > c_1 + d_1a_{li}^0 \quad \text{for all } a_{li}^0 \geq a_{li}^0.$$

This immediately implies  $w_{l2}(2, a_{li}^0) = c_2 + d_2E[a_{jl}|a_{jl} \geq a_{li}^0]$ .

The function  $\Delta\pi_l^p(a_{li}^0)$  obviously characterizes the additional second-period profit from promoting a worker with ability  $a_{li}^0$ . From the above first-order condition to firm  $i$ 's maximization problem, it follows that the promotion threshold fulfills  $\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$ . Because

$$\begin{aligned} \Delta\pi_l^p(a^e, a_{li}^0) &= (1-\gamma)[(1+S)(c_2 - c_1 + (d_2 - d_1)a^e) - (w_{l2}(2, a_{li}^0) - w_{l2}(1, a_{li}^0))] \\ &= (1-\gamma)[w_{l2}(1, a_{li}^0) - w_{l2}(2, a_{li}^0)] < 0, \end{aligned}$$

$\Delta\pi_l^p(a_{li}^0, a_{li}^0) = 0$ , and  $\Delta\pi_l^p(a_{li}^0, a_{li}^0)$  is strictly increasing in  $a_{li}^0$ , it follows that  $a_{li}^0 > a^e$ .

<sup>30</sup> We obtain  $(1-\gamma)n_{li}f_l'(a_{li}^0)[(1+S)(c_1 + d_1a_{li}^0) - w_{il2}(1) - (1+S)(c_2 + d_2a_{li}^0) + w_{il2}(2)] + (1-\gamma)n_{li}f_l(a_{li}^0)(1+S)(d_1 - d_2) < 0$  as the second-order condition. The second-order condition is always fulfilled, as the first term is equal to zero as a result of the first-order condition, and the second term is negative.



- (iii) Next, we turn to part (b) and determine the allocation of workers to firms in  $\tau = 1$ . It is well established that the set of firms' equilibrium labor demands (at given period-1 wages  $w_{l1}$ ) equals the set of labor demands that maximize the firms' aggregate profits (see, e.g., Proposition 5.E.1 in Mas-Colell, Whinston, and Green, 1995). Also, as the labor supply is fully inelastic, we have  $\sum_{i=1}^N n_{li} = n_l$  for  $l \in \{A, B\}$ , implying that  $\sum_{i=1}^N (n_{Ai} w_{A1} + n_{Bi} w_{B1})$  is constant, that is, the firms' total wage costs are fixed. It follows that equilibrium labor demands maximize the firms' total expected output (net of period-2 wage costs), which is given by

$$\sum_{i=1}^N (n_{Ai} Y_{A1} + n_{Bi} Y_{B1}) + (1 - \gamma) S \sum_{i=1}^N (n_{Ai} Y_{A2}(a_{Ai}^0(n_{Ai}, n_{Bi})) + n_{Bi} Y_{B2}(a_{Bi}^0(n_{Ai}, n_{Bi}))).$$

Notice that  $Y_{l2}(a_{li}^0(n_{Ai}, n_{Bi}))$  can be stated as

$$\begin{aligned} Y_{l2}(a_{li}^0(n_{Ai}, n_{Bi})) &= c_2 + (c_1 - c_2)F_l(a_{li}^0) + d_1 E(a_{jl}) + (d_2 - d_1)(1 - F_l(a_{li}^0)) \cdot E[a_{jl}|a_{jl} \geq a_{li}^0] \\ &= c_2 + (c_1 - c_2)F_l(a_{li}^0) + d_1 E(a_{jl}) + (d_2 - d_1) \int_{a_{li}^0}^{\bar{a}_{jl}} a_{jl} f_l(a_{jl}) da_{jl}, \end{aligned}$$

which is independent of  $n_{li}$ . Hence, the preceding expression can be restated as

$$(Y_{A1} + (1 - \gamma)SY_{A2}(a_{Ai}^0(n_{Ai}, n_{Bi}))) \sum_{i=1}^N n_{Ai} + (Y_{B1} + (1 - \gamma)SY_{B2}(a_{Bi}^0(n_{Ai}, n_{Bi}))) \sum_{i=1}^N n_{Bi}. \quad (A3)$$

Differentiating  $Y_{l2}(a_{li}^0(n_{Ai}, n_{Bi}))$  with respect to  $a_{li}^0$  leads to

$$\frac{\partial Y_{l2}}{\partial a_{li}^0} = (c_1 - c_2 - (d_2 - d_1) \cdot a_{li}^0) \cdot f_l(a_{li}^0),$$

and thus,  $\frac{\partial Y_{l2}}{\partial a_{li}^0} < 0$  for all  $a_{li}^0 > a^e$ . It follows that any allocation of workers to firms with  $\sum_{i \in I_l} n_{li} = n_l$  for  $l \in \{A, B\}$ , where  $I_l$  is the set of firms for which the threshold-value  $a_{li}^0$  is lowest, maximizes (A3) and thus, can result in  $\tau = 1$ .

- (iv) Finally, we determine period-1 wages, that is, we prove the proposition's part (a). In  $\tau = 1$ , labor demand is completely elastic, whereas labor supply is fully inelastic. It follows that period-1 wages are determined such that firms' expected profit over both periods is zero, and we obtain  $w_{l1} = Y_{l1} + (1 - \gamma)SY_{l2}(a_{li}^0(n_{Ai}, n_{Bi}))$ .  $\square$

*Proof of Proposition 2.* First, we note that, analogous to our argumentation in the basic model, if  $S$  is sufficiently high, no equilibrium exists in which one of the external firms ever manages to hire a worker away from the first-period employer (unless the worker switches firms for exogenous reasons). We can always choose  $S_2$  such that this is fulfilled. In the following, let  $\alpha \in (0, 1)$  with  $\alpha \leq \hat{\alpha}$  (note that  $\hat{\alpha} > 0$  because  $E[a_{jl}|a_{jl} \geq a_l^0] > a^e$ ). We use backward induction to determine optimal behavior.

Proof of (c). We begin by showing that the firms have no incentive to deviate from the job assignment of newly hired external workers in  $\tau = 2$  as described in part (c) if they follow the proposed behavior up to this decision. Each firm originally hired a continuum of workers from group  $A$  or group  $B$  of measure  $n_A/N$  or  $n_B/N$ , respectively. It follows that each firm loses a continuum of workers from group  $l$  that were assigned to job 1 of measure  $\gamma F_l(a_l^0) \frac{n_l}{N}$  and to job 2 of measure  $\gamma(1 - F_l(a_l^0)) \frac{n_l}{N}$  in period 2. Because all firms were initially of equal size and they all choose the same promotion thresholds, each firm simultaneously manages to hire the very same measures of workers from the same group and assigned to the same job. In equilibrium, all firms can correctly anticipate the chosen promotion standards  $a_l^0$ . Because  $E[a_{jl}|a_{jl} \geq a_l^0] \geq a^e$  for both  $l$  (for group  $A$ , this is true because  $\alpha \leq \hat{\alpha}$ , for group  $B$ , the condition holds as  $a_B^0 > a^e$ ), each firm would find it optimal to assign all newly hired workers who were originally assigned to job 2 to the high-level job as well. As the quota constraint was originally met and the newly hired workers simply replace the leaving workers, such assignment is always possible without violating the quota constraint. Similarly, because  $E[a_{jl}|a_{jl} < a_l^0] < E[a_{jl}] < a^e$  for both  $l$ , each firm always assigns all newly hired workers that were originally assigned to job 1 to the low-level job as well.

Proof of (b). The next decision relates to the firms' period-2 wage offers. First, it is easy to see that a firm's wage offers to the own employees are always optimal, as matching the external firms' offers represents the cheapest way to retain the workers. Consider now the wage offers to the external firms' workers. Again, suppose that all firms follow the behavior described in the proposition up to this decision. If firms continued to stick to the proposed equilibrium behavior, each firm would earn zero profit from the newly hired workers. The reason is that each worker who was originally assigned to job  $k$  will be assigned to the very same job when switching firms (as explained in the proof of (c)), implying that the period-2 wage offers specified in the proposition are equal to the workers' expected period-2 outputs. When lowering some or all of the wage offers, a firm would never have a chance at hiring the workers whose offers were lowered, meaning that the measure of the hired workers would decline. In addition, the firm might have to reassign some of the remaining workers to the job at which they are least productive, in case the quota constraint were violated (e.g., if only the wage offers to workers from group  $A$  were lowered). Hence, such deviation from the equilibrium can never be profitable. Increasing the wage offers to some or all of the workers would not be profitable either, as the firm would then pay those workers a wage

above their expected output. This could only be profitable when this would relax the quota constraint, making it easier to assign other workers to the job at which they are most productive. As the newly hired workers are already assigned to the jobs at which they are most productive and the existing workers cannot be reassigned, this is not the case.

Proof of (a). Finally, we turn to the firms' promotion decisions at the end of  $\tau = 1$ . Let  $i$  be an arbitrary firm and denote the firm's optimal promotion standards by  $a_{Ai}^\alpha$  and  $a_{Bi}^\alpha$ , respectively. The quota constraint is given by

$$q_{Ai}(a_{Ai}^\alpha, a_{Bi}^\alpha) := \frac{n_{Ai}(1 - F_A(a_{Ai}^\alpha))}{n_{Ai}(1 - F_A(a_{Ai}^\alpha)) + n_{Bi}(1 - F_B(a_{Bi}^\alpha))} \geq \alpha$$

$$\Leftrightarrow (1 - \alpha)n_{Ai}(1 - F_A(a_{Ai}^\alpha)) - \alpha n_{Bi}(1 - F_B(a_{Bi}^\alpha)) \geq 0.$$

In analogy to the proof of Proposition 1, one can show that the optimal promotion standards are determined by the conditions

$$\Delta\pi_A^p(a_{Ai}^\alpha, a_{Ai}^\alpha) + \mu_i(1 - \alpha) = 0 \text{ and } \Delta\pi_B^p(a_{Bi}^\alpha, a_{Bi}^\alpha) - \mu_i\alpha = 0,$$

where  $\mu_i \geq 0$  is the Lagrange multiplier to  $(1 - \alpha)n_{Ai}(1 - F_A(a_{Ai}^\alpha)) - \alpha n_{Bi}(1 - F_B(a_{Bi}^\alpha)) \geq 0$ . It follows that  $\Delta\pi_A^p(a_{Ai}^\alpha, a_{Ai}^\alpha) \leq 0$  and  $\Delta\pi_B^p(a_{Bi}^\alpha, a_{Bi}^\alpha) \geq 0$  (with both inequalities being strict if  $\mu_i > 0$  and  $\alpha < 1$ ).

In addition, if  $S_2(< S)$  is sufficiently high, we have both

$$\Delta\pi_i^p(a^e, a^e) = (1 - \gamma)[(1 + S)(c_2 - c_1 + (d_2 - d_1)a^e) - (w_{I2}(2, a^e) - w_{I2}(1, a^e))]$$

$$= (1 - \gamma)[w_{I2}(1, a^e) - w_{I2}(2, a^e)] < 0 \text{ and}$$

$$\Delta\pi_i^p(\bar{a}_i, \bar{a}_i) = (1 - \gamma)[(1 + S)(c_2 - c_1 + (d_2 - d_1)\bar{a}_i) - (w_{I2}(2, \bar{a}_i) - w_{I2}(1, \bar{a}_i))]$$

$$= (1 - \gamma)[(1 + S)(c_2 - c_1 + (d_2 - d_1)\bar{a}_i) - (c_2 + d_2\bar{a}_i - c_1 - d_1E[a_{jI}])] > 0.$$

Because  $a^e < a_i^0 < \bar{a}_i$ ,  $\Delta\pi_i^p(a, a)$  is continuous in  $a$ , and we assumed (throughout Section 4) that the solution for  $a_i^0$  to the condition  $\Delta\pi_i^p(a_i^0, a_i^0) = 0$  is unique, it immediately follows that  $\Delta\pi_i^p(a, a) < 0$  for all  $a < a_i^0$  and  $\Delta\pi_i^p(a, a) > 0$  for all  $a > a_i^0$ . Under consideration of  $\Delta\pi_A^p(a_{Ai}^\alpha, a_{Ai}^\alpha) \leq 0$  and  $\Delta\pi_B^p(a_{Bi}^\alpha, a_{Bi}^\alpha) \geq 0$ , it follows  $a_{Ai}^\alpha \leq a_A^0$  and  $a_{Bi}^\alpha \geq a_B^0$ .

Finally, as all firms hired a continuum of workers from groups  $A$  and  $B$  of measures  $\frac{n_A}{N}$  and  $\frac{n_B}{N}$ , there is always an equilibrium in which  $\mu_i$  is the same for all firms and they all choose the same promotion standards.  $\square$

*Proof of Proposition 3.*

- (a) The statement immediately follows from the wage formula in part (b) of Proposition 2 because  $c_1 + d_1E[a_{jI}|a_{jI} < a]$  as well as  $c_2 + d_2E[a_{jI}|a_{jI} \geq a]$  are increasing functions in  $a$ , and  $a_A^\alpha \leq a_A^0$  with the inequality being strict when  $\alpha$  is so high that firms are forced to change the promotion standards to fulfill the quota constraint, that is,  $\alpha > n_A(1 - F_A(a_A^0))/(n_A(1 - F_A(a_A^0)) + n_B(1 - F_B(a_B^0)))$ . Furthermore, we have to consider that  $\alpha < 1$  and  $\alpha \leq \hat{\alpha}$  (so that Proposition 2 can be applied).
- (b) If  $a_{jA} \in [\underline{a}_A, a_A^\alpha]$ , the corresponding worker is neither promoted if the promotion standard is  $a_A^0$  nor if the promotion standard is  $a_A^\alpha$ . Consequently, the period-2 wage difference amounts to

$$w_{A2}(1, a_A^\alpha) - w_{A2}(1, a_A^0) = d_1(E[a_{jA}|a_{jA} < a_A^\alpha] - E[a_{jA}|a_{jA} < a_A^0]) < 0.$$

The argumentation is analogous if  $a_{jA} \in [a_A^0, \bar{a}_A]$ .

Finally, if  $a_{jA} \in [a_A^\alpha, a_A^0]$ , the worker is not promoted if the promotion standard is  $a_A^0$ , but is promoted if the promotion standard is  $a_A^\alpha$ . The resulting wage difference corresponds to

$$w_{A2}(2, a_A^\alpha) - w_{A2}(1, a_A^0)$$

$$= c_2 - c_1 + d_1(E[a_{jA}|a_{jA} \geq a_A^\alpha] - E[a_{jA}|a_{jA} < a_A^0]) + (d_2 - d_1)E[a_{jA}|a_{jA} \geq a_A^\alpha]$$

$$> c_2 - c_1 + (d_2 - d_1)E[a_{jA}|a_{jA} \geq a_A^\alpha] \geq 0.$$

The last inequality results because  $E[a_{jA}|a_{jA} \geq a_A^\alpha] \geq a^e$ .  $\square$

*Proof of Proposition 4.* Because  $a_A^\alpha$  is a continuous function of  $\alpha$  and  $\hat{\alpha} < \infty$ , it immediately follows that  $a_A^{\hat{\alpha}} = \hat{a}$ .

- (a) It is obvious that the workers' expected period-2 wage is given by  $Y_{A2}$ . From the derivative of  $Y_{A2}$  with respect to the promotion standard, which has been determined in part (iii) of the proof of Proposition 1, we obtain the following necessary condition of an extremum:

$$\left. \frac{\partial Y_{A2}}{\partial a} \right|_{a=a^*} = 0 \Leftrightarrow a^* = \frac{c_1 - c_2}{d_2 - d_1} = a^e.$$

As  $\partial^2 Y_{A2} / \partial a^2|_{a=a^e} = -(d_2 - d_1)f_A(a^e) + (c_1 - c_2 - (d_2 - d_1)a^e)(\partial f_A / \partial a)|_{a=a^e} < 0$  (because  $(c_1 - c_2) - (d_2 - d_1)a^e = 0$ ), a local maximum of  $Y_{A2}$  at  $a^e$  immediately results. Because  $a^e$  is the only zero of  $\partial Y_{A2} / \partial a$  in  $(\hat{a}, \bar{a}_A)$  and  $Y_{A2}$  is continuous over  $[\hat{a}, \bar{a}_A]$ ,  $a^e$  is the global maximum of  $Y_{A2}$ .

(b) Notice that

$$c_2 + d_2 E[a_{jA}|a_{jA} \geq \hat{a}] = c_2 + d_2 a^e = c_1 + d_1 a^e = c_1 + d_1 E[a_{jA}|a_{jA} \geq \hat{a}].$$

By the law of total expectation, it then follows that

$$\begin{aligned} Y_{A2}(\hat{a}) &= F_A(\hat{a})(c_1 + d_1 E[a_{jA}|a_{jA} < \hat{a}]) + (1 - F_A(\hat{a}))(c_2 + d_2 E[a_{jA}|a_{jA} \geq \hat{a}]) \\ &= F_A(\hat{a})(c_1 + d_1 E[a_{jA}|a_{jA} < \hat{a}]) + (1 - F_A(\hat{a}))(c_1 + d_1 E[a_{jA}|a_{jA} \geq \hat{a}]) \\ &= c_1 + d_1 E[a_{jA}]. \end{aligned}$$

As  $\partial Y_{A2}/\partial a|_a < 0$  for all  $a \in (a^e, \bar{a}_A) \supset (a_A^0, \bar{a}_A)$ , we have

$$Y_{A2}(\hat{a}) = c_1 + d_1 E[a_{jA}] = Y_{A2}(\bar{a}_A) < Y_{A2}(a_A^0).$$

□

*Proof of Proposition 5.* First, we note that, if  $S > S_2$ , no equilibrium exists in which one of the external firms ever manages to hire a worker away from the first-period employer (unless the worker switches firms for exogenous reasons). We use backward induction to show that none of the parties wishes to deviate from the proposed behavior.

Proof of (c). The proof is completely analogous to the proofs of the corresponding parts of Proposition 2.

Proof of (a) and (b). With a similar argumentation as in the proof of Proposition 1, the allocation of workers to firms maximizes  $NTO(\alpha, n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN})$  s.t.  $\sum_{i=1}^N n_{li} = n_l$  for  $l \in \{A, B\}$ . As  $NTO$  is a concave function of  $(n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN})$ , the maximizer of  $NTO$  s.t.  $\sum_{i=1}^N n_{li} = n_l$  for  $l \in \{A, B\}$  is given by the first-order conditions to the problem of maximizing the (also concave) Lagrangian

$$L := NTO(\alpha, n_{A1}, \dots, n_{AN}, n_{B1}, \dots, n_{BN}) + \lambda_1 \left( \sum_{i=1}^N n_{Ai} - n_A \right) + \lambda_2 \left( \sum_{i=1}^N n_{Bi} - n_B \right).$$

These conditions can be stated as

$$\begin{aligned} 0 &= \frac{\partial L}{\partial n_{Ai}} = Y_{A1} + (1 - \gamma)SY_{A2}(a_{Ai}^\alpha(n_{Ai}, n_{Bi})) \\ &\quad + (1 - \gamma)S \left( n_{Ai} \frac{\partial Y_{A2}(a_{Ai}^\alpha(n_{Ai}, n_{Bi}))}{\partial n_{Ai}} + n_{Bi} \frac{\partial Y_{B2}(a_{Bi}^\alpha(n_{Ai}, n_{Bi}))}{\partial n_{Ai}} \right) + \lambda_1, \\ 0 &= \frac{\partial L}{\partial n_{Bi}} = Y_{B1} + (1 - \gamma)SY_{B2}(a_{Bi}^\alpha(n_{Ai}, n_{Bi})) \\ &\quad + (1 - \gamma)S \left( n_{Ai} \frac{\partial Y_{A2}(a_{Ai}^\alpha(n_{Ai}, n_{Bi}))}{\partial n_{Bi}} + n_{Bi} \frac{\partial Y_{B2}(a_{Bi}^\alpha(n_{Ai}, n_{Bi}))}{\partial n_{Bi}} \right) + \lambda_2, \end{aligned}$$

for  $i \in \{1, \dots, N\}$ . It is straightforward to check that a symmetric solution with  $n_{l1} = \dots = n_{lN} = \frac{n_l}{N}$  for  $l \in \{A, B\}$  exists to the set of first-order conditions. It follows that an equilibrium exists in which all firms hire a continuum of workers from each group  $l$  of measure  $\frac{n_l}{N}$ .

Finally, a worker's first-period wage is equal to the marginal revenue product of labor. A firm's revenue (net of period-2 wages) is given by

$$n_{Ai} Y_{A1} + n_{Bi} Y_{B1} + (1 - \gamma)S(n_{Ai} Y_{A2}(a_{Ai}^\alpha(n_{Ai}, n_{Bi})) + n_{Bi} Y_{B2}(a_{Bi}^\alpha(n_{Ai}, n_{Bi}))).$$

Thus, the marginal revenue product of labor for group  $l$  is

$$Y_{l1} + (1 - \gamma)S \left( Y_{l2}(a_{li}^\alpha(n_{Ai}, n_{Bi})) + n_{li} \frac{\partial Y_{l2}(a_{li}^\alpha(n_{Ai}, n_{Bi}))}{\partial n_{li}} + n_{\ell i} \frac{Y_{\ell 2}}{\partial a_{\ell i}^\alpha}(a_{\ell i}^\alpha(n_{Ai}, n_{Bi})) \right),$$

with  $\ell \in \{A, B\} \setminus \{l\}$ . Inserting  $\frac{n_A}{N}$  and  $\frac{n_B}{N}$  for  $n_{Ai}$  and  $n_{Bi}$ , we obtain the period-1 wage stated in the proposition. □

*Proof of Proposition 6.* Recall that period-2 wages are independent of  $S$  regardless of whether or not the positive discrimination policy is in place. We will prove the proposition by showing that there exists a quota  $\alpha$  such that the difference in period-1 wages for workers from group  $A$  when comparing the situations with and without the policy is increasing without bound in  $S$ .

The corresponding difference in period-1 wages can be stated as

$$w_{A1}^* - w_{A1} = (1 - \gamma)S \left( Y_{A2}(a_A^\alpha(n_A/N, n_B/N)) + \frac{1}{N} \left( n_A \frac{\partial Y_{A2}}{\partial a_A^\alpha} \frac{\partial a_A^\alpha}{\partial n_{Ai}} + n_B \frac{\partial Y_{B2}}{\partial a_B^\alpha} \frac{\partial a_B^\alpha}{\partial n_{Ai}} \right) - Y_{A2}(a_A^0) \right).$$

Hence, we need to show that there exists  $\alpha$  such that

$$Y_{A2}(a_A^\alpha(n_A/N, n_B/N)) + \frac{1}{N} \left( n_A \frac{\partial Y_{A2}}{\partial a_A^\alpha} \frac{\partial a_A^\alpha}{\partial n_{Ai}} + n_B \frac{\partial Y_{B2}}{\partial a_B^\alpha} \frac{\partial a_B^\alpha}{\partial n_{Ai}} \right) - Y_{A2}(a_A^0) > 0.$$

According to Proposition 4(a), there exists  $\alpha^e \in (0, 1)$  such that  $a_A^{\alpha^e} = a^e$  and  $\frac{\partial Y_{A2}}{\partial a_A^\alpha} \big|_{\alpha=\alpha^e} = 0$ . Suppose that  $\alpha = \alpha^e$ . Then, we obtain

$$\begin{aligned} & Y_{A2}(a_A^\alpha(n_A/N, n_B/N)) + \frac{1}{N} \left( n_A \frac{\partial Y_{A2}}{\partial a_A^\alpha} \frac{\partial a_A^\alpha}{\partial n_{Ai}} + n_B \frac{\partial Y_{B2}}{\partial a_B^\alpha} \frac{\partial a_B^\alpha}{\partial n_{Ai}} \right) - Y_{A2}(a_A^0) \\ &= Y_{A2}(a^e) + \frac{n_B}{N} \frac{\partial Y_{B2}}{\partial a_B^\alpha} \frac{\partial a_B^\alpha}{\partial n_{Ai}} - Y_{A2}(a_A^0). \end{aligned}$$

This expression is strictly positive, because  $Y_{A2}(a^e) > Y_{A2}(a_A^0)$  and  $\frac{\partial Y_{B2}}{\partial a_B^\alpha} \frac{\partial a_B^\alpha}{\partial n_{Ai}} \geq 0$ . The latter inequality holds, as  $\frac{\partial a_B^\alpha}{\partial n_{Ai}} \leq 0$  if  $S$  is sufficiently high (a formal proof of this condition is available from the authors upon request).  $\square$

## References

- AIGNER, D.J. AND CAIN, G.G. "Statistical Theories of Discrimination in Labor Markets." *Industrial and Labor Relations Review*, Vol. 30 (1977), pp. 175–187.
- ATHEY, S., AVERY, C., AND ZEMSKY, P. "Mentoring and Diversity." *American Economic Review*, Vol. 90 (2000), pp. 765–786.
- BECKER, G.S. *The Economics of Discrimination*. Chicago: University of Chicago Press, 1957.
- BERNHARDT, D. "Strategic Promotion and Compensation." *Review of Economic Studies*, Vol. 62 (1995), pp. 315–339.
- BERTRAND, M., BLACK, S.E., JENSEN, S., AND LLERAS-MUNEY, A. "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway." *Review of Economic Studies*, Vol. 86 (2019), pp. 191–239.
- BJERK, D. "Glass Ceilings or Sticky Floors? Statistical Discrimination in a Dynamic Model of Hiring and Promotion." *The Economic Journal*, Vol. 118 (2008), pp. 961–982.
- BLACK, D.A. "Discrimination in an Equilibrium Search Model." *Journal of Labor Economics*, Vol. 13 (1995), pp. 309–333.
- BOGNANNO, M.L. AND MELERO, E. "Promotion Signals, Experience, and Education." *Journal of Economics and Management Strategy*, Vol. 25 (2016), pp. 111–132.
- BURGER, R. AND JAFTA, R. "Affirmative Action in South Africa: An Empirical Assessment of the Impact on Labour Market Outcomes." In G. Brown, A. Langer, and F. Stewart, eds., *Affirmative Action in Plural Societies*. Basingstoke, UK: Palgrave, 2012.
- CASSIDY, H., DEVARO, J., AND KAUFMAN, A. "Promotion Signaling, Gender, and Turnover: New Theory and Evidence." *Journal of Economic Behavior & Organization*, Vol. 126 (2016), pp. 140–166.
- COATE, S. AND LOURY, G.C. "Antidiscrimination Enforcement and the Problem of Patronization." *American Economic Review*, Vol. 83 (1993a), pp. 92–98.
- . "Will Affirmative-Action Policies Eliminate Negative Stereotypes?" *American Economic Review*, Vol. 83 (1993b), pp. 1220–1240.
- DATO, S., GRUNEWALD, A., KRÄKEL, M., AND MÜLLER, D. "Asymmetric Employer Information, Promotions, and the Wage Policy of Firms." *Games and Economic Behavior*, Vol. 100 (2016), pp. 273–300.
- DEVARO, J., GHOSH, S., AND ZOGHI, C. "Job Characteristics and Labour Market Discrimination in Promotions: New Theory and Empirical Evidence." Working Paper, 2012.
- DEVARO, J. AND KAUFMAN, A. "An 'Opposing Responses' Test of Classic versus Market-Based Promotion Tournaments." *Journal of Labor Economics*, Vol. 34 (2016), pp. 747–779.
- DEVARO, J. AND WALDMAN, M. "The Signaling Role of Promotions: Further Theory and Empirical Evidence." *Journal of Labor Economics*, Vol. 30 (2012), pp. 91–147.
- ESTEVA, F., GALL, T., LEGROS, P., AND NEWMAN, A.F. "College Admission and High School Integration." Department of Economics—FEA/USP Working Paper no. 2014–26, 2014.
- EWERHART, C. "Unique Equilibrium in Rent-Seeking Contests with a Continuum of Types." *Economics Letters*, Vol. 125 (2014), pp. 115–118.
- FANG, H. AND MORO, A. "Theories of Statistical Discrimination and Affirmative Action: A Survey." In J. Benhabib, A. Bisin, and M. Jackson, eds., *Handbook of Social Economics*. Amsterdam: Elsevier, AI. Pages 133–200, 2011.
- FANG, H. AND NORMAN, P. "Government-Mandated Discriminatory Policies: Theory and Evidence." *International Economic Review*, Vol. 47 (2006), pp. 361–389.
- FRYER, R.G. "Belief Flipping in a Dynamic Model of Statistical Discrimination." *Journal of Public Economics*, Vol. 91 (2007), pp. 1151–1166.
- FRYER, R.G. AND LOURY, G.C. "Affirmative Action and Its Mythology." *Journal of Economic Perspectives*, Vol. 19 (2005), pp. 147–162.
- GHOSH, S. AND WALDMAN, M. "Standard Promotion Practices Versus Up-or-Out Contracts." *RAND Journal of Economics*, Vol. 41 (2010), pp. 301–325.
- GREENWALD, B.C. "Adverse Selection in the Labor Market." *Review of Economic Studies*, Vol. 53 (1986), pp. 325–347.

- GÜRTLER, M. AND GÜRTLER, O. "The Optimality of Heterogeneous Tournaments." *Journal of Labor Economics*, Vol. 33 (2015), pp. 1007–1042.
- HOLMSTRÖM, B. "Managerial Incentive Problems—A Dynamic Perspective." In *Essays in Economics and Management in Honor of Lars Wahlbeck*. Helsinki, 1982. Reprinted in *Review of Economic Studies*, Vol. 66 (1999), pp. 169–182.
- HOLZER, H. AND NEUMARK, D. "Assessing Affirmative Action." *Journal of Economic Literature*, Vol. 38 (2000), pp. 483–568.
- KRISHNAMURTHY, P. AND EDLIN, A. "Affirmative Action and Stereotypes in Higher Education Admissions." NBER Working Paper no. 20629, 2014.
- KURTULUS, F.A. "Affirmative Action and the Occupational Advancement of Minorities and Women During 1973–2003." *Industrial Relations*, Vol. 51 (2012), pp. 213–246.
- LANG, K. AND LEHMANN, J.-Y.K. "Racial Discrimination in the Labor Market: Theory and Empirics." *Journal of Economic Literature*, Vol. 50 (2012), pp. 959–1006.
- LEONARD, J.S. "Wage Disparities and Affirmative Action in the 1980's." *American Economic Review*, Vol. 86 (1996), pp. 285–289.
- MAS-COLELL, A., WHINSTON, M.D., AND GREEN, J.R. *Microeconomic Theory*. Oxford: Oxford University Press, 1995.
- MILGROM, P. AND OSTER, S. "Job Discrimination, Market Forces, and the Invisibility Hypothesis." *Quarterly Journal of Economics*, Vol. 102 (1987), pp. 453–476.
- MORO, A. AND NORMAN, P. "Affirmative Action in a Competitive Economy." *Journal of Public Economics*, Vol. 87 (2003), pp. 567–594.
- OWAN, H. "Promotion, Turnover, Earnings, and Firm-Sponsored Training." *Journal of Labor Economics*, Vol. 22 (2004), pp. 955–978.
- PHELPS, E.S. "The Statistical Theory of Racism and Sexism." *American Economic Review*, Vol. 62 (1972), pp. 659–661.
- PINKSTON, J.C. "A Model of Asymmetric Employer Learning with Testable Implications." *Review of Economic Studies*, Vol. 76 (2009), pp. 367–394.
- SCHÖNBERG, U. "Testing for Asymmetric Employer Learning." *Journal of Labor Economics*, Vol. 25 (2007), pp. 651–691.
- SHANKAR, K. "Multi-Skilling, Specialization and Lateral Job Mobility." Working Paper, 2016.
- SMITH, J.P. AND WELCH, F. "Affirmative Action and Labor Markets." *Journal of Labor Economics*, Vol. 2 (1984), pp. 269–301.
- WALDMAN, M. "Job Assignments, Signalling, and Efficiency." *RAND Journal of Economics*, Vol. 15 (1984), pp. 255–267.
- . "Classic Promotion Tournaments Versus Market-Based Tournaments." *International Journal of Industrial Organization*, Vol. 31 (2013), pp. 198–210.
- . "The Dual Avenues of Labor Market Signaling." *Labour Economics*, Vol. 41 (2016), pp. 120–134.
- ZÁBOJNÍK, J. "Promotion Tournaments in Market Equilibrium." *Economic Theory*, Vol. 51 (2012), pp. 213–240.
- ZÁBOJNÍK, J. AND BERNHARDT, D. "Corporate Tournaments, Human Capital Acquisition, and the Firm Size-Wage Relation." *Review of Economic Studies*, Vol. 68 (2001), pp. 693–716.