ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF BUSINESS, ECONOMICS AND LAW
UNIVERSITY OF GOTHENBURG
259

Telling Talent:

Essays on Discrimination and Promotion Contests

Timm Behler

UNIVERSITY OF
GOTHENBURG

ISBN 978-91-88199-77-5 (printed)
ISBN 978-91-88199-78-2 (pdf)
ISSN 1651-4289 (printed)
ISSN 1651-4297 (online)

Printed in Sweden,
Stema 2024

SVANENMÄRKET
Trycksak
3041 0234

To my family and friends.

Acknowledgements

Doing this PhD was quite a journey, which, over the past years, came with many ups and downs. Luckily, I have met many people along the way whose support and friendship helped to push through.

My biggest gratitude goes out to my two amazing supervisors, Dawei Fang and Olof Johansson-Stenman. Our many meetings always helped me to find the right path forward, and without them I would have felt so much more lost while writing the chapters of this thesis or looking for a job. You're both great, and I am grateful to have benefited from that! A special thank you also goes out to Katarina Nordblom, whose mentorship over the first two years of the program is not forgotten! I also want to mention Martin Dufwenberg, who gave so many comments on my chapters and was an immense help throughout the job market experience. I hope to see all of you again.

Louis-Pierre Lepage, Eirik Gaard Kristiansen, and Johan Stennek discussed my chapters in three final seminars. Their comments have prompted many changes in this dissertation—some of them last minute before print—and I can really say that they have hugely improved all of the chapters. For this, I am grateful to them. In advance, I also thank all the members of my defense committee: Markus Dertwinkel-Kalt, Anja Prummer, Erik Mohlin, and, once more, Johan Stennek.

Throughout the PhD, my work has not only benefitted from my supervisors and discussants, but also from the all the comments I received from the faculty here in Gothenburg. Some of you were an immense help on the job market as well. I want to thank Randi Hjalmarsson, Andreas Dzemski, Joe Vecci, Eva Ranehill, Li Chen, Andreea Mitrut, Henrik Petri, as well as many people who gave me extremely insightful comments while I presented my job market paper for the 100th time at some internal seminar.

I am also very grateful to all the other PhD students and non-econ friends in Gothenburg, at home, or elsewhere. Cecilia,[1] Tim, Andreas, Ville,

[1] Extra thanks go to you for being my spinning partner together with Petter who is mentioned further down below.

Jimmy, Vorada, Axel, Chi, Yannik, Giulia, Mattias, Xiaoxu, Ruijie, Richard, Elliott, Stefano, Juanita, Ziheng, and Chenlu—thanks to all of you. Thanks to Patrik for talks about music and wine, but also for being a great coauthor; to Daniel for all the beers and talks about whatever really, and just for the friendship; and to Jens, similarly, for being a great friend with whom I even had the pleasure to share multiple offices over the years—and yes, of course for being a great coauthor as well![2] Thanks to Sonny, Josefine, Petter,[3] Anna, Aigerim, Simon, Niklas, Johannes, Kevin, Paul, Lenny, Franzi, Chiara, Alex, Maja, and all the ones I am forgetting right now. All of you are important to me, and I am grateful for all the time spent together. And thank you Tati, without your support and encouragement over the years I do not think I would have been able to write this thesis.

Lastly, I thank my parents Christine and Helmut, my sister Katharina, and my grandmother Gisela for all their support over the years, and for believing in me.

* * *

See you soon friends!

Gothenburg, March, 2024
Timm Behler

[2]Our offices were always amongst the most desirable in the department, and I believe that this is largely due to the many plants we have accumulated over the years. At this point I can admit that they only survived due to Jens.

[3]See Footnote 1 on the previous page.

Contents

List of Figures

List of Tables

Introduction

Many of our interactions are shaped by the beliefs we hold about the people we interact with. Because of this, we often try to gather information about others. This information shapes our beliefs, which then inform the decisions we make. For example, believing that one group is inherently worse than another might lead us to discriminate against that group. Similarly, believing that one employee is better than another might lead a manager to promote that employee over the other. This thesis deals with how people collect and interpret information about others, broadly speaking, and applies this to discrimination and promotion contests.

The first chapter, "Salience-Based Stereotyping," deals with the question of how members of a socially dominant group form incorrect and overly negative beliefs about members of a socially dominated group. In my model, everyone is endowed with productivity, which depends on an intrinsic part as well as an additive education parameter. In the socially dominated group, there are less educated individuals—for example, due to historic discrimination. An agent from the dominant group wants to learn the average intrinsic productivity of the dominated group. Whenever he observes a productivity signal from one of the educated individuals among the dominated group, however, the relatively low share of educated individuals among them makes education very salient to him. As a result, he overweights the importance that education had in generating the observed productivity. Conversely, this implies that he will underweight the importance of intrinsic productivity. This will, in the long run, lead to overly negative beliefs about this intrinsic productivity parameter, based on which the dominant group discriminates against the dominated. Beyond this, the model can also explain other things, like in-group bias or stereotype substitution, as well as make additional predictions regarding the effects of affirmative action.

The second chapter, "Meritocracy in Hierarchical Organizations," is coauthored with Patrik Reichert and deals with the question whether competitive promotion schemes are—as often believed—really best in selecting talented members of an organization. More specifically, we consider an organization with three tiers. The middle tier receives ability signals of the members of the bottom tier and then proceeds to promote one of them to the middle tier. Afterwards, the incumbent and the entrant at the middle tier compete over a promotion to the highest tier. In our case—unlike under the standard modeling paradigm—the competition on the organization's middle tier creates moral hazard: The incumbent has an incentive to promote a subordinate who has low ability as this makes competing against this person in the future easier. This way the incumbent can secure her own career prospects—at least to some extent. We show that, as a result, uncompetitive promotion mechanisms which are based on seniority rather than performance might outperform competitive schemes in their selection power. At the same time, we identify a novel trade-off between the selection of talent onto the middle and top tiers of the hierarchy. Whether this trade-off occurs, and whether more noisy environments favor or disfavor the competitive scheme depends crucially on the distribution of talent at the lowest tier of the organization.

The third chapter, "Promotions Under Excessive Workplace Surveillance," deals with a similar question as the previous chapter: How can promotions in an organization be designed such that the (expected) talent—which is ex-ante unknown to the principal—of whomever she chooses to promote is maximized. This time, I consider an organization with only two tiers and examine the extent to which the principal should monitor the workers. More monitoring implies that the principal gathers less noisy information about the performance of each worker. As it turns out, expected talent is maximized when some noise is introduced to the evaluation. Would the evaluation be completely noiseless, a lower ability employee would drop out of the competition since she would have no chance of winning. This means that also an employee with high ability would reduce her performance because she is now without competition. Once the high ability agent has reduced her performance however, the low ability agent has a winning chance again and therefore an incentive to enter the competition after all. This logic implies that equilibrium can only exist in mixed strategies. These introduce endogenous noise to the performance, which can distort the relation between performance and ability even more than the exogenous noise introduced by less monitoring.

Chapter 1

Salience-Based Stereotyping

Timm Behler

Abstract

I propose a model in which an agent from a socially dominant group learns about the average productivity of a socially dominated group. Productivity is determined by an intrinsic productivity parameter and education. The agent updates his beliefs overweighting the importance of education based on how salient it appears to him among a given group. The model can account for several stylized facts: i) The agent will hold persistently negative beliefs about the dominated group; ii) he will be subject to in-group bias, exhibiting overly positive beliefs about his own group; and iii) adding a new dominated group can improve the dominant group's perception of the old dominated group. Additionally, it predicts that iv) stereotyping is particularly extreme if the agent learns mostly from "tokens"—in which case the agent's beliefs get more negative even though his sample mean increases—and that v) affirmative action aimed at one of two dominated groups can help that group but hurt the other. Lastly, the model provides a novel connection between taste-based, accurate statistical, and inaccurate statistical discrimination, leading to a feedback loop between discrimination in education and inaccurate beliefs about the dominated group.

1.1 Introduction

Stereotypes are often inaccurate yet persistent. Devine and Elliot (1995), for example, show that the content of the stereotypes black Americans have to face has changed very little over the last century. Such persistence is difficult to reconcile with standard economic learning models: With sufficient access to data, inaccurate stereotypes, for example about a group's average productivity, should converge to the truth. Based on insights from social psychology, this paper provides a possible explanation for why stereotypes are hard to unlearn, even with infinite access to data, and explores implications thereof.

The explanation rests on a psychologically well-founded observation, which—to the best of my knowledge—has not been integrated into the economics literature:[1] Individuals focus on features that are inconsistent with the expectation of the stereotyped group to "explain away" incongruities between what they believe and what they observe. I incorporate this idea into an economic model of stereotyping: An agent from a socially dominant group sequentially learns about the average productivity of a socially dominated group.[2] Productivity depends on a productivity parameter whose distribution is intrinsic to the social group, and on whether an individual is educated or not. Based on how representative of the dominated group education appears to the agent, he over- or underaccounts for its role in generating the observed data. I show, among other things, how this can lead the agent to hold wrong beliefs about both the agent's own and other social groups that persist even in the long run: Beliefs about the stereotyped group will be negatively biased whereas beliefs about the agent's own group will be positively biased. These predictions are consistent with empirical observations and an assumption in several theories of discrimination in the psychology literature.[3]

Indeed, the logic of this paper is related to the social-cognitive approach to stereotyping, advocated for by many social psychologists (see, e.g., Sidanius and Pratto 2001 for an overview). Following Allport, Clark, and Pettigrew

[1]For evidence from psychology, see, for example, Erber and Fiske (1984), Neuberg and Fiske (1987), Pratto and Bargh (1991), or Hilton and Von Hippel (1996).

[2]I borrow the terms "socially dominant" and "socially dominated" from the psychology literature on "social dominance" (Sidanius and Pratto 2001). The idea is that many societies are hierarchical in nature, whereby some groups (the dominant groups) control more resources (e.g., the labor market) than others (the dominated groups), or have otherwise more influence on social outcomes.

[3]See, for example, Sumner (1906), Allport, Clark, and Pettigrew (1954), or Tajfel and Turner (1982).

(1954), its main conclusion is that stereotypes are the result of unassuming learning processes. They argue, however, that individuals often have to simplify their data to make sense of them, or have a predisposition towards other imperfect learning rules, which leads them to hold potentially wrong and overly negative beliefs about other social groups.

The social-cognitive approach contrasts the paper with the classical economics literature on stereotyping and discrimination, where discrimination is assumed to be based either on *true* differences between groups ("statistical discrimination:" see, e.g., Phelps 1972) or on a decision maker's distaste against a specific social group ("taste-based discrimination:" see, e.g., Becker 1957). Only more recently have economists incorporated the idea that stereotyping and discrimination can be based on wrong beliefs (see, e.g., Bordalo et al. 2016; Sarsons 2017; Bohren et al. 2019; Heidhues, Kőszegi, and Strack 2023b).[4,5,6] My model, too, can be interpreted in this context of "inaccurate statistical discrimination" in that stereotypes are not understood as an intrinsic distaste against a certain group or a reaction to true group differences, but rather as an *incorrect* belief in certain group differences.

Model Overview. In section 3.2 of this paper, I introduce the formal model. An agent who belongs to a socially *dominant* group sequentially observes the productivities of the members of a socially *dominated* group. The observations come in two flavors, as members of the dominated group can be one of two types: They can be *educated* or *uneducated*. Types are observable to the agent. An individual's productivity is a random variable. Its distribution depends in part on an intrinsic productivity parameter whose (unconditional) distribution is common among all members of the dominated group, and on an productivity parameter whose (conditional) distribution depends on whether a person is educated or uneducated. I make two assumptions: First, educated individuals have, on average, higher productivity than uneducated individuals; second, the agent believes that being educated is less likely for members of the dominated group than for the dominant group. Such a belief would be correct for different racial

[4]In some of these papers, stereotypes may still contain a "kernel of truth," but stereotypes will amplify these real group differences. See also Judd and Park (1993) for work on this kernel-of-truth hypothesis from psychology.

[5]One reason for why economists have incorporated the concept of inaccurate statistical discrimination into their models only recently may be that it is difficult to disentangle from other forms of discrimination. Recent work tackles this (see, e.g., Feld et al. 2022).

[6]See also Section 4 in Onuchic (2022) for an overview of recent contributions to the literature on inaccurate statistical discrimination.

groups in the US, where both black Americans and Hispanics have lower rates of college completion than white Americans (Pew Research Center 2019), or regarding gender differences in specific educational domains like engineering (Pew Research Center 2021).

The agent makes one crucial mistake when updating his beliefs and is, in this sense, only boundedly rational. In a logic similar to that in Bordalo et al. (2016),[7,8] the agent (subconsciously) evaluates the representativeness of being educated for the dominated group by comparing the likelihood of a member of the dominated group being educated with that of the dominant group. Because the agent believes being educated to be less likely for members of the dominated group, he perceives education to be less representative of them. Drawing an observation from an educated member of the dominated group is, in this sense, unexpected, which makes the respective individual's educational attainment salient to him. To capture the psychological logic I am interested in, I assume that the resulting salience of education leads the agent to overweight its importance for generating the observed data with respect to the true statistical model of the world. The salience-based bias I introduce allows the agent's misspecification to be context-dependent. Indeed, it is the context dependency of the bias which allows me to explain several stylized facts within one unified framework—and make additional predictions.

From a behavioral perspective, my model is most closely related to Bordalo et al. (2016) due to the use of the representativeness heuristic to derive stereotypes. Their paper, however, does not include learning and, instead, the representativeness of certain characteristics distorts the recall of a distribution. Among the papers on biased beliefs about social groups that consider learning, my paper is most closely related to Heidhues, Kőszegi, and Strack (2023b), some of whose predictions overlap with mine—most notably regarding negative stereotypes and in-group biases. Their main mechanism, however, is different and based on overconfidence rather representativeness. As it turns out, this distinction is important. For

[7]This idea is related to other work that incorporates salience into economic models (see, e.g., Bordalo, Gennaioli, and Shleifer 2022 for an overview). Here, a decision maker—for concreteness, we can think of her as a consumer—overweights the importance of salient product traits (e.g., price vs. quality) when she evaluates the utility that a purchase decision will generate. As a result, purchase decisions may be both context dependent and suboptimal when compared to an evaluation according to the agent's true utility function.

[8]The authors of these previous papers introduce this logic to provide an economic formalization of the representativeness heuristic by Kahneman and Tversky (1972). For other economic research around the representativeness heuristic see Barberis, Shleifer, and Vishny (1998) or Dumm et al. (2020).

example, in my model it is straightforward to allow for the analysis of common anti-discriminatory policies, such as affirmative action, because they directly affect representation—i.e., the foundation of the bias in my model. Overconfidence, however, is not directly affected by such policies. Indeed, overconfidence is known to be a particularly persistent bias and tough to combat with policies, making the results in Heidhues, Kőszegi, and Strack (2023b) more pessimistic than mine.

Outline and Results. In section 1.3, I first apply the model to show how a negative stereotype about the dominated group's productivity can both emerge and persist. The mechanism described above leads the agent to overweight the importance of education for generating the observed data. Because educated individuals have on average a higher productivity than ordinary individuals, he will often misattribute high signals to a high education. Conversely, the agent will underaccount for the role of the individual's intrinsic productivity, leading to a negative stereotype—even in the long run.

The logic also sheds light on the observation that different (potentially incorrect) stereotypes can interact with and cause each other (see, e.g., Hamilton and Rose 1980 for experimental work from psychology): The negative bias against the socially dominated group is more pronounced the lower the agent believes the share of educated individuals among the dominated group to be. If the agent holds, for example, a negative stereotype regarding the educational attainment of the dominated group, this stereotype will increase the salience of education and, hence, lead to a more severely negative stereotype regarding the dominated group's productivity.

After deriving the baseline result, I briefly interpret it in light of the discussion around "tokens." Tokens are members of an underrepresented group that are placed in prestigious positions—e.g., on a corporate board—simply so that the responsible organization can be perceived as inclusive. While the social psychology literature discusses several reasons why such tokens may increase the extent to which members of a disadvantaged group are negatively stereotyped, my model provides a novel explanation: Intuitively, individuals who are tokens are often highly educated. Additionally, since many real-world societies face segregation between dominant and dominated groups, tokens are typically figures through which the dominant group learns about some average characteristic of the dominated group. When the agent learns about the dominated group mostly through tokens, he therefore learns mostly from educated people. The more he learns from educated people, the more biased are his long-run beliefs since the agent only misperceives the signals generated by educated people. Perhaps counterintuitively, because educated people have a higher expected productivity than uneducated people, this also implies that the agent's belief about the dominated group's average productivity can *decrease* as the average productivity in his sample *increases*.

In section 1.4, I show how my model can help to explain in-group bias—i.e., overly positive views that the an agent from the socially dominant group has bout the average productivity of his own group. To do this, I allow the agent to be uncertain about the average productivity of the socially dominant group and learn about it over time too. A logic that is, in a sense, reverse to that described above implies that the agent will perceive being educated to be very representative of the dominant group, and, hence, it will *not* be salient to him. As a result, he underweights the importance of education for the generated data, which leads him to have an inflated view about the average productivity of the dominant group. Such an in-group bias is in line with empirical observations (see Mullen, Brown, and Smith 1992 for a meta-analysis). The result is related to an idea from sociology and psychology, where socially dominant groups have privileges—e.g., because of historic and ongoing discrimination—that they do not necessarily recognize (see, e.g., Collins 2018). In my model, because it is so expected for members of the dominant group to be educated, they fail to account for the positive effect of education when updating their beliefs, generating the positive bias.

In section 1.5, I investigate what happens when a second, new dominated group enters the society. We can think of such a group as a more recent cohort of immigrants. Assuming that the agent believes that the second dominated group, too, is less likely to be educated than the dominant group, its introduction affects how the old dominated group is evaluated: The agent will now compare the old dominated group to both the dominant group and the new dominated group. Education is now less salient among the old dominated group than before and, as a result, the long-run beliefs about the old dominated group's average productivity will be less negatively biased than before. As such, my model can also provide an explanation for the empirical observation that upon the entry of a new dominated group, stereotypes regarding the old dominated group often decrease and seem to center around members of the new group instead (see, e.g., Fouka, Mazumder, and Tabellini 2022).

In section 1.6, I stay in the framework with two dominated groups to show how affirmative action can benefit one dominated group while hurting the other. I introduce a policy maker who implements an affirmative action policy which increases the share of educated individuals among the old dominated group but not among the new. The reason could be either that the affirmative action policy is aimed at one of the groups specifically, or, equivalently, that it only reaches one of them—e.g., due to an asymmetry in the language barriers the two groups face. Such a policy can—but must not—decrease the extent of stereotyping faced by the old group: Education will be less salient among the old group and, hence, the agent overweights it less in the updating process, suggesting less biased beliefs in the long run. There is, however, also a counteracting effect: When more people are educated, the agent more often processes the productivity signals in a biased way, suggesting more biased beliefs in the long run. The effect of affirmative action on the old dominated group is therefore ambiguous and depends on the share of educated people among the dominated group before the affirmative action policy is implemented.

At the same time, however, the new dominated group is always harmed by the policy: The comparison group of the new dominated group becomes on average more educated, which increases the salience of education among the new group. This implies that the importance of education will be more overweight compared to the case without affirmative action. In this sense, the new dominated group is negatively affected by affirmative action. This suggests that, while affirmative action can be beneficial for some groups, it must be designed carefully as to make sure that no group that is already marginalized is harmed by it.

Finally, in section 1.7, I provide a discussion of the results around the idea of "legitimizing myths" (Sidanius and Pratto 2001)—i.e., beliefs that help members of the dominant group to justify a social hierarchy. Based on this discussion, I also provide a simple formal argument about how stereotypes might interact with more active forms of discrimination (e.g., the decision *not* to select a qualified applicant from a dominated group) and how stereotyping and this type of discrimination can help each other to co-exist. I further show how both classical types of discrimination—taste-based and statistical—can cause stereotyping and, as a result, lead to inaccurate statistical discrimination. section 1.8 discusses related literature; literature will also be discussed throughout the text where useful. section 3.7 concludes. All proofs are in subsection 2.A.4.

1.2 Model

1.2.1 The Society

Consider an infinite population of individuals, each denoted by i. Each individual is defined by three parameters: (i) a productivity parameter, $a_i \in$

$\mathbb{R}$; (ii) an education level, $e_i \in \{u, e\}$ (*uneducated* or *educated*, respectively); and (iii) a social group membership $s_i \in \{w, b\}$. Borrowing from the psychology literature, I say that all individuals i for whom $s_i = w$ form the *socially dominant* group and those for whom $s_i = b$ form the *socially dominated* group. Often it is useful to think of the dominated group as an ethnical minority (e.g. *b*lack people in the US) while thinking of the dominant group as an ethnical majortiy (e.g. *w*hite people in the US). I will refer to members of the dominant group as *he* and to members of the dominated group as *she*. An individual's *type*, denoted θ_i, refers to the pair consisting of his or her education level and his or her social group; an individual's producticity is *not* part of his or her type. Formally, we have $\theta_i = (s_i, e_i) \in \Theta := \{w, b\} \times \{u, e\}$.

Finally, I denote the share of educated individuals among the socially dominated group by $\beta_b \in (0, 1)$. Similarly, I denote the share of educated people among the socially dominant group by $\beta_w \in (0, 1)$. Crucially, I assume that the share of highly educated people among the socially dominant group is larger than the share of highly educated people among the dominated group, $\beta_b \leq \beta_w$. This captures, for example, historical discrimination. It is also precisely in this way that we can think of one group as being dominated and one being dominant; the dominant group may have some decision power over who is able to get educated and therefore may be responsible for the lower rate of education among the dominated group.

1.2.2 Productivity

Each individual is endowed with the productivity parameter a_i. Specifically, we have that an individual's productivity is given by

$$a_i = a_{s_i} + \rho a_{e_i}, \tag{1.1}$$

where a_{s_i} is called the individual's *intrinsic productivity parameter* and a_{e_i} is called the individual's *education-based productivity parameter*; they are realizations of draws from the distributions G_{s_i} and H_{e_i}, respectively. The distribution G_{s_i} from which a_{s_i} is drawn can differ across social groups. For now I make the simplifying assumption that G_{s_i} is a normal distribution with mean μ_{s_i}, and variance $1/\lambda_{s_i}$—i.e.,[9]

$$a_{s_i} \sim \mathcal{N}(\mu_{s_i}, 1/\lambda_{s_i}). \tag{1.2}$$

[9]In subsection 1.A.2, I consider more general distributions and show that the main insight of the paper carries over.

We can therefore interpret λ_{s_i} as a precision parameter. Notice that the parameters—μ_{s_i} and λ_{s_i}—can, in principle, differ between the two social groups, but they need not do so. The model can therefore account for stereotyping in cases both with and without true underlying differences in the productivity.

The distribution H_{e_i} from which the individual's education-based productivity parameter, a_{e_i}, is drawn is specified as follows: In the case where $e_i = u$, I assume that that H_u is a Dirac measure at the point zero. If $e_i = e$ instead, I assume that H_e is a Dirac measure at the point $\alpha > 0$. Finally, the parameter $\rho \in \mathbb{R}_+$ simply captures how important education is for an individual's productivity compared to the individual's social group membership. For example, a higher ρ suggests that education has a relatively bigger impact.

We can put these distributional assumptions together to get the distribution of an individuals' productivity parameter, a_i, conditional on his or her education level, e_i. Specifically, we have that

$$a_i \mid e_i = u \sim \mathcal{N}(\mu_{s_i}, 1/\lambda_{s_i}) \quad \text{and} \quad a_i \mid e_i = e \sim \mathcal{N}(\mu_{s_i} + \rho\alpha, 1/\lambda_{s_i}), \quad (1.3)$$

where $\mu_{s_i} + \rho\alpha > \mu_{s_i}$—i.e., fixing an individual's social group, an educated individual has on expectation a higher productivity than an uneducated individual. For future convenience, I will denote the mean of the second distribution in expression (1.3) by $\mu_{s_i,e} := \mu_{s_i} + \rho\alpha$.

1.2.3 The Agent

There is an agent from the dominant group who, in each of infinitely many periods $t = 1, 2, ...$, randomly draws a member of the dominated group from the population. The agent can observe the individual's realized productivity together with her type.[10] I denote the productivity of the individual who was drawn in period t as a_i^t. Similarly, I denote her type as θ_i^t. We can interpret these productivity signals in different ways. For example, the observed individual may be a coworker of the agent, and her observed productivity may be her performance at the job. Otherwise, the observed individual may be a public figure, and her observed productivity may be conveyed through a news article or a public statement made by the individual.

The agent does not (ex ante) know the mean of the dominated group's intrinsic productivity parameter, μ_b, and is interested in learning it. For simplicity, I assume that the agent knows all other parameters of the model.

[10] I assume for now that the agent can perfectly (i.e. without noise) observe individuals' productivities. This is easily extended to the case where he only receives a noisy signal of the productivities (see subsection 1.A.2).

The agent has a prior belief about μ_b. I denote this prior belief by $\tilde{\mu}_b^0$, and assume that it is distributed according to

$$\tilde{\mu}_b^0 \sim \mathcal{N}(\bar{\tilde{\mu}}_b, 1/\lambda_b^0), \tag{1.4}$$

where $\bar{\tilde{\mu}}_b$ may or may not be equal to the group's true mean, μ_b. Ex ante, the agent's subjective distribution of the dominated group's intrinsic productivity parameter must reflect the uncertainty in his prior belief about its mean. Hence, it is given by

$$a_b \sim \mathcal{N}(\bar{\tilde{\mu}}_b, 1/\lambda_b + 1/\lambda_b^0). \tag{1.5}$$

1.2.4 Updating Procedure

I assume that the agent is a *salient thinker*: He updates his beliefs in a Bayesian way but may use a misspecified statistical model of the world when doing so. The misspecification is based on how salient highly educated individuals among the dominated group appear to him. This is my only deviation from an otherwise standard Bayesian learning model.

More specifically, instead of the model given in expression (1.1), he updates his beliefs as if the model would be given by

$$a_i = a_{s_i} + \tilde{\rho}_{s_i} a_{e_i}, \tag{1.6}$$

where $\tilde{\rho}_{s_i} \neq \rho$. We can interpret $\tilde{\rho}_{s_i}$ in different ways. For example, we can either think of it as simply a wrong (but persistent) belief or a deviation from (rational) Bayesian updating. In this paper, I prefer the latter interpretation and assume that $\tilde{\rho}_{s_i}$ is a distorted version of ρ that depends on a representativeness parameter, which is specified in more detail below (Definition 1 and Assumption 1). This makes the misspecified model of the world from expression (1.6) endogenous to other parameters of the model, which is in contrast to most of the related learning literature where misspecifications are usually taken to be exogenous.[11]

Similar to the model of stereotyping by Bordalo et al. (2016) (whose model is in turn based on the representativeness heuristic due to Kahneman and Tversky (1972)), I assume that the representativeness of being highly educated for either of the two social groups is given as follows.

[11] A very recent exception is Esponda, Oprea, and Yuksel (2023), although they are, unlike this paper, not interested in long-run learning outcomes.

Definition 1.1 (Representativeness)**.** *The representativeness of an individual of either social group being highly educated compared to the other social group is given by*

$$R_{s_i} := \frac{\Pr(e_i = e \mid s_i)}{\Pr(e_i = e \mid -s_i)} = \frac{\beta_{s_i}}{\beta_{-s_i}}, \tag{1.7}$$

where $-s_i := \{w, b\} \setminus s_i$.

In words, this implies that being highly educated is more representative for one social group compared to the other the higher the respective likelihood ratio in expression (1.7) is. Keeping this definition in mind, we can further specify the relationship between ρ, $\tilde{\rho}_{s_i}$, and R_{s_i}. Specifically, I assume the following:

Assumption 1.1 (Representativeness and Salience-Based Weighting)**.** *For all* $s_i \in \{w, b\}$*, we have that: (i)* $\tilde{\rho}_{s_i}$ *is continuously differentiable and decreasing in* R_{s_i}*; (ii)* $R_{s_i} = 1 \Rightarrow \tilde{\rho}_{s_i} = \rho$*; (iii)* $R_{s_i} < 1 \Rightarrow \tilde{\rho}_{s_i} > \rho$*; (iv)* $R_{s_i} > 1 \Rightarrow \tilde{\rho}_{s_i} < \rho$*.*

The idea is that the less representative a high education is for one of the social groups, the more this trait sticks out—i.e., the more *salient* it appears to the agent. Based on this salience, the agent distorts the importance of education for having generated the observed signal. Such a salience-based weighting of certain traits is common in the economics literature on salience (Bordalo, Gennaioli, and Shleifer 2012b, 2013b; Bordalo et al. 2016; Bordalo, Gennaioli, and Shleifer 2022).

Notice that the misspecification described by expression (1.6) is only relevant in the periods in which the agent draws a signal produced by an educated individual. In any other period, when he draws a signal generated by an uneducated individual, he evaluates this signal using the correct model.

1.3 Persistent Stereotypes

In this section, I show how my model can help to explain persistently wrong and overly negative beliefs about the socially dominated group. After deriving the baseline result, I interpret it in light of how stereotypes in different domains can interact with each other, and how stereotyping can be particularly bad if the agent learns mostly from "tokens."

1.3.1 Persistent Stereotypes as Long-Run Beliefs

We are interested in the agent's long-run beliefs about the average value of the dominated group's intrinsic productivity, μ_b. We know that in each period t, the agent receives a signal drawn from one of two possible distributions. Specifically, with frequency β_b, he receives signals drawn from $\mathcal{N}(\mu_b + \rho\alpha, 1/\lambda_b)$ which he will interpret as if they were drawn from $\mathcal{N}(\mu_b + \tilde{\rho}_b\alpha, 1/\lambda_b)$. With frequency $1 - \beta_b$, instead, he receives signals drawn from $\mathcal{N}(\mu_b, 1/\lambda_b)$. Because the agent also observes the type of the individual who produced the signal, he knows in any given period from which distribution the signal in that period was drawn. In the case where he receives a signal from an uneducated member of the socially dominated group, he can directly incorporate this signal into his posterior. In the case where he receives a signal from an educated member of the socially dominated group, he first has to back out the parameter of interest, a_b, before he can incorporate the signal into his posterior.

To derive an explicit expression of the agent's period-t beliefs, I introduce some additional notation. Specifically, I denote by $\mathcal{S}^t_{b,e}$ the set of all signals that where generated by an educated member of the dominated group until period t. Analogously, $\mathcal{S}^t_{b,u}$ denotes the set of all signals generated by an uneducated member of the dominated group until period t. Finally, I denote the cardinality of the respective sets by $|\mathcal{S}^t_{b,e}|$ and $|\mathcal{S}^t_{b,u}|$, such that $|\mathcal{S}^t_{b,e}| + |\mathcal{S}^t_{b,u}| = t$. Standard Bayesian reasoning implies that the agent's posterior belief about μ_b in period t will be given by a normal distribution with mean

$$\bar{\tilde{\mu}}^t_b := \frac{|\mathcal{S}^t_{b,e}|(\bar{\tilde{s}}_{b,e} - \tilde{\rho}_{s_i}\alpha)\lambda_b + |\mathcal{S}^t_{b,u}|\bar{\tilde{s}}_{b,u}\lambda_b + \tilde{\mu}^0_b\lambda^0_b}{|\mathcal{S}^t_{b,e}|\lambda_b + |\mathcal{S}^t_{b,u}|\lambda_b + \lambda^0_b} \tag{1.8}$$

and variance

$$\frac{1}{\bar{\tilde{\lambda}}^t_b} := \frac{1}{|\mathcal{S}^t_{b,e}|\lambda_b + |\mathcal{S}^t_{b,u}|\lambda_b + \lambda^0_b}. \tag{1.9}$$

Hereby,

$$\bar{\tilde{s}}_{b,e} := \frac{1}{|\mathcal{S}^t_{b,e}|} \sum_{s^t_{b,e} \in \mathcal{S}^t_{b,e}} s^t_{b,e} \tag{1.10}$$

is the average of all signals observed at time t that were produced by an individual of type (b, e). The $s^t_{b,t}$ in expression (1.10) denotes the specific productivity observed in period t. Analogously,

$$\bar{\tilde{s}}^t_{b,u} := \frac{1}{|\mathcal{S}^t_{b,u}|} \sum_{s^t_{b,u} \in \mathcal{S}^t_{b,u}} s^t_{b,u} \tag{1.11}$$

is the average of all signals observed at time t that were produced by an individual of type (b, u). The $s^t_{b,u}$ in expression (1.11) denotes the specific productivity observed in period t.

To apply these observations to stereotyping, and to simplify the following discussion around it, I formally define how I think of stereotypes in the remainder of this paper.

Definition 1.2 (Stereotypes). *A stereotype about the dominated group's intrinsic productivity parameter exists in period t if we have that $\bar{\tilde{\mu}}^t_b \neq \mu_b$. I say that the stereotype is negative if $\bar{\tilde{\mu}}^t_b < \mu_b$. The stereotype is positive if $\bar{\tilde{\mu}}^t_b > \mu_b$. The extent of the stereotype refers to the absolute difference between beliefs and the truth, $|\bar{\tilde{\mu}}^t_b - \mu_b|$.*

Intuitively, I think of a stereotype as a wrong belief about the dominated group's intrinsic productivity, which can be either negative or positive. Using the specification of the agent's posterior given in expressions (1.8) through (1.11) together with Definition 1.2 allows me to state my first result, which characterizes the agent's long-run beliefs about the dominated group's intrinsic productivity.

Proposition 1.1 (Persistent Stereotypes). *Suppose that $\beta_b < \beta_w$. As $t \to \infty$, the agent's belief about the average of the dominated group's intrinsic productivity parameter converges in distribution to the value*

$$\tilde{\mu}_b := \mu_b - \beta_b(\tilde{\rho}_b\alpha - \rho_b\alpha) < \mu_b \;\; \forall \beta_b \in (0, \beta_w). \tag{1.12}$$

That is, there is a negative stereotype about the dominated group's average productivity.

Proposition 1.1 states that the agent's belief about average intrinsic productivity of the socially dominated group will be negatively biased in the long-run if the share of educated individuals is smaller among the dominated group than among the dominant group. This is a plausible assumption for many dominated groups that are subjected to stereotyping and discrimination—for example, women ins specific domains, or black Americans more generally (Pew Research Center 2019, 2021). The result also implies that the belief about the average productivity over all members of the dominated group—whether they are educated or not—will be negatively biased.[12] A natural interpretation of this result, in line with Definition 1.2

[12]The overall average belief about the productivity of the dominated group is given by

$$\mathbb{E}[a_m] = \beta_b(\tilde{\mu}_b + \tilde{\rho}_b\alpha) + (1 - \beta_b)\tilde{\mu}_b,$$

which, because $\tilde{\mu}_b < \mu_b$, is smaller than the true population mean.

above, is that there will be a persistent negative stereotype.

Because Proposition 1.1 provides a closed-form solution for the agent's long-run belief, it is possible to provide some comparative statics. I first consider an increase of the share of educated individuals among the dominant group, β_w.

Corollary 1.1 (Share of Educated Individuals Among Dominant Group.)**.** *The agent's long run bias increases as the share of educated individuals among the dominant group, β_w, increases.*

Corollary 1.1 states that when the share of educated individuals among the dominant group is higher, this will result in an increase of the extent of the stereotype against the dominated group. The logic is that being educated becomes more salient among the dominated group, which leads to a more biased updating process. Perhaps intuitively, in the real world too, an increase of educated people among the agent's own group might contribute to a perceived superiority, which can amplify the agent's perception of group differences.

Similarly, we can investigate the effect of an increase of the share of educated individuals among the dominated group, β_b. The effect of such an increase is more subtle, however. The reason is that β_b enters the agent's long-run bias directly, but also indirectly through $\tilde{\rho}_b$. Both of these effects go in opposite directions. The next corollary states this in more detail.

Corollary 1.2 (Share of Educated Individuals Among Dominated Group)**.** *Consider an increase of the share of educated individuals among the dominated group, $\beta_b \in [0, \beta_w]$. Keep all other parameters fixed. There exists a threshold level $\hat{\beta}_b \in (0, \beta_w)$ such that for all $\beta_b < \hat{\beta}_b$, in the long-run, the extent of the negative stereotype is increasing in β_b. Conversely, for all $\beta_b > \hat{\beta}_b$, the extent of the stereotype is decreasing in β_b.*

Corollary 1.2 states that the direction of the effect of an increase of the share of educated members of the dominated group depends crucially on whether this share is low or high to begin with. Specifically, when the share is relatively low, then an increase will increase the extent of negative stereotyping the dominated group faces. If the share of educated individuals is relatively high, however, a further increase will decrease the extent of negative stereotyping faced by the dominated group. This has some important real-world implications. For example, it suggests that an increase in social mobility among marginalized groups can either decrease or increase stereotyping, depending on how easy it was for them to reach a high educational attainment to begin with. Another natural interpretation is affirmative action. Here, the result suggests that affirmative action can decrease stereotyping if the share of educated members of the dominated group is already relatively high, but not if it is very low to begin with. I will discuss implications in the context of affirmative action in more detail in section 1.6.

I finish this section with a remark on the difference in persistence of negative vs. positive initial stereotypes.

Remark 1.1 (Positive vs. Negative Stereotypes)**.** *Suppose that* $\beta_b < \beta_w$. *Negative stereotypes are persistent, positive stereotypes are not.*

Remark 1.1 states that negative stereotypes are persistent: They will not converge to the true parameter value—even though an agent's long run belief can, in principle, be higher than his prior. At the same time, positive stereotypes are *not* persistent: They will converge to a value below the true value in the long-run. Indeed, overly positive initial beliefs can turn into negative stereotypes fast, depending on the precision of the agent's prior and the signals he receives. As such, my model may help to explain the casual observation that many stereotypes about dominated groups—such as ethnic minorities—used in everyday life are negative rather than positive. This idea is a building block in many social psychological and sociological theories of inter-group conflict. In this sense, my paper connects more to the previous research by Heidhues, Kőszegi, and Strack (2023b), who also focus on overly negative beliefs held by an individual about another social group. As such, both papers analyze potential drivers of beliefs that can generate inter-group conflict and discrimination. Another recent paper along this line is due to Lepage (2021). The crucial difference to my paper is that in Lepage (2021), overly positive beliefs converge to the truth, whereas in my model, they turn into negative stereotypes. Glaeser (2005) also focuses on negative stereotypes. In his model, politicians can supply the population with hate-inducing stories about minorities, leading to negative stereotypes. Finally, Hübert and Little (2023), provide a theory in which, due to over-policing in certain regions, more crime is detected among a disadvantaged vs. an advantaged social group. Police fails to account for their selective policing when updating. This leads the police to hold overly negative beliefs about the disadvantaged group, which induces them to police disadvantaged individuals even more. Ultimately this results in a particularly vicious feedback loop, which amplifies negative beliefs over time.

All of the above papers are opposed to more general theories of cognitive limitations that lead to simplified stereotypes, whether they are positive or negative. In Bordalo et al. (2016), for example, stereotypes are only characterized by overly simplified beliefs that blow up true group differences, but they need not only be negative. As we will see in section 1.4, my model can, in principle, also account for positive stereotypes—i.e., overly positive beliefs—about the productivity of a social group. Indeed, only groups who are socially dominated—in the sense that $\beta_b \leq \beta_w$—will be exposed to negative stereotyping.

1.3.2 On the Interplay Between Stereotypes in Different Domains

Both Proposition 1.1 and Remark 1.1 require that $\beta_b < \beta_w$. In this case, the likelihood ratio defined in expression (1.7) implies that being educated is a more representative trait for the dominant group than for the dominated group. As mentioned in the introduction, this assumption is correct for several group differences in the US (Pew Research Center 2019, 2021). As a result, education is more salient when observed on a member of the dominated group and, hence, overweighted in the updating process. This logic drives the results. It is worth noting, however, that it is not necessary that there are actually fewer educated people among the dominated group. Instead, it is sufficient for the agent to *believe* that this is the case. This hints at an interesting interplay between two different stereotypes. Suppose that the agent has a belief about β_b, which is given by $\tilde{\beta}_b < \beta_b$. We can interpret such a belief as a negative stereotype that the dominated group is less educated than it actually is. With this additional assumption, Proposition 1 implies the following.

Corollary 1.3 (Interplay Between Stereotypes in Different Domains)**.** *Suppose that $\tilde{\beta}_b < \beta_b$. The extent of the negative stereotype is increasing as $\tilde{\beta}_b$ decreases.*

Corollary 1.3 implies that a negative stereotype in one domain (e.g., education) can increase stereotyping in another domain (e.g., intrinsic productivity). This effect is on top of any mechanical effect that a low level of education already has on the beliefs about the expected average productivity of a group. This observation is relevant, because it helps to shed light on another mechanism that helps stereotypes to persist: one stereotype can be used to arrive at another stereotype and vice versa. I will further highlight this theme in my discussion around legitimizing myths in section 1.7.

1.3.3 Tokens

In this section, I apply my results to analyze the effects that *tokens* can have on stereotyping. According to the MacMillan dictionary, a token is "someone who is included in a group to make people believe that the group is trying to be fair and include all types of people when this is not really true." In social psychology, there has long been a debate about the effect of tokens on stereotyping. Indeed, social psychologists know that tokens can increase stereotyping through at least two distinct channels: First, they can increase the salience[13] of the marginalized group, making it more easy for other people to connect them to stereotypes (Sidanius and Pratto 2001). Second, being a token can lead to a high social pressure to perform, which can actually lower performance (Word, Zanna, and Cooper 1974). In this sense, tokenism can lead to self-fulfilling stereotypes. The model I present in this paper can be used to derive a different, to the best of my knowledge novel, explanation for how tokens can increase stereotyping.

Typically, tokens still require some high achievement (e.g., high educational attainment) to obtain their high position. That is, in my model, tokens are most likely to be highly educated members of the dominated group. Asking what happens if the agent learns mostly from tokens is therefore akin to asking what happens if he samples more often from the pool of educated individuals than their true share in society would suggest. In my model, we can formalize this by assuming that, while a share β_b of the dominated group is educated, the share of educated individuals in the agent's sample is given by $\check{\beta}_b$. Then, by the logic described above, if the agent learns mostly from tokens, this corresponds to an increase in $\check{\beta}_b$. The following corollary of Proposition 1 summarizes the implication.

Corollary 1.4 (Tokens)**.** *The extent of the negative stereotype regarding the dominated group's intrinsic productivity increases as* $\check{\beta}_b$ *increases.*

That is, learning mostly from tokens leads to more negative stereotyping. To put the result's relevance into perspective, note that many real-world societies face segregation between the dominated and the dominant group. Hence, the dominant group will often learn from the few members of the dominated group that are in some respect similar to themselves—for example, because they have reached similar positions.[14] Notice that for the result to

[13]Notice that in this context, "salience" refers to the salience of being in the dominated group, whereas "salience" in my model refers to the salience of a specific individual's educational attainment.

[14]For a more in-depth analysis of the effect that segregation can have on beliefs about social groups, see Frick, Iijima, and Ishii (2022).

hold, the members of the dominated group from which the agent samples need not necessarily be tokens. It is sufficient that his sample is skewed, for whatever reason, towards educated members of the dominated group.

The result abstracts away from potential positive effects that tokens can have in the real world. For example, whereas they might increase stereotyping, they can also provide role models to other members of the dominated group, ultimately helping them to achieve similar positions and increasing the share of educated people among the dominated group (see, e.g., Marx, Ko, and Friedman 2009). In my model, this would lower the salience of education, and, given that β_m would be sufficiently high, lead to lower stereotypes. Still, previous research suggests that the potential positive effect on encouraging other members of the dominated group is only minor (Sidanius and Pratto 2001).

The intuition behind Corollary 4 is straightforward in my model. However, I want to point out one immediate implication on the relationship between the sample's average productivity and the agent's belief about the population productivity, which is, perhaps, surprising:

Remark 1.2 (Sample Average and the Extent of Stereotyping)**.** *The extent of the stereotype against the dominated group can increase as the average productivity of the agent's sample increases.*

This is precisely the case when $\check{\beta}_b$ increases: In this scenario, the share of educated individuals among the sample increases which also increases the average productivity of the sample. At the same time, by Corollary 1.4, this increases the extent of negative stereotyping, leading to the described relationship.

1.4 In-Group Bias

So far, I have assumed that the agent, who is a member of the socially dominant group, perfectly knows the productivity distribution of other members of his own social group. For various reasons—for example, because the agent can easily interpret signals he receives about the individuals of his own social group—this may be plausible. Yet, a central feature of many theories from social psychology is that individuals tend to overestimate the quality of individuals with whom they share their social group. This is called in-group bias. Indeed, my model can also be applied to an agent from the socially dominant group who learns about the average intrinsic productivity of the members of his own social group. We can simply assume that he is uncertain of this average, and that in each period t, he observes a member of his own group in addition to the signals he receives about the dominated group's productivity. Signals are distributed as described in section 3.2 and all notation is analogous to that in section 1.3. I simply replace the indices b with w.

In the case where the agent learns about his own group, a reverse logic to that discussed in section 1.3 will come into play: Because education is a representative trait of the dominant group, it is less salient to the agent when learning about the average productivity of other members of the dominant group compared to the dominated group. As a result, by Definition 1 and Assumption 1, he will put *less* weight on education when evaluating other members of his own group. This will lead him to develop overly positive beliefs about the average productivity of the dominant group; an in-group bias emerges—and persists. The following proposition summarizes this.

Proposition 1.2 (In-Group Bias)**.** *Suppose that* $\beta_b < \beta_w$. *Then, as* $t \to \infty$*, the agent's belief about the average of the dominant group's intrinsic productivity parameter converges in distribution to the value*

$$\tilde{\mu}_w := \mu_w - \beta_w(\tilde{\rho}_w\alpha - \rho\alpha) > \mu_w \quad \forall\beta_w \in (\beta_b, 1]. \tag{1.13}$$

That is, there is a positive bias in favor of the socially dominant group.

The intuition behind this result is, in a sense, reverse to that of Proposition 1.1: Education is such a common trait for members of the socially dominant group that it does not get much attention in the updating process—i.e., the agent will underaccount for the role of education and, hence, develop an inflated view about the average productivity of members of the dominant group. This result is related to an idea from sociology and psychology, where it is acknowledged that socially dominant groups have privileges that they do not necessarily recognize (see, e.g., Collins 2018). It is well documented, for example, that discrimination against black Americans exists in the labor market, yielding them in a clear disadvantage compared to white Americans regarding many social and economic outcomes (Bertrand and Mullainathan 2004; for similar evidence from the Czech Republic and Germany, see Bartoš et al. 2016). Yet, the majority of white Americans do not believe that they benefit from any such preferential treatment (Pew Research Center 2017). In my model, because members of the dominant group are privileged to have a high share of educated individuals among them, they underaccount for education's importance and hence develop overly positive views about their group.

The result can also shed some light on a potential connection between beliefs about social groups and overconfidence. In the social psychology literature, it is often argued that members of a social group reflect traits of that group onto themselves (e.g., Sidanius and Pratto 2001). An in-group bias is therefore understood as a means to keep up a positive self-view. Indeed, in my model, it is possible to assume that the agent has only imperfect knowledge about his own productivity. In this case, he may use information about the average productivity of other members of the dominant group to infer his own productivity. Since the logic of my model helps the agent to develop an inflated view about the average productivity of members of the dominant group, it will, in turn, also inflate the view about his own productivity. The agent will become overconfident.[15] Indeed, recent research provides experimental evidence of this connection between in-group bias and overconfidence (Flores and Fonseca 2022).

A similar link is analyzed in much more detail in Heidhues, Kőszegi, and Strack (2023b), although they use a logic that is reverse to mine. In their model, an overconfident agent learns about both the average productivity of other members of his social group and the extent to which society discriminates against his social group. Because he is overconfident, he can only explain the surprisingly low success he has in life if he believes that discrimination against his group takes place. If discrimination takes place, any negative life-outcomes of other members of his group are likely to be due to discrimination, too, which leads him to believe that the average "caliber" of the other group members is higher than it actually is—i.e., there is an in-group bias.

1.5 A Second Dominated Group

So far, the model was build around a society comprised of two social groups, the socially dominated and the socially dominant group. In real-world societies, however, there are typically more than two social groups. To account for this, I introduce a second dominated group. For example, one can think of the two dominated groups as immigrant groups from different regions of the world. Crucially, however, for simplification, I assume that

[15]We could come up with a simple formalization where the agent's prior belief is given by his belief about the group average, $\tilde{\mu}_w$. He then receives a signal about his own productivity according to $s_i \sim \mathcal{N}(a_i - \rho a_{e_i}, 1/\lambda_i)$. His posterior is then given by

$$\tilde{a}_i \sim \mathcal{N}\left(\frac{\tilde{\mu}_w \lambda_w + s_i \lambda_i}{\lambda_w + \lambda_i}, \frac{1}{\lambda_w + \lambda_i}\right),$$

where, fixing s_i, the mean is positively biased compared to the case where the agent would have correct beliefs about his social group.

the two groups are disjoint—i.e., a member of the first dominated group cannot be a member of the second dominated group. This assumption is plausible if we indeed think of the two groups as being immigrants from different regions or of a different cohort. It is less plausible if we think of one of the groups being based on immigration whereas the other is based on gender. This simplification precludes an analysis around the question of "intersectionality." In future work, I plan to relax this assumption.

To facilitate the analysis of a society with two dominated groups, I introduce some additional notation. The *old* dominated group will now be denoted b_1, whereas the *new* dominated group will be denoted b_2. As before, members of either dominated group can be educated or uneducated. This extends the set of possible types to be given by $\Theta^+ := \{b_1, b_2, w\} \times \{e, u\}$.

With two dominated groups in the society, we also need to consider the relative group sizes. The population share of the old dominated group is denoted by α_1. The population share of the new dominated group is denoted by α_2. Correspondingly, the population share of the dominant group is given by $1 - \alpha_1 - \alpha_2$. Additionally, the share of educated members of the old dominated group will now be denoted by β_{b_1}. The share of educated members of the new dominated group will be denoted by β_{b_2}. All other new notation will follow the same pattern of indexation—i.e., the means and variances of the respective intrinsic productivity distributions will be denoted μ_{b_1}, μ_{b_2}, $1/\lambda_{b_1}$, $1/\lambda_{b_2}$, and so on.

Additionally, the old definition of what determines how representative education is among a social group—specified in expression (1.7)—needs to be extended to allow for more than two social groups.[16] Specifically, I now assume that the representativeness of education for either of the dominated groups is given according to the following definition.

Definition 1.3 (Representativeness with Multiple Dominated Groups)**.** *The representativeness of education for dominated group k compared to the rest of the population is given by the likelihood ratio*

$$R_{b_k} := \frac{\beta_{b_k}}{\frac{1-\alpha_1-\alpha_2}{1-\alpha_k}\beta_w + \frac{\alpha_{h\neq k}}{1-\alpha_k}\beta_{b_{h\neq k}}} \quad \forall k, h \in \{1, 2\}. \tag{1.14}$$

That is, education is more representative of dominated group k the higher the likelihood ratio in expression (1.14) is, and vice versa.

The only difference between the new likelihood ratio in expression (1.14) and the old likelihood ration in expression (1.7) is that, in (1.7), represen-

[16] And indeed, the following definition can be readily extended to the case with even more groups.

tativeness was determined only by a comparison of the share of educated members of the single dominated group and the share of educated members of the dominant group. Now, instead, it is determined by a comparison of the share of educated members of one of the dominated groups and the average of the share of educated members of the dominant group and the other dominated group.[17] This specification can reflect that, when one dominated group has a low share of educated individuals, the relative share of educated individuals in the other dominated group compared to *all* other groups increases and, hence, education becomes a less salient feature among that group.

Suppose that the share of educated individuals among the new dominated group is lower than among the dominant group—i.e., $\beta_{b_2} < \beta_w$. This decreases the average share of educated individuals in the overall population compared to the case without the new dominated group. As a result, the low share of educated individuals among the old dominated group is less salient. Hence, it will be weighted less in the updating process than in the case with only two social groups. This implies that stereotyping against the old dominated group decreases. The following proposition summarizes this.

Proposition 1.3 (A New Dominated Group)**.** *Suppose that* $\beta_{b_k} < \beta_w \ \forall k \in \{1,2\}$*. As* $t \to \infty$*, the agent's belief about the average of dominated group k's intrinsic productivity parameter converges in distribution to the value*

$$\tilde{\mu}_{b_k} := \mu_{b_k} - \beta_{b_k}(\tilde{\rho}_{b_k}\alpha - \rho_{b_k}\alpha) < \mu_{b_k} \quad \forall \beta_{b_k} \in (0, \beta_w). \tag{1.15}$$

(i) We have that $\tilde{\mu}_{b_1} \in (\tilde{\mu}_b, \mu_b)$*—i.e., the extent of the negative stereotype against the old dominated group is lower than in the case without the new dominated group. (ii) Suppose that* $\mu_{b_1} = \mu_{b_2}$ *and* $\lambda_{b_1} = \lambda_{b_2}$*. We have that* $\tilde{\mu}_{b_1} > \tilde{\mu}_{b_2}$ *if and only if* $\beta_{b_1} > \beta_{b_2}$ *and vice versa.*

Let us put the result into perspective with the help of some empirical patterns. Proposition 1.3 states that when a new dominated group enters the society, this decreases the extent of stereotyping against the old dominated group (part (i)). This is in line with empirical observations. For example, Fouka, Mazumder, and Tabellini (2022) show that the migration of black Americans into the northern part of the United States reduced the

[17]This is of course not the only way one could formalize the representativeness of education in this case. For example, one could argue that the dominant group compares the dominated groups only to itself and not to both itself and the other dominated group. However, the specification I suggest captures the empirical pattern that the introduction of an even more distant option reduces the perceived distance between some baseline and a less different option.

stereotyping against Irish and Italian immigrants that was previously strong. Instead, stereotyping concentrated around black Americans afterwards. In my model, this is the case if the share of educated individuals among the new dominated group is smaller than among the old dominated group (part (ii) of Proposition 1.3). Indeed, the logic for which Fouka, Mazumder, and Tabellini (2022) argue is similar to the one I explore: The introduction of black Americans into previously whiter areas increased the *perceived* similarity between long-time residents and the older immigrants, reducing the extent of stereotyping they faced. This logic is one possible interpretation of the formalization of representativeness in expression (1.14).

1.6 Affirmative Action

We remain in the framework with two dominated groups, as introduced in section 1.5. In this section, I look at the effect of an affirmative action policy. Specifically, I consider a policy that increases the share of educated individuals among the old dominated group, while it leaves the share of educated individuals among both other groups unaffected.[18] We could interpret the policy as a (successful) attempt to provide better education specifically aimed at the old dominated group (or, equivalently, aimed at both dominated groups but only successful for the first—e.g., due to asymmetric language or cultural barriers).[19]

Looking at expression (1.14), it is easy to see that such a policy will make education a more representative trait of the old dominated group vis-a-vis both other groups. At the same time, it makes education a less representative trait among the new dominated group. In my framework, this implies that the impact of education may be more or less over- or underweighted when updating about the old dominated group's productivity compared to the case without affirmative action, based on how high the share of educated people among the old dominated group was to begin with—this follows from Corollary 2. At the same time, it implies that the impact of education will always be more overweighted when updating about the new dominated

[18]One may argue that an increase in the share of educated individuals for one group may go hand in hand with a decrease in the share of educated individuals for the other groups. I abstract away from this for simplicity.The results would be qualitatively unchanged from the ones stated in this section.

[19]I analyze this extreme case where the new dominated group's share of educated people is not affected by the affirmative action policy at all for clarity. The results carry over to the more general case, however, as long as the new dominated group is affected *less* than the old dominated group (or vice versa).

group's productivity compared to the case without affirmative action. This is because for the new dominated group, the increase of β_{b_1} affects only the representativeness of educated members of their group, while leaving the *share* of educated individuals of their group from which the dominant group learns unaffected. The following two propositions summarize the effects. The first proposition holds also for the model with only one dominated group.

Proposition 1.4 (Affirmative Action and the Old Dominated Group). *Suppose that a policy maker implements an affirmative action policy that increases β_{b_1} while it leaves all other shares unaffected. Then, as $t \to \infty$, the agent's belief about the average of old dominated group's intrinsic productivity parameter converges in distribution to some value $\tilde{\mu}_{b_1}^{aa}$, about which we can say the following*

1. *Suppose the original share of educated individuals is relatively low, $\beta_{b_1} < \hat{\beta}_{b_1}$.*
   a) *If the increase of β_{b_1} is sufficiently small, the long-run belief $\tilde{\mu}_{b_1}^{aa}$ is more negatively biased compared to the case without affirmative action.*
   b) *If the increase of β_{b_1} is sufficiently large, the long-run belief $\tilde{\mu}_{b_1}^{aa}$ is less negatively biased compared to the case without affirmative action.*
2. *Suppose the original share of educated individuals is relatively high, $\beta_{b_1} > \hat{\beta}_{b_1}$. Then, the long-run belief $\tilde{\mu}_{b_1}^{aa}$ is less negatively biased compared to the case without affirmative action.*

Proposition 1.5 (Unintended Consequences of Affirmative Action for the New Dominated Group). *Again, suppose that a policy maker implements an affirmative action policy that increases β_{b_1} while it leaves all other shares unaffected. Then, as $t \to \infty$, the agent's belief about the average of the new dominated group's intrinsic productivity parameter converges in distribution to some value $\tilde{\mu}_{b_2}^{aa}$. This value is lower compared to the case without affirmative action.*

Proposition 1.4 and Proposition 1.5 together imply that while an affirmative action policy targeted at one dominated group can improve that group's standing in society, it can harm the standing of another dominated group. With this result, I contribute to the literature that discusses the effects of affirmative action and, in particular, its unintended consequences (see, e.g., Coate and Loury 1993, Fryer Jr and Loury 2005, or Fershtman and Pavan

2021). The result highlights how acknowledging the co-existence of multiple dominated groups can be important to evaluate the overall benefits of an affirmative action policy in a specific context. Such spillovers have, to the best of my knowledge, not been analyzed previously.

Additionally, affirmative action can also decrease the in-group bias if the agent learns also about his own group. The next remark summarizes this.

Remark 1.3 (Affirmative Action and In-Group Bias)**.** *Suppose that the agent also learns about the average productivity of his own group and that there is an affirmative action policy targeting either one or both of the dominated groups. His in-group bias will be lower than in the case without affirmative action.*

Importantly, Remark 1.3 implies that while an affirmative action policy may have the unintended consequence of increasing stereotyping against the dominated group that was not targeted, it can have the (potentially) positive effect of debiasing the agent's view about his own group. This effect can—for example, in the labor market—help either of the dominated groups.

1.7 Legitimizing Myths

In this section, I first discuss the previous results in light of the concept of legitimizing myths. I further provide a formal extension of the model used in the previous sections to highlight the main points of the discussion. This helps to understand how taste-based, statistical, and inaccurate statistical discrimination can be intergenerationally related.

1.7.1 Discussion

In the previous sections, I have discussed how a salience-based learning model can lead to negative stereotypes about a dominated group and, beyond this, explain several stylized facts and make new predictions. In this section, I want to discuss these results in light of the concept of legitimizing myths (see Sidanius and Pratto 2001). Legitimizing myths are, roughly speaking, commonly held beliefs that help to justify a specific social order—e.g., discrimination based on race or gender.

One crucial assumption of the model is that the share of educated individuals among the dominated group is lower than among the dominant group, or that the agent at least thinks that this is the case. This creates the salience effects, which drive the results in this paper. More specifically, members from the dominant group overestimate their group's average intrinsic productivity, whereas they underestimate the dominated group's average intrinsic productivity.

In this sense, my paper is only about beliefs, rather than discriminatory actions. However, we can discuss the results while keeping more active forms of discrimination in mind; in principle, the agent might be opposed to discrimination based on taste. But the overly negative beliefs he develops about the dominated group can help him to justify why discrimination takes place and, hence, provides an argument to him why discrimination is fair or justified. For example, the agent might be aware that it is harder for specific groups to be admitted to a university and, hence, receive a high level of education. However, given his negative beliefs, he can justify this, and, if one would extend the model to allow for actions, might participate in discrimination himself.

A similar argument can be used to make a connection between historic discrimination and present day stereotypes. Indeed, we can interpret the model as suggesting that historic discrimination may be a reason for present day stereotyping. Intuitively, historic discrimination is one of the reasons why some groups are underrepresented in schools, universities, and prestigious jobs, or otherwise have worse life-outcomes (see, e.g., Rubio 2022). This under-representation, in my model, leads to a salient trait when corresponding individuals are sampled and, hence, is the root of the negative stereotype. Following the above logic, this provides the agent with a reason to uphold discrimination. Discrimination is, hence, a vicious cycle that creates negative stereotypes, which create discrimination, which creates negative stereotypes, and so on.

1.7.2 Extension

I now provide a simple formal argument of the interplay between stereotypes and discrimination highlighted in the discussion above. Again, the model has stayed silent about the precise reasons for why we would expect a low share of members of the dominated group among the educated people, even though I have previously hinted at some possible explanations: Assume that members of the dominant group are capable of deciding which other person will *become* educated—for example, because they are on the admission committee of a university or school. Then, if these people are negatively biased against the dominated group, they will be less likely to select a member from the dominated group. This, then, leads to a lower share of the dominated group among the highly educated and, hence, implies that beliefs among future generations will be negatively biased against the dominated group, too. While analyzing actions is not the main purpose of my paper, in this section, I present a simple model with overlapping generations to illustrate the above-described dynamic.

Suppose that each member of a continuum of the dominant group lives for three periods. In the first period, they are an *observer*, and learn about the average productivity of the dominated group according to the rules laid out so far in the paper—i.e., their belief at the end of period 1 is given by expression (1.12). In the second period, they become *selectors* and observe two additional signals: one about the intrinsic productivity of a randomly drawn member of the dominated group and one about the intrinsic productivity of a randomly drawn member of the dominant group. We can think of these two individuals as *applicants* to a university. The selector selects one of the two applicants to admit them to university in the third period of his life (afterwards he dies). During the third period of his life, a new generation is born and starts their life as observers. They observe the dominated group based on how the previous generation of the dominant group sorted them into educated and uneducated. Afterwards they become selectors themselves, and so on. Each generation of selectors is naive: They ignore *how* the individuals they observe have been sorted into the ordinary and exceptional groups.[20] An individual generation is denoted $\tau = 1, 2, 3, ...$

The dominated applicant and the dominant applicant generate the signals

$$s_b = a_b + \epsilon_b, \quad \text{and} \quad s_w = a_w + \epsilon_w, \tag{1.16}$$

respectively. Hereby, the distribution of both ϵ_b and ϵ_w is the same, $\epsilon_b, \epsilon_w \sim \mathcal{N}(0, 1/\lambda)$. Because both a_b and a_w are random variables, from the perspective of the selector, the signals are distributed according to $s_b \sim \mathcal{N}(\tilde{\mu}_b, 1/\lambda + 1/\lambda_b)$ and $s_w \sim \mathcal{N}(\mu_w, 1/\lambda + 1/\lambda_w)$. After receiving the signals, they form a posterior belief about a specific individual. Let the posterior means of each individual's productivity be denoted by $\tilde{a}_b$ and $\tilde{a}_w$, respectively.

I allow for taste-based discrimination of the dominant group against the dominated group, captured by the discrimination parameter $d \in \mathbb{R}_+$. Specifically, I assume that the selector selects a member of the dominated group over a member of the dominant group if and only if

$$\tilde{a}_b - d > \tilde{a}_w.$$

[20]The case where the dominant group is not naive and can take into account the sorting is, in principle, captured by the generalization in subsection 1.A.2. However, this generalization does not allow for a closed-form solution, which is important to derive the comparative statics this section is built around.

Before I state any results related to the above extension, the following definition regarding the term *inaccurate statistical discrimination* is useful:

Definition 1.4 (Inaccurate Statistical Discrimination)**.** *In any period* $\tau = 1, 2, 3, ...$*, there is inaccurate statistical discrimination if for an arbitrary member of the dominated group with intrinsic productivity* a_b^i*, the belief* $\tilde{\mu}_b^\tau$ *is such that the probability of being selected over an arbitrary member of the dominant group with the same intrinsic productivity,* $a_w^j = a_b^i$*, is lower than if the belief would be correct,* $\tilde{\mu}_b^\tau = \mu_b$*.*

Therefore, a sufficient condition for inaccurate statistical discrimination to occur in generation τ is that there is a negative stereotype against the dominated group, $\tilde{\mu}_b^\tau < \mu_b$, whereby $\tilde{\mu}_b$ denotes generation τ's belief about the dominated group's mean intrinsic productivity. With this in mind, I can state my next result.

Proposition 1.6 (Stereotypes and Inaccurate Statistical Discrimination)**.** *A sufficient condition for there to be negative stereotyping (and therefore inaccurate statistical discrimination) against members of the dominated group in any generation* $\tau = 2, 3, ...$ *is that the first generation of the dominant group hires more members of the dominant group than of the dominated group.*

Proposition 1.6 states that if the first generation of the dominant group selects less members of the dominated group than of the dominant group, there will be both stereotyping and inaccurate statistical discrimination against the dominated group in any subsequent generation. The proposition stays silent about the reasons why the first generation may have such a bias in selecting. Possibilities are that the first generation participates in taste-based discrimination, accurate statistical discrimination, or already in inaccurate statistical discrimination. The next proposition investigates the inter-generational relationship between these different forms of discrimination.

Proposition 1.7 (Statistical, Taste-Based, and Inaccurate Statistical Discrimination)**.** *Suppose that in the first generation,* $\tau = 1$*, the members of the dominated group are either subjected to taste-based discrimination* ($d > 0$) *or accurate statistical discrimination* ($\tilde{\mu}_b^{\tau=1} = \mu_b < \mu_w$)*. Then, any subsequent generation of the dominant group will hold a negative stereotype against the dominated group (and therefore participate in inaccurate statistical discrimination).*

Proposition 1.7 makes a novel connection between the two classical forms of discrimination considered in the economics literature (taste-based and statistical discrimination) and the more recently considered type of discrimination, inaccurate statistical discrimination, which is the focus of this paper. In my model, inaccurate statistical discrimination emerges because the socially dominated group is underrepresented among the group of educated people. As Proposition 7 shows, it does not matter *why* they are underrepresented; the reasons for the underrepresentation may very well be taste-based or statistical discrimination. However, in my model, any such classical form of discrimination will, through its negative effect on the share of the dominated group among the educated, lead to inaccurate statistical discrimination against the dominated group by *any* subsequent generation of the dominant group. In this sense, this simple extension of my main model illustrates how historic discrimination, can lead to stereotyping in present periods, which, in turn, helps to uphold the discriminatory systems established in the past. Discrimination and stereotyping, hence create a vicious cycle, or a feedback loop, that is hard to interrupt.

1.8 Related Literature

In this paper, I have provided an argument for how biased beliefs about social groups can emerge and persist, and lead to discrimination. Individuals are subject to a biased learning rule that is based on how salient education appears among a given social group. The paper is therefore mainly related to three strands of literature, (i) discrimination, (ii) biased learning, and (iii) salience effects.

Discrimination. As discussed briefly in the introduction, the economics literature on discrimination can roughly be separated in statistical discrimination (see, e.g., Phelps 1972; Arrow 1973; Aigner and Cain 1977; Morgan and Várdy 2009) and taste-based discrimination (see, e.g., Becker 1957; Prendergast and Topel 1996; Lagerlöf 2020; Pikulina and Ferreira 2023). More recently, however, a third type of discrimination has been getting more attention, *inaccurate statistical discrimination* (Bordalo et al. 2016; Sarsons 2017; Bohren et al. 2019; Mengel and Campos Mercade 2021; Barron et al. 2022; Heidhues, Kőszegi, and Strack 2023b; Bohren, Imas, and Rosenberg 2019; Sarsons et al. 2021; Siniscalchi and Veronesi 2021). This literature deviates from the earlier economics literature in that it posits that discrimination can be based on incorrect beliefs.

This idea relates the respective literature to the corresponding social psychology literature. Here, advocates of the social-cognitive approach to stereotyping (see Allport, Clark, and Pettigrew 1954), have argued that stereotypes are the result of, in principle, innocuous but overly-simplified and thereby incorrect information processing rules. Hamilton and Gifford (1976), for example, provide evidence that people stereotype others because they learn in an associative way—i.e., they link unusually negative traits and relatively unusual people (e.g., minorities) together. Similarly, Levine and Campbell (1972) argue that when groups are overrepresented among certain roles in society, people tend to erroneously believe that all members of the group take on that specific role. In a sense, the Bordalo et al. (2016) model of stereotyping follows a similar logic in that the frequency with which members of a group take on roles among which they are overrepresented are overestimated when a person recalls the distribution of types. Snyder and Uranowitz (1978) provide evidence on how biased recall of information can affect stereotyping. It is natural to conjecture that if present stereotyping leads individuals to recall certain information, this recalled information will also be incorporated in the individuals' posteriors about the stereotyped group, and therefore learning will occur in a biased way.

Biased Learning. Overall, the above mentioned references suggest that stereotyping is due to biased (or otherwise incorrect) processing of the information individuals receive about others. This is true in my model as well. As such, the paper is also related to the literature on biased learning. Typically, papers in this literature assume that individuals receive biased signals but do not take the bias into account when updating. Some research is further complicated by modeling agents who can influence the signals they receive through their actions. See, for example, Esponda and Pouzo (2016), Fudenberg, Romanyuk, and Strack (2017), Heidhues, Kőszegi, and Strack (2021), or Heidhues, Kőszegi, and Strack (2023a). To consider a more concrete example, Heidhues, Kőszegi, and Strack (2018), model an overconfident agent who manages a team and does not know the other team members' abilities. Due to his overconfidence, the surprisingly low output he observes is misattributed to the abilities of his team members. The abilities of his team members are therefore underestimated by the agent. As a response, he takes on more control of the team, further worsening its performance and leading him away from learning the true value of the fundamental. Such models have also been applied to stereotyping. Hübert and Little (2023), for example, provide a model in which police officers police members of one subpopulation inefficiently much. As a result, they detect more crime among that group, leading them to believe that members of this group are more criminal than others. This leads to even more policing and the beliefs grow increasingly negative. Other models[21] of biased learning do not necessarily consider actions—for example, Bohren and Hauser (2021), Bohren (2016),[22] Schwartzstein (2014), Fryer Jr and Jackson (2008), or Fryer Jr, Harms, and Jackson (2019). Such models, too have been applied to stereotyping. Heidhues, Kőszegi, and Strack (2023b), for example, show how overconfidence together with uncertainty about the extent and direction of discrimination can lead individuals to underestimate other groups average productivity while overestimating their own. These results are similar to the ones obtained in my paper. While I do not allow the agent to affect the signals he receives through his own actions, the main contribution of my paper from the perspective of the biased learning literature is that I endogenize the bias, which allows to explain several stylized facts in one unified framework.

Salience Effects. More precisely, I endogenize the bias based on a salience function. The paper is therefore also related to the literature on salience effects. See, for example, Bordalo, Gennaioli, and Shleifer (2012b), Bordalo, Gennaioli, and Shleifer (2012a), Bordalo, Gennaioli, and Shleifer (2013b), Bordalo, Gennaioli, and Shleifer (2013a), or Dertwinkel-Kalt and Köster (2020). In my model, most importantly, this implies that learning is potentially biased and the direction of the bias, as well as the frequency with which it occurs, is context dependent. While some parameter may be overweight in the evaluation of one group, it can be underweight in the evaluation of another; and indeed, there is experimental evidence that depending on the context, salience effects can lead an agent to either under- or overvalue relevant traits (Bushong, Rabin, and Schwartzstein 2021). Overall this logic is analogous to the other papers on salience effects. The main difference is that most of these papers are interested in how salience affects economic choice, whereas in my paper, salience affects the processing of information. In the end, however, as outlined in section 1.7, any biased belief will ultimately also affect people's actions. Hence, my model can also be interpreted as one where an agent's actions are affected through salience—even though it requires the intermediate step of forming biased beliefs.

[21]And others are general enough to capture both classes of model-misspecification—e.g., Frick, Iijima, and Ishii (2023).

[22]In this paper, learning depends on other people's actions but not on those of the agent.

1.9 Conclusion

The paper has shown how incorporating a salience-based updating rule into a normal Bayesian learning model can account for several stylized facts and make additional predictions about the nature of stereotyping. As such, the model can provide a unified account for several patterns that have not previously been connected in the same sense. Yet, the model makes several simplifying assumptions, some of which I relax in subsection 1.A.2, where I consider more general productivity distributions. Other simplifications, I plan to relax in ongoing work. For example, I only allow for the coexistence of several dominated groups that do not overlap. In reality, however, it is well possible for an individual to be marginalized through different dimensions—e.g., through both race and gender. Incorporating this could lead to further, economically interesting results.

Another potentially interesting extension is regarding the question how stereotypes can interact with people's beliefs in discrimination. Suppose the agent wants to learn the reason for why the share of educated individuals among the dominated group is lower than among the dominant group. Suppose further that the agent entertains two possible explanations: a lower intrinsic productivity vs. historic discrimination. Suppose that the true reason is historic discrimination. The question is now whether he can correctly learn that the low share of educated individuals is due to discrimination. Intuitively, because of his biased learning, his belief about the average intrinsic productivity becomes more and more negatively biased and, hence, over time, the agent believes less and less that discrimination is a decisive factor. Hence, over time, generally, he comes to believe less and less in discrimination, which is in line with empirical patterns. For example, whereas 11 percent of white Americans belief that there is no racism against black people in the US, only 3 percent of black Americans agree (Pew Research Center 2022). These numbers clearly suggest bias in the evaluation of the extent of discrimination against black people in the US.

1.A Appendix

1.A.1 Proofs

Proof of Proposition 1.1

We have a misspecified learning model. We know from Berk (1966) that, as $t \to \infty$, the agent's long-run beliefs will concentrate on the set of minimizers of the Kullback-Leibler divergence between the true statistical model of the world and the agent's misspecified one. In my model, the agent does not sample from only one distribution but rather from two. The distribution from which he samples depends on whether he is sampling an educated or an uneducated member of dominated group. In this case, we can write the Kullback-Leibler divergence as the weighted sum—weighted based on how often the agent samples from each of the two distributions relative to the other—of the Kullback-Leibler divergences of the two different distributions. To proceed, I introduce some additional notation. Let us call the conditional distribution of productivity for an educated member of the dominated group $G_{b,e}$. Similarly, denote the conditional productivity distribution of an uneducated member of the dominated group $G_{b,u}$. I denote the distributions' probability density functions by $g_{b,e}$ and $g_{b,u}$, respectively. The CDF and PDF of what the agent believes the signals of educated members of the dominated group to be drawn are denoted $\tilde{G}_{b,e}$ and $\tilde{g}_{b,e}$, respectively. All parameters of the misspecified distribution will be denoted by a "tilde" as well. All of these distributions, including the misspecified ones, are normal. Furthermore, notice that the Kullback-Leibler divergence between two normal distributions $h_o \sim \mathcal{N}(\mu_0, \sigma_0^2)$ and $h_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ is given by[23]

$$D(h_0 \mid\mid h_1) = \frac{1}{2}\left[\frac{\sigma_0^2}{\sigma_1^2} + \frac{(\mu_1 - \mu_0)^2}{\sigma_1^2} - 1 + \ln\frac{\sigma_1^2}{\sigma_0^2}\right].$$

[23]See, for example, https://en.wikipedia.org/wiki/Normal_distribution.

Applying this to our problem, we know that we need to find the minimizer $\tilde{\mu}_m$ over μ_m^* of

$$\begin{aligned}
D((1-&\beta_b)g_{b,u} + \beta_b g_{b,e} \mid\mid \ (1-\beta_b)g_{b,u} + \beta_b \tilde{g}_{b,e}) \\
&= \frac{(1-\beta_b)}{2}\left[\frac{\sigma_{b,u}^2}{\sigma_{b,u}^2} + \frac{(\mu_b^* - \mu_b)^2}{\sigma_{b,u}^2} - 1 + \ln\frac{\sigma_{b,u}^2}{\sigma_{b,u}^2}\right] \\
&\quad + \frac{\beta_b}{2}\left[\frac{\sigma_{b,e}^2}{\tilde{\sigma}_{b,e}^2} + \frac{(\mu_b^* + \tilde{\rho}_b\alpha - \mu_b - \rho\alpha)^2}{\tilde{\sigma}_{b,e}^2} - 1 + \ln\frac{\tilde{\sigma}_{b,e}^2}{\sigma_{b,e}^2}\right] \\
&= \frac{(1-\beta_b)}{2}\left[\frac{(\mu_b^* - \mu_b)^2}{\sigma_{b,u}^2}\right] \\
&\quad + \frac{\beta_b}{2}\left[\frac{\sigma_{b,e}^2}{\tilde{\sigma}_{b,e}^2} + \frac{(\mu_b^* + \tilde{\rho}_b\alpha - \mu_b - \rho\alpha)^2}{\tilde{\sigma}_{b,e}^2} - 1 + \ln\frac{\tilde{\sigma}_{b,e}^2}{\sigma_{b,e}^2}\right].
\end{aligned}$$

We can rewrite this term by replacing the variances with the inverses of the corresponding precision parameters. This yields

$$\begin{aligned}
D((1-\beta_b)g_{b,u}+&\beta_b g_{b,e} \mid\mid \ (1-\beta_b)g_{b,u} + \beta_b \tilde{g}_{b,e}) \\
&= \frac{(1-\beta_b)}{2}\lambda_b(\mu_b^* - \mu_b)^2 \\
&\quad + \frac{\beta_b}{2}\left[\frac{\lambda_b}{\lambda_b} + \lambda_b(\mu_b^* + \tilde{\rho}_b\alpha - \mu_b - \rho\alpha)^2 - 1 + \ln\frac{\lambda_b}{\lambda_b}\right] \\
&= \frac{(1-\beta_b)}{2}\lambda_b(\mu_b^* - \mu_b)^2 + \frac{\beta_b}{2}\left[\lambda_b(\mu_b^* + \tilde{\rho}_b\alpha - \mu_b - \rho\alpha)^2\right].
\end{aligned}$$

To minimize this expression, we take the derivative with respect to μ_b^* and set it equal to zero. That is, the first order condition is:

$$\begin{aligned}
\frac{\partial D(\cdot \mid\mid \cdot)}{\partial \mu_b^*} = 0 &\Leftrightarrow (1-\beta_b)\lambda_b(\mu_b^* - \mu_b) + \beta_b\lambda_b(\mu_b^* + \tilde{\rho}_b\alpha - \mu_b - \rho\alpha) = 0 \\
&\Leftrightarrow \lambda_b(\mu_b^* - \mu_b) + \beta_b\lambda_m(\tilde{\rho}_b\alpha - \rho\alpha) = 0
\end{aligned}$$

Solving this for μ_b^* yields

$$\mu_b^* := \tilde{\mu}_b = \frac{\lambda_b\mu_b - \beta_b\lambda_b(\tilde{\rho}_b\alpha - \rho\alpha)}{\lambda_b} = \mu_b - \beta_b(\tilde{\rho}_b\alpha - \rho\alpha),$$

which is equivalent to expression (1.12) as stated in Proposition 1. Similarly, we have that $\tilde{\mu}_b < \mu_b$ because $\beta_b(\tilde{\rho}_b\alpha - \rho_b\alpha) > 0$—which is because, for $\beta_b < \beta_w$, we have that $\tilde{\rho}_b > \rho_b$. Hence, we only need to show that the solution is indeed a minimum. We get the second order derivative by taking the second derivative of the respective term. This yields

$$\frac{\partial^2 D(\cdot \mid\mid \cdot)}{\partial \mu_b^{*2}} = \lambda_b > 0.$$

The Kullback-Leibler divergence is therefore strictly convex and, hence, we have indeed found a minimum. □

Proof of Corollary 1.1

We are interested in the effect on an increase of the share of educated members of the dominant group on the long-run beliefs in expression (1.12). It suffices to show that

$$\frac{d\tilde{\mu}_b}{d\beta_w} = -\beta_b \alpha \frac{\partial \tilde{\rho}_b}{\partial \beta_w} < 0.$$

This is indeed the case because, by Definition 1 together with Assumption 1, we have that $\partial \tilde{\rho}_b / \partial \beta_w > 0$. □

Proof of Corollary 1.2

Consider expression (1.12). We want to show that there exists a non-empty set of values of β_b such that marginally below any value in that set, $\tilde{\mu}_b$ is increasing in β_w, whereas marginally above any such value, the opposite is true.[24] This statement follows from the combination of several observations. First, let us consider the extreme cases—i.e., the lowest and highest values that β_b can take on—and evaluate the agent's long-run bias at these values. The lowest possible value is $\beta_b = 0$. If we evaluate $\tilde{\mu}_b$ at this point, we can see that it is equal to μ_b—i.e., the long-run belief is unbiased in this case. This is because in this scenario, the agent samples only from the distribution he evaluates in an unbiased way. The highest possible value is $\beta_b = \beta_w$. In this case, too, we have that the long-run belief is unbiased, $\tilde{\mu}_b = \mu_b$. This is because in this case, the exceptional trait is neither particularly representative of the dominant group or the dominated group and, hence, any signal is evaluated in an unbiased way. By Proposition 1, we know that for any value $\beta_b \in (0, \beta_w)$, the long-run belief will be negatively biased. Together with continuity, this implies that there must be points on $(0, \beta_w)$

[24]This is because Assumption 1 which specifies the salience-based weighting of education in the updating process is quite general. If we would assume a specific functional form, we could simply take the derivative.

for which $\tilde{\mu}_b$ is either decreasing, increasing, or unaffected by a change in β_b. Denote the set of the points β_b for which $d\tilde{\mu}_b/d\beta_b = 0$ by A.

we want to establish that there is only one point on $(0, \beta_w)$ for which $d\tilde{\mu}_b/d\beta_b = 0$—i.e., that A is a singleton. We can do this by taking the total derivative of $\tilde{\mu}_b$ with respect to β_b and setting it equal to zero,

$$\frac{d\tilde{\mu}_b}{d\beta_b} = \underbrace{\frac{\partial\tilde{\mu}_b}{\partial\tilde{\rho}_b}\frac{d\tilde{\rho}_b}{d\beta_b}}_{\text{indirect effect}} + \underbrace{\frac{\partial\tilde{\mu}_b}{\partial\beta_b}\frac{d\beta_b}{d\beta_b}}_{\text{direct effect}} = 0. \tag{1.17}$$

Any point β_b in A must satisfy condition (1.17). By explicitly looking at the different derivatives in the expression, we can see that the indirect effect is given by

$$\frac{\partial\tilde{\mu}_b}{\partial\tilde{\rho}_b}\frac{d\tilde{\rho}_b}{d\beta_b} = \underbrace{-\beta_b\alpha}_{<0}\underbrace{\frac{d\tilde{\rho}_b}{d\beta_b}}_{<0} > 0,$$

whereas the direct effect is given

$$\frac{\partial\tilde{\mu}_b}{\partial\beta_b}\frac{d\beta_b}{d\beta_b} = -\underbrace{(\tilde{\rho}_b\alpha - \rho\alpha)}_{>0} < 0.$$

This means that both the indirect and the direct effect of β_b on $\tilde{\mu}_b$ are monotonic and go in opposing directions. This implies that there is at most one point satisfying the condition that (1.17) is equal to zero. Indeed, because we know from the previous step of this proof that the set A is non-empty, this implies that it is a singleton.

Taken together, the twp steps of the proof imply that the indirect and the direct effect of an increase of β_b on $\tilde{\mu}_b$ have the same strength at exactly one point on $(0, \beta_w)$. This point is the threshold given in the corollary, $\hat{\beta}_b$. □

Proof of Remark 1.1

Follows immediately from observing that the prior, as in standard Bayesian inference problems, does not play a role for the long-run belief. □

Proof of Corollary 1.3

Suppose that the agent samples from ordinary and exceptional members of the dominated group according to their true shares—i.e., according to β_b. However, he believes that the share of exceptional people among the dominated group is given by $\tilde{\beta}_b < \beta_b$. Then, this affects the expression in

(1.12) only through an increase in $\tilde{\rho}_b$. This follows from Definition 1 together with Assumption 1. Specifically, we have

$$\frac{\partial}{\partial \tilde{\rho}_b} \left[\mu_b - \beta_b(\tilde{\rho}_b \alpha - \rho\alpha) \right] < 0, \tag{1.18}$$

which implies that the effect of an increase of $\tilde{\rho}_b$ is an increase of the extent of the negative stereotype. □

Proof of Corollary 1.4

Notice that in the scenario described in Corollary 4, the true share of educated people among the dominated group does not change—i.e., β_b is fixed—while the share of educated people among the subset of the dominated group from which the agent samples increases from β_b to $\check{\beta}_b$. This implies that, in expression (1.12), $\tilde{\rho}_b$ is fixed, whereas β_b increases. Again, we can then see that

$$\frac{\partial \tilde{\mu}_b}{\partial \beta_b} = -(\tilde{\rho}_b \alpha - \rho\alpha) < 0,$$

because $\tilde{\rho}_b > \rho$. This means that an increase in β_b leads to a decrease of $\tilde{\mu}_b$—i.e., the extent of the negative stereotype increases. □

Proof of Proposition 1.2

The proof of Proposition 2 is analogous to that of Proposition 1. We simply have to replace β_b, μ_b, and λ_b in expression (1.12) with the respective terms for the dominant group to obtain

$$\tilde{\mu}_w = \mu_w - \beta_w(\tilde{\rho}_w \alpha - \rho\alpha).$$

We know that $\tilde{\mu}_w > \mu_w$ because when $\beta_w > \beta_b$—which is the case discussed in the proposition—we have that $\tilde{\rho}_w < \rho$. □

Proof of Proposition 1.3

If we have that $\beta_{b_2} < \beta_w$, the average share of exceptional individuals over all groups is lowered through the entry of the second dominated group. By Definition 2, this makes education more representative of the old dominated group than it was before the new dominated group entered the society. Then the effect of the entry of the new dominated group simply corresponds to a decrease of $\tilde{\rho}_{b_1}$ as compared to $\tilde{\rho}_b$ in expression (1.12), which implies that the long-run belief $\tilde{\mu}_{b_1}$ must be higher than the long-run belief about the corresponding group's average intrinsic productivity parameter before the new dominated group entered entered, $\tilde{\mu}_b$. This proofs part (i) of the proposition. Part (ii) follows because we know that if $\beta_{b_1} > \beta_{b_2}$, then $\tilde{\rho}_{b_1} < \tilde{\rho}_{b_2}$, and hence learning will be more biased when the agent learns about the new dominated group. Hence, the extent of the stereotype will be larger for that group. In the case where $\beta_{b_1} < \beta_{b_2}$, the opposite is true. □

Proof of Proposition 1.4

If a policy maker introduces an affirmative action policy that corresponds to an increase of β_{b_1}, then education becomes more representative of the old dominated group according to Definition 2. Then, by Assumption 1, this implies that $\tilde{\rho}_{b_1}$ decreases compared to the case without affirmative action. This makes the bias smaller whenever the agent samples from the distribution he misinterprets—i.e., when he samples an educated member of the old dominated group. At the same time, he more often samples from a distribution he misinterprets. This corresponds to the exact same ambiguous effects described in Corollary 2, and, hence, whether he benefits are not from an affirmative action policy depends on whether the starting share of educated individuals among the old dominated group is above or below the threshold $\hat{\beta}_{b_1}$ and on how much the affirmative action policy increases this share. Consider case (1) in the proposition—i.e., the case where β_{b_1} is below the threshold. Below the threshold, an increase in β_{b_1} leads to a more negatively biased long-run belief (case (1a)). If the increase is so substantial, however, that the share of educated people is shifted sufficiently high over the threshold beyond which a further increase in β_{b_1} increases the long-run beliefs, then affirmative action will lead to less biased long-run beliefs (case (1b)).

If the share β_{b_1} is already above the threshold, then a further increase will lead to a lower extent of stereotyping and, hence, so will affirmative action (case (2)). This follows from the proof of Corollary 2. □

Proof of Proposition 1.5

For the new dominated group, if β_{b_1} increases while β_{b_2} is fixed, education becomes less representative of that group. This corresponds to an increase of $\tilde{\rho}_{b_2}$ compared to $\tilde{\rho}_b$ in expression (1.12)—i.e., the indirect effect is in play, while the direct effect is absent. As a result, the long-run beliefs about the new dominated group's average intrinsic productivity parameter will be more biased than without affirmative action aimed at old dominated group—i.e., affirmative action aimed at the old group will increase the extent of stereotyping the new group faces. □

Proof of Proposition 1.6

First, notice that whenever the (incorrect) model is normal—as is the case here—its mean is the minimizer of the Kullback-Leibler divergence when we want to estimate the mean of any other (correct) model. This is shown in Berk (1966). Hence, by our Proposition 1, having $\beta_b < \beta_w$ is sufficient to generate negative long-run stereotypes when the model is normal. I therefore proceed by establishing that we will have indeed $\beta_b < \beta_w$ in any period under the assumption stated in the proposition.

By standard Bayesian reasoning, after observing the two signals, the selector's posterior regarding the dominated applicant's expected productivity is given by

$$a_b \mid s_b \sim \mathcal{N}\left(\frac{\tilde{\mu}_b \lambda_b + s_b \lambda_{s_b}}{\lambda_b + \lambda_{s_b}}, \frac{1}{\lambda_b + \lambda_{s_b}}\right) \quad \text{with} \quad \lambda_{s_b} := \frac{1}{\frac{1}{\lambda} + \frac{1}{\lambda_b}}. \tag{1.19}$$

Analogously, the selector's posterior regarding the dominant applicant's expected productivity is given by

$$a_w \mid s_w \sim \mathcal{N}\left(\frac{\mu_w \lambda_w + s_w \lambda_{s_w}}{\lambda_w + \lambda_{s_w}}, \frac{1}{\lambda_w + \lambda_{s_w}}\right) \quad \text{with} \quad \lambda_{s_w} := \frac{1}{\frac{1}{\lambda} + \frac{1}{\lambda_w}}. \tag{1.20}$$

Recall that the selector may also participate in taste-based discrimination against the dominated group. This is captured by the taste-based discrimination parameter $d \in \mathbb{R}_+$. It enters the selector's utility such that he selects the dominated applicant over the dominant applicant if and only if

$$\frac{\tilde{\mu}_b \lambda_b + s_b \lambda_{s_b}}{\lambda_b + \lambda_{s_b}} - d > \frac{\mu_w \lambda_w + s_w \lambda_{s_w}}{\lambda_w + \lambda_{s_w}}. \tag{1.21}$$

That is, the dominated applicant's posterior expected productivity must outperform the dominant applicant's posterior expected productivity by a margin d for them to be selected. We can rewrite the selector's decision rule to obtain

$$s_b > \underbrace{\left[\frac{\mu_w \lambda_w + s_w \lambda_{s_w}}{\lambda_w + \lambda_{s_w}}(\lambda_b + \lambda_{s_b}) - \mu_b \lambda_b\right] \frac{1}{\lambda_{s_b}}}_{:=\Phi} + d\frac{\lambda_b + \lambda_{s_b}}{\lambda_{s_b}}, \tag{1.22}$$

where we know that s_w is a normally distributed random variable and, hence, so is Φ:

$$\Phi \sim \mathcal{N}\left(\left[\mu_w(\lambda_b + \lambda_{s_b}) - \mu_b \lambda_b\right] \frac{1}{\lambda_{s_b}}, \lambda_{s_w} \left(\frac{\lambda_b + \lambda_{s_b}}{\lambda_w + \lambda_{s_b}}\right)^2 \left(\frac{1}{\lambda_{s_b}}\right)^2\right). \tag{1.23}$$

To proceed, it is useful to define the random variable

$$\Delta s := s_b - \Phi. \tag{1.24}$$

Because both s_b and Φ are normal, Δs is normal too. Specifically, we have that

$$\Delta s \sim \mathcal{N}\left(\mu_b - [\mu_w(\lambda_b + \lambda_{s_b}) - \mu_b\lambda_b]\frac{1}{\lambda_{s_b}}, \lambda_{s_w}\left(\frac{\lambda_b + \lambda_{s_b}}{\lambda_w + \lambda_{s_b}}\right)^2\left(\frac{1}{\lambda_{s_b}}\right)^2 + \frac{1}{\lambda_{s_b}}\right). \tag{1.25}$$

I denote the corresponding cumulative distribution function of Δs by $H(\cdot)$. More concretely, because $H(\cdot)$ is a function that depends on parameters which may differ across generations τ, I write $H(\cdot)$ with the parametrization from generation τ as $H_\tau(\cdot)$. We also know, by combining expressions (1.22) and (1.24), that the dominated applicant will be selected if

$$\Delta s > d\frac{\lambda_b + \lambda_{s_b}}{\lambda_{s_b}}. \tag{1.26}$$

Then, using our definition of $H_\tau(\cdot)$, the probability that an arbitrary applicant from the dominated group outperforms an arbitrary applicant from the dominant group in an arbitrary generation τ can simply be written by $1 - H_\tau\left(d\frac{\lambda_b+\lambda_{s_b}}{\lambda_{s_b}}\right)$. Because in any generation τ there are infinitely many selectors, we can apply a law of large numbers, yielding the following result:

Lemma 1. *For any observer in generation $\tau + 1$, the share of educated individuals among the dominated group from which he learns is given by $\beta_b^{\tau+1} = H_\tau\left(d\frac{\lambda_b+\lambda_{s_b}}{\lambda_{s_b}}\right)$.*

Proof: Follows from the formal arguments right above the lemma. □

We now only need to consider the first generation: if $H_{\tau=1}(d) > 1 - H_{\tau=1}(d)$ holds, then in the subsequent generation $\tau+1$ we have $\beta_b^{\tau+1} < \beta_w^{\tau+1}$. By Proposition 1, this is sufficient to generate a negative stereotype against the dominated group. A negative stereotype implies that the belief regarding the dominated group's mean productivity is negatively biased, $\tilde{\mu}_b^{\tau+1} < \mu_b$. Given that the belief regarding the dominant group's mean productivity is correct (or even upwards biased as in Proposition 2), this implies that $H_{\tau+1=2}(d) > 1 - H_{\tau+1=2}(d)$ too—i.e., there will be inaccurate statistical discrimination. Thereafter, the logic repeats for all following generations. □

Proof of Proposition 1.7

From Proposition 6, we know that a sufficient condition for inaccurate statistical discrimination and negative stereotyping against the dominated group to take place in any generation $\tau = 2, 3, ...$ is that $H_{\tau=1}(d) > 1 - H_{\tau=1}(d)$. As a benchmark case, consider a world where all beliefs regarding both groups are identical and correct, and in which there is no taste-based discrimination. Then, clearly, we are in a scenario in which $H_{\tau=1}(0) = 1 - H_{\tau=1}(0)$. We then only need to verify that i) an increase in d (away from this benchmark) leads to an increase in $H_{\tau=1}(d)$, $\partial H_{\tau=1}(d)/\partial d > 0$, which—all else equal—is true; ii) we need to check that $\partial H_{\tau=1}(d)/\partial \mu_b > 0$. This is true, because we know that for a fixed d, $H_{\tau=1}(d)$ increases when its mean increases. Taking the derivative of $\mathbb{E}[\Delta s] = \mu_b - [\mu_w(\lambda_b + \lambda_{s_b}) - \mu_b\lambda_b]\frac{1}{\lambda_{s_b}}$ with respect to $\tilde{\mu}_b$ yields

$$\frac{\partial \mathbb{E}[\Delta s]}{\partial \tilde{\mu}_b} = 1 + \frac{\lambda_b}{\lambda_{s_b}} > 0, \tag{1.27}$$

and therefore

$$\frac{\partial H_{\tau=1}(d)}{\partial \tilde{\mu}_b} > 0, \tag{1.28}$$

which establishes the result. □

1.A.2 Generalization

In this appendix, I first generalize the main insight from the main text to allow for more general productivity distributions. I show that, under some regularity conditions, the agent's long-run beliefs abut the socially dominated group will be negatively biased—i.e., there will be a negative stereotype (Proposition 8). Afterwards, I provide an example to illustrate that the generalization can capture scenarios in which the individuals are sorted into becoming educated or staying uneducated based on their realized intrinsic productivity parameter (Corollary 5).

General Result

In the main body of the text, I have only considered cases where productivities as well as the signals thereof are normally distributed. In this extension, I allow for more general distributions. I retain the assumption, however, that individuals, if they are educated, simply receive an additive "boost" of $\rho\alpha$ in addition to their realized intrinsic productivity.

Notation mostly follows the main text: I assume that the intrinsic productivity over *all* members of the dominated group is distributed according to G_b with PDF g_b. The productivity of uneducated members of the dominated group is distributed according to $G_{b,u}$ with PDF $g_{b,u}$. The productivity of educated members of the dominated group is distributed according to $G_{b,e}$ with PDF $g_{b,e}$. All densities are strictly positive and all distributions have finite but non-zero variance. These assumptions take care of assumptions (ii) and (iii) in Berk (1966), which is necessary so that we can later apply his convergence result.

The agent is interested in learning the mean of G_b—as in the main text—still denoted μ_b. I assume that $G_{b,u}$ provides some information about μ_b. Similarly, I assume that $G_{b,e}$ provides information about μ_b. However, in the case of samples drawn from $G_{b,e}$, this information is entangled with the impact of education. Specifically, I assume that the mean of $G_{b,e}$ (denote $\mu_{b,e}$) is a function of μ_b and $\rho\alpha$, where these two terms are to some extent substitutes. More formally, I make the following assumption:

Assumption 1.2. *We have that $\mu_{b,e}$ is a function of both μ_b and $\rho\alpha$, such that (i) $\mu_{b,e}$ is continuously differentiable in both μ_b and $\rho\alpha$, (ii) $\partial\mu_{b,e}/\partial\mu_b > 0$, (iii) $\partial\mu_{b,e}/\partial\rho\alpha > 0$, and (iv) $\forall\ \mu_b, \rho, \rho^* \ \exists!\ \mu_b^* : \mu_{b,e}(\mu_b, \rho\alpha) = \mu_{b,e}(\mu_b^*, \rho^*\alpha)$.*

In words, what this implies is that if ρ is misperceived by the agent—which is the case in my model—there always exists some possible value of μ_b that the agent can entertain such that his belief about $\mu_{b,e}$ can remain unaffected with respect to the correctly specified model. Assumption 1.2 takes care of assumption (i) in Berk (1966), and makes one further assumption that implies that the asymptotic carrier is a singleton—i.e., that there is convergence opposed to cycling. I further make the assumption that the agent knows all other parameters of the distributions G_b, $G_{b,u}$, and $G_{b,e}$ and is only unsure about μ_b. Furthermore, Assumption 1.1 from the main text still holds—i.e., the agent still distorts ρ to $\tilde{\rho}$ in line with some salience-based weighting function.

Most importantly, the above described generalization captures scenarios in which whether an individual is getting educated depends on her realized intrinsic productivity. Whenever this is the case, we have to make stronger assumptions regarding the agent's beliefs about the share of educated people among the dominated and dominant groups respectively. In the main text, I have mostly assumed that the agent just has a (correct) point-belief about these shares. In a model with selection as described above, however, this implies that the agent could always infer the mean intrinsic productivity from the shares of educated people. Hence, the learning problem I consider in this paper effectively ceases to exist. To combat this, I assume instead that the agent is not sure about the exact shares of educated and uneducated individuals among the dominated and the dominant groups, respectively. Specifically, I assume that the agent's belief regarding the share of educated individuals among the dominated group is described by the distribution χ_b with support on $(0,1)$—i.e., $\tilde{\beta}_b \sim \chi_b$. Similarly, I assume that the agent's belief about the share of educated individuals among the dominant group is described by the distribution χ_w with support on $(0,1)$—i.e., $\tilde{\beta}_w \sim \chi_w$. With this in mind, we have to restate Definition 1 on representativeness to allow for the uncertainty in β_b and β_w. I say the following.

Definition 1.5. *Being educated is representative of the dominant group if*

$$\frac{\chi_b(\beta)}{\chi_w(\beta)} > 1 \;\; \forall \beta \in (0,1). \tag{1.29}$$

In words, being educated is representative of the dominant group if χ_w first order stochastically dominates χ_b. Finally, for simplification, I make the assumption that the agent treats his belief about the shares as fixed and does not attempt to learn them more precisely.[25] Notice that the new definition also implies that all signs in Assumption 1.1 have to be reversed.

To continue with the analysis, no—like in the model presented in the main text—we know that with a probability of $1-\beta_b$, the agent draws an observation from an uneducated member of the dominated group. With probability β_b instead, he draws an observation from an educated member of the dominated group. The agent's misspecification (i.e., him updating his belief as if ρ were given by $\tilde{\rho}$) is only present in the latter. As such, we can still interpret the model as a learning model with *occasional* model misspecification. Because whether the agent draws an uneducated or educated individual, however, is random, we can restate the problem into one where we say that the agent observes, in each period t, a vector $(\mathbb{I}_{b,u}, a^t)$, where $\mathbb{I}_{b,u}$ is an indicator variable equal to one if the type of the individual drawn in that period is (b,u)—i.e., with probability $1-\beta_b$—and equal to zero otherwise—i.e., with probability β_b (for convenience I will sometimes drop the indices of $\mathbb{I}$.). The distribution of a^t is then best defined conditional on $\mathbb{I}_{b,u}$, such that $a^t \mid \mathbb{I}_{b,u} = 1 \sim g_{b,u}(\cdot)$ and $a^t \mid \mathbb{I}_{b,e} = 0 \sim g_{b,e}(\cdot)$. I denote the joint distributions by $g(s,\mathbb{I})$ and $\tilde{g}(s,\mathbb{I})$, where the first corresponds to

[25] In principal, we could also avoid the problem simply by assuming that the agent has point beliefs about the respective shares of educated individuals but is naive about how these shares are connected to the population mean. Otherwise we could also assume that the agent knows the ratio of educated individuals among both groups but not the respective shares. This would combat the problem too.

the true model of the world (i.e., with ρ) and the latter corresponds to the misspecified model (i.e., with ρ replaced by $\tilde{\rho}$). In other words, we have reformulated the problem into one where the agent's model is not occasionally but always misspecified. This allows us to use the well known result that, as $t \to \infty$, the agent's belief converges in distribution to the parameter that minimizes the Kullback-Leibler divergence between the correct and the misspecified model—if this parameter is unique (see Berk (1966) for an early treatment of this from statistics, or White (1982) for an econometric treatment). This yields the following result:

Proposition 1.8 (Stereotypes Under General Distributions)**.** *Under Assumption 2: As $t \to \infty$, the agent's belief converges in distribution to a Dirac measure at some point $\mu_b^* < \mu_b$—i.e., there is a negative stereotype against the dominated group in the long run.*

Proof: Given that Assumption 1.2 holds, the convergence proof follows immediately from the main Theorem in Berk (1966). We need only establish that the minimizer of the Kullback-Leibler divergence is indeed smaller than the true value—i.e., $\mu_b^* < \mu_b$. For this it is useful to write down the Kullback-Leibler divergence. We have

$$\begin{aligned} D(g(s,\mathbb{I}) \;||\; \tilde{g}(s,\mathbb{I})) &= \mathbb{E}_{s\sim g}\left[\log \frac{g(s,\mathbb{I})}{\tilde{g}(s,\mathbb{I})}\right] \\ &= \sum_{\mathbb{I}\in\{0,1\}} \int_{s\in\mathbb{R}} g(s,\mathbb{I}) \log \frac{g(s,\mathbb{I})}{\tilde{g}(s,\mathbb{I})} ds \\ &= \beta_m \int_{s\in\mathbb{R}} g(s,0) \log \frac{g(s,0)}{\tilde{g}(s,0)} ds \\ &\quad + (1-\beta_m) \int_{s\in\mathbb{R}} g(s,1) \log \frac{g(s,1)}{\tilde{g}(s,1)} ds, \end{aligned}$$

where the last step is possible because $\mathbb{I} \sim \text{Bernoulli}(\beta_b)$. We can see from the abvove expression that we can express the Kullback-Leibler divergence as the weighted sum of the Kullback-Leibler divergences of the conditional distributions. This implies that the minimizer, too, must be a weighted sum of the minimizers of the two divergences of the conditional distributions. Hence, I can proceed by considering the individual minimizers separately. Recall that the model is only misspecified in the case where $\mathbb{I} = 1$—i.e., we have that $g(s,0) = \tilde{g}(s,0)$. This implies that the minimizer of the second term of the above Kullback-Leibler divergence, denote $\tilde{\mu}_b^{**}$, is minimized by setting $\tilde{\mu}_b^{**} = \mu_b$. Hence, for the proposition to hold, it is sufficient if the minimizer of the first term of the above divergence, denote $\tilde{\mu}_b^{***}$, is smaller than the true mean, $\tilde{\mu}_b^{***} < \mu_b$. The Kullback-Leibler divergence between two distributions is always minimized in the case where the two distributions coincide. Because the agent knows all parameters of the distribution except the mean, the Kullback-Leibler divergence will be minimized by setting a belief μ_b^* such that the means of the true distribution and the incorrect distribution are the same. The belief that will yield the two distributions identical must be smaller than the true value. This follows immediately from the assumption that both $\partial\mu_{b,e}(\cdot)/\partial\rho\alpha > 0$ and $\partial\mu_{b,e}(\cdot)/\partial\mu_b > 0$, and that $\forall\ \mu_b, \rho, \rho^*\ \exists!\ \mu_b^* : \mu_{b,e}(\mu_v, \rho\alpha) = \mu_{b,e}(\mu_b^*, \rho^*\alpha)$. If this is true, then a belief $\tilde{\rho} > \rho$ can only be consistent with the observed data by forming a belief $\tilde{\mu}_b < \mu_b$—i.e., if there is a negative stereotype. □

The main point of this generalization is to demonstrate that the main insight of the paper—the existence of negative stereotypes in the long run—carries over to other economically relevant scenarios. For example, the main text assumes that there is no correlation between the intrinsic productivity of an individual and that individual's probability of becoming educated. Indeed, any individual, conditional on their social group membership, has the same probability of becoming educated. If we would relax this—i.e., assume that selection into education is partially based on intrinsic productivity—than clearly among the sample of educated individuals, intrinsic productivity would be distributed differently than among the sample of uneducated individuals. This is precisely what the above extension attempts to capture by allowing for more general distributions of intrinsic productivity that differ between educated and uneducated individuals of one social group.

Example: Hard Selection Rules and the Truncated Normal

As mentioned above, the main point of the generalization is to show that the model can in principle account for selection into the two different levels of education. Here I want to briefly outline more concretely a simple model with such selection that would be in line with the above generalization.

Suppose that $G_b \sim \mathcal{N}(\mu_b, \sigma_b^2)$. The selection rule is as follows: Any individual from the dominated group whose intrinsic productivity is above a threshold $\underline{a}$ is selected into the group of educated individuals and receives a productivity boost of $\rho\alpha$. Any individual below the threshold is selected into the group of uneducated individuals and does not receive a productivity boost. Then we have that $g_{b,u}$ is a truncated normal distribution with minimum $-\infty$ and maximum $\underline{a}$. Its mean is given by

$$\mathbb{E}[a_b \mid a_b < \underline{a}] = \mu_b - \sigma_b \frac{\varphi((\underline{a} - \mu_b)/\sigma_b)}{\Phi((\underline{a} - \mu_b)/\sigma_b)}, \tag{1.30}$$

where $\varphi(\cdot)$ is the PDF of the standard normal distribution and $\Phi(\cdot)$ is its CDF. Similarly, we have that $g_{b,e}$ is a truncated normal distribution with minimum $\underline{a} + \rho\alpha$ and maximum ∞. Its mean is given by

$$\mathbb{E}[a_b + \rho\alpha \mid a_b > \underline{a}] = \mu_b + \sigma_b \frac{\varphi((\underline{a} - \mu_b)/\sigma_b)}{1 - \Phi((\underline{a} - \mu_b)/\sigma_b)} + \rho\alpha \tag{1.31}$$

I additionally assume that the agent cannot precisely observe the individual's productivities and instead observes them with additive iid noise $\varepsilon \in \mathbb{R}$ with zero mean, finite but non-zero variance, and strictly positive density everywhere. This assumption is necessary as otherwise the distributions proposed here would violate Assumptions (ii) and (iii) in Berk (1966), and, hence, we could not apply his convergence result.

To continue, notice that both of the above means, (1.30) and (1.31), carry information about the original distribution's mean, μ_m. Even though statistics has a long standing tradition on research regarding the difficulties associated with recovering information about population parameters when the sample is truncated, recovering the population moments is in principle possible, even though there may not be an analytical solution (see, e.g., Pearson 1902; Fisher 1931; Cohen Jr 1949; Kontonis, Tzamos, and Zampetakis 2019). In our case, the following holds:

Corollary 1.5 (The Truncated Normal Distribution)**.** *Suppose that G_b, $G_{b,e}$, and $G_{b,u}$ are distributed as described above. Then, as $t \to \infty$ the agent's belief converges in distribution to a Dirac measure at the point*

$$\mu_b - (\tilde{\rho}\alpha - \rho\alpha) < \mu_b \tag{1.32}$$

That is, there is a negative stereotype against the dominated group in the long run.

Proof: From Proposition 1.8, we know that under Assumption 1.2 the minimizer of the Kullback-Leibler divergence is unique, and such that it is negatively biased with respect to the true value. Hence, to show that Corollary 1.5 follows from Proposition 1.8 under the respective distributional assumptions, we need to verify that the Assumption 1.2 is indeed fulfilled in the proposed model. As mentioned above, finding an analytical solution—which would otherwise be the most straightforward way to prove the statement—is not generally possible.

In our case, however, we can restate the problem in a simpler way, thereby circumventing the problem of analytically finding a solution to the minimization problem. First notice that the population mean can be expressed as follows:

$$\mu_b = \beta_b \underbrace{\left[\mu_b + \sigma_b \frac{\varphi((\underline{a}-\mu_b)/\sigma_b)}{1-\Phi((\underline{a}-\mu_b)/\sigma_b)}\right]}_{=\mathbb{E}[a_b \mid a_b>\underline{a}]} + (1-\beta_b) \underbrace{\left[\mu_b - \sigma_b \frac{\varphi((\underline{a}-\mu_b)/\sigma_b)}{\Phi((\underline{a}-\mu_b)/\sigma_b)}\right]}_{=\mathbb{E}[a_b \mid a_b<\underline{a}]}. \tag{1.33}$$

That is, the population mean is the weighted average of the averages of the two truncated normals introduced above. Hence, to find the long-run belief about the population mean, we can consider the long-run beliefs about the means of the two truncated distributions separately and then take the weighted average of these two beliefs. If the result is (i) unique and (ii) negatively biased, this proves the corollary.

First, notice that if the agent would separately estimate the mean intrinsic productivity among the subpopulation of uneducated people among the dominated group, $\mathbb{E}[a_b \mid a_b < \underline{a}]$, this estimate would coincide with the correct value. This is because we can find the long-run belief as the unique minimizer of the Kullback-Leibler divergence. Since in this case the model used by the agent is correct, the Kullback-Leibler divergence is uniquely minimized at the correct parameter value, $\mathbb{E}^*[a_b \mid a_b < \underline{a}] = \mathbb{E}[a_b \mid a_b < \underline{a}]$.

Second, notice that if the agent would separately estimate the mean intrinsic productivity parameter among the subpopulation of educated people among the dominated group, $\mathbb{E}[a_b \mid a_b > \underline{a}]$, his estimate would converge in distribution to a Dirac measure at a point strictly below the true mean. To see this, let us write the true mean productivity (i.e., intrinsic mean productivity + $\rho\alpha$) of these individuals as

$$\mathbb{E}[a_b + \rho\alpha \mid a_b > \underline{a}] = \mu_{b,e}^{int} + \rho\alpha. \tag{1.34}$$

The agent instead thinks that this expression is given by

$$\mathbb{E}[a_b + \tilde{\rho}\alpha \mid a_b > \underline{a}] = \tilde{\mu}_{b,e}^{int} + \tilde{\rho}\alpha. \tag{1.35}$$

Now we must check if there exists a unique $\tilde{\mu}_{b,e}^{int}$ such that

$$\mathbb{E}[a_b + \rho\alpha \mid a_b > \underline{a}] = \mu_{b,e}^{int} + \rho\alpha = \tilde{\mu}_{b,e}^{int} + \tilde{\rho}\alpha = \mathbb{E}[a_b + \tilde{\rho}\alpha \mid a_b > \underline{a}] \tag{1.36}$$

Rearranging gives

$$\mu_{b,e}^{int^*} := \tilde{\mu}_{b,e}^{int} = \mu_{b,e}^{int} - (\tilde{\rho}\alpha - \rho\alpha). \tag{1.37}$$

If the agent believes the true mean intrinsic productivity among the educated members of the dominated group to be $\mu_{b,e}^{int^*}$, then this implies that his belief about the mean productivity of this population (again, intrinsic productivity + $\rho\alpha$ would be correct. In this sense his model would appear to him is coinciding with the true model. Therefore the Kullback-Leibler divergence is uniquely minimized at this point. This point is negatively biased, which follows from inspecting expression (1.37) and remembering that $\tilde{\rho} > \rho$.

Finally, putting the two subpopulation estimates together, and using expression (1.33), we know that the unique minimizer of the Kullback-Leibler divergence between the correct and incorrect unconditional distributions (i.e., over the whole population) must be given by

$$\begin{aligned}\tilde{\mu}_b &= \beta_b \left\{\mathbb{E}[a_b \mid a_b > \underline{a}] - (\tilde{\rho}\alpha - \rho\alpha)\right\} + (1-\beta_b)\mathbb{E}[a_b \mid a_b < \underline{a}] \\ &= \beta_b \mathbb{E}[a_b \mid a_b > \underline{a}] + (1-\beta_b)\mathbb{E}[a_b \mid a_b < \underline{a}] - (\tilde{\rho}\alpha - \rho\alpha) \\ &= \mu_b - (\tilde{\rho}\alpha - \rho\alpha) < \mu_b\end{aligned}$$

where the last inequality follows because $\tilde{\rho} > \rho$. □

1.B References

Aigner, D. J., and G. G. Cain. 1977. "Statistical theories of discrimination in labor markets." *Ilr Review* 30 (2): 175–187.

Allport, G. W., K. Clark, and T. Pettigrew. 1954. "The Nature of Prejudice."

Arrow, K. J. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets,* 1–33. Princeton University Press.

Barberis, N., A. Shleifer, and R. Vishny. 1998. "A model of investor sentiment." *Journal of Financial Economics* 49 (3): 307–343.

Barron, K., R. Ditlmann, S. Gehrig, and S. Schweighofer-Kodritsch. 2022. "Explicit and implicit belief-based gender discrimination: A hiring experiment."

Bartoš, V., M. Bauer, J. Chytilová, and F. Matějka. 2016. "Attention discrimination: Theory and field experiments with monitoring information acquisition." *American Economic Review* 106 (6): 1437–1475.

Becker, G. S. 1957. *The Economics of Discrimination.* University of Chicago press.

Berk, R. H. 1966. "Limiting Behavior of Posterior Distributions when the Model is Incorrect." *The Annals of Mathematical Statistics* 37 (1): 51–58. https://doi.org/10.1214/aoms/1177699597.

Bertrand, M., and S. Mullainathan. 2004. "Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination." *American Economic Review* 94 (4): 991–1013.

Bohren, J. A. 2016. "Informational herding with model misspecification." *Journal of Economic Theory* 163:222–247.

Bohren, J. A., K. Haggag, A. Imas, and D. G. Pope. 2019. *Inaccurate statistical discrimination.* Technical report.

Bohren, J. A., and D. N. Hauser. 2021. “Learning with heterogeneous misspecified models: Characterization and robustness.” *Econometrica* 89 (6): 3025–3077.

Bohren, J. A., A. Imas, and M. Rosenberg. 2019. “The dynamics of discrimination: Theory and evidence.” *American Economic Review* 109 (10): 3395–3436.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer. 2016. “Stereotypes.” *The Quarterly Journal of Economics* 131 (4): 1753–1794.

Bordalo, P., N. Gennaioli, and A. Shleifer. 2012a. “Salience in experimental tests of the endowment effect.” *American Economic Review* 102 (3): 47–52.

———. 2012b. “Salience theory of choice under risk.” *The Quarterly Journal of Economics* 127 (3): 1243–1285.

———. 2013a. “Salience and asset prices.” *American Economic Review* 103 (3): 623–28.

———. 2013b. “Salience and consumer choice.” *Journal of Political Economy* 121 (5): 803–843.

———. 2022. “Salience.” *Annual Review of Economics* 14:521–544.

Bushong, B., M. Rabin, and J. Schwartzstein. 2021. “A model of relative thinking.” *The Review of Economic Studies* 88 (1): 162–191.

Coate, S., and G. C. Loury. 1993. “Will affirmative-action policies eliminate negative stereotypes?” *The American Economic Review,* 1220–1240.

Cohen Jr, A. C. 1949. “On estimating the mean and standard deviation of truncated normal distributions.” *Journal of the American Statistical Association* 44 (248): 518–525.

Collins, C. 2018. “What is white privilege, really.” *Teaching Tolerance* 60:1–11.

Dertwinkel-Kalt, M., and M. Köster. 2020. “Salience and skewness preferences.” *Journal of the European Economic Association* 18 (5): 2057–2107.

Devine, P. G., and A. J. Elliot. 1995. “Are racial stereotypes really fading? The Princeton trilogy revisited.” *Personality and social psychology bulletin* 21 (11): 1139–1150.

Dumm, R. E., D. L. Eckles, C. Nyce, and J. Volkman-Wise. 2020. "The representative heuristic and catastrophe-related risk behaviors." *Journal of Risk and Uncertainty* 60:157–185.

Erber, R., and S. T. Fiske. 1984. "Outcome dependency and attention to inconsistent information." *Journal of Personality and Social Psychology* 47 (4): 709.

Esponda, I., R. Oprea, and S. Yuksel. 2023. "Seeing what is representative." *The Quarterly Journal of Economics*, forthcoming.

Esponda, I., and D. Pouzo. 2016. "Berk–Nash equilibrium: A framework for modeling agents with misspecified models." *Econometrica* 84 (3): 1093–1130.

Feld, J., E. Ip, A. Leibbrandt, and J. Vecci. 2022. "Identifying and overcoming gender barriers in tech: A field experiment on inaccurate statistical discrimination."

Fershtman, D., and A. Pavan. 2021. "" Soft" Affirmative Action and Minority Recruitment." *American Economic Review: Insights* 3 (1): 1–18.

Fisher, R. A. 1931. "Properties and applications of Hh functions." *Mathematical tables* 1 (815-852): 2.

Flores, L., and M. A. Fonseca. 2022. "Do In-Group Biases Lead to Overconfidence in Performance? Experimental Evidence." *Available at SSRN 4287592.*

Fouka, V., S. Mazumder, and M. Tabellini. 2022. "From immigrants to Americans: Race and assimilation during the Great Migration." *The Review of Economic Studies* 89 (2): 811–842.

Frick, M., R. Iijima, and Y. Ishii. 2022. "Dispersed behavior and perceptions in assortative societies." *American Economic Review* 112 (9): 3063–3105.

———. 2023. "Belief convergence under misspecified learning: A martingale approach." *The Review of Economic Studies* 90 (2): 781–814.

Fryer Jr, R., and M. O. Jackson. 2008. "A categorical model of cognition and biased decision making." *The BE Journal of Theoretical Economics* 8 (1).

Fryer Jr, R. G., P. Harms, and M. O. Jackson. 2019. "Updating beliefs when evidence is open to interpretation: Implications for bias and polarization." *Journal of the European Economic Association* 17 (5): 1470–1501.

Fryer Jr, R. G., and G. C. Loury. 2005. "Affirmative action and its mythology." *Journal of Economic Perspectives* 19 (3): 147–162.

Fudenberg, D., G. Romanyuk, and P. Strack. 2017. "Active learning with a misspecified prior." *Theoretical Economics* 12 (3): 1155–1189.

Glaeser, E. L. 2005. "The political economy of hatred." *The Quarterly Journal of Economics* 120 (1): 45–86.

Hamilton, D. L., and R. K. Gifford. 1976. "Illusory correlation in interpersonal perception: A cognitive basis of stereotypic judgments." *Journal of Experimental Social Psychology* 12 (4): 392–407.

Hamilton, D. L., and T. L. Rose. 1980. "Illusory correlation and the maintenance of stereotypic beliefs." *Journal of Personality and Social Psychology* 39 (5): 832.

Heidhues, P., B. Kőszegi, and P. Strack. 2018. "Unrealistic expectations and misguided learning." *Econometrica* 86 (4): 1159–1214.

———. 2021. "Convergence in models of misspecified learning." *Theoretical Economics* 16 (1): 73–99.

———. 2023a. "Misinterpreting Yourself." *Working Paper.*

———. 2023b. "Overconfidence and Prejudice." *Working Paper.*

Hilton, J. L., and W. Von Hippel. 1996. "Stereotypes." *Annual Review of Psychology* 47 (1): 237–271.

Hübert, R., and A. T. Little. 2023. "A Behavioural Theory of Discrimination in Policing." *The Economic Journal,* uead043.

Judd, C. M., and B. Park. 1993. "Definition and assessment of accuracy in social stereotypes." *Psychological Review* 100 (1): 109.

Kahneman, D., and A. Tversky. 1972. "Subjective probability: A judgment of representativeness." *Cognitive psychology* 3 (3): 430–454.

Kontonis, V., C. Tzamos, and M. Zampetakis. 2019. "Efficient truncated statistics with unknown truncation." In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS),* 1578–1595. IEEE.

Lagerlöf, J. N. 2020. “Strategic gains from discrimination.” *European Economic Review* 122:103354.

Lepage, L.-P. 2021. *Bias formation and hiring discrimination.* Technical report. Working Paper.

Levine, A., and T. Campbell. 1972. *Ethnocentrism: Theories of conflict, ethnic attitudes, and group behaviour.*

Marx, D. M., S. J. Ko, and R. A. Friedman. 2009. “The “Obama effect”: How a salient role model reduces race-based performance differences.” *Journal of Experimental Social Psychology* 45 (4): 953–956.

Mengel, F., and P. Campos Mercade. 2021. “Non Bayesian Statistical Discrimination.” *Available at SSRN 3843579.*

Morgan, J., and F. Várdy. 2009. “Diversity in the Workplace.” *American Economic Review* 99 (1): 472–85.

Mullen, B., R. Brown, and C. Smith. 1992. “Ingroup bias as a function of salience, relevance, and status: An integration.” *European Journal of Social Psychology* 22 (2): 103–122.

Neuberg, S. L., and S. T. Fiske. 1987. “Motivational influences on impression formation: outcome dependency, accuracy-driven attention, and individuating processes.” *Journal of Personality and Social Psychology* 53 (3): 431.

Onuchic, P. 2022. “Recent contributions to theories of discrimination.” *arXiv preprint arXiv:2205.05994.*

Pearson, K. 1902. “On the systematic fitting of curves to observations and measurements: part II.” *Biometrika* 2 (1): 1–23.

Pew Research Center. 2017. *Views about whether whites benefit from societal advantages split sharply along racial and partisan lines.* https://www.pewresearch.org/fact-tank/2017/09/28/views-about-whether-whites-benefit-from-societal-advantages-split-sharply-along-racial-and-partisan-lines/.

———. 2019. *10 facts about today’s college graduates.* https://www.pewresearch.org/fact-tank/2022/04/12/10-facts-about-todays-college-graduates/#:~:text=Among%20adults%20ages%2025%20and,2021%20Current%20Population%20Survey%20data..

Pew Research Center. 2021. *STEM Jobs See Uneven Progress in Increasing Gender, Racial and Ethnic Diversity.* https://www.pewresearch.org/science/2021/04/01/stem-jobs-see-uneven-progress-in-increasing-gender-racial-and-ethnic-diversity/.

———. 2022. *Black Americans differ from other U.S. adults over whether individual or structural racism is a bigger problem.* https://www.pewresearch.org/fact-tank/2022/11/15/black-americans-differ-from-other-u-s-adults-over-whether-individual-or-structural-racism-is-a-bigger-problem/.

Phelps, E. S. 1972. "The statistical theory of racism and sexism." *The American Economic Review* 62 (4): 659–661.

Pikulina, E., and D. Ferreira. 2023. "Subtle Discrimination." *Working Paper.*

Pratto, F., and J. A. Bargh. 1991. "Stereotyping based on apparently individuating information: Trait and global components of sex stereotypes under attention overload." *Journal of Experimental Social Psychology* 27 (1): 26–47.

Prendergast, C., and R. H. Topel. 1996. "Favoritism in organizations." *Journal of Political Economy* 104 (5): 958–978.

Rubio, M. 2022. *From plantations to prisons: The race gap in incarceration after the abolition of slavery in the US.* Technical report. ECONtribute Discussion Paper.

Sarsons, H. 2017. "Interpreting signals in the labor market: evidence from medical referrals." *working paper.*

Sarsons, H., K. Gërxhani, E. Reuben, and A. Schram. 2021. "Gender differences in recognition for group work." *Journal of Political economy* 129 (1): 101–147.

Schwartzstein, J. 2014. "Selective attention and learning." *Journal of the European Economic Association* 12 (6): 1423–1452.

Sidanius, J., and F. Pratto. 2001. *Social dominance: An intergroup theory of social hierarchy and oppression.* Cambridge University Press.

Siniscalchi, M., and P. Veronesi. 2021. *Self-image bias and lost talent.* Technical report. National Bureau of Economic Research.

Snyder, M., and S. W. Uranowitz. 1978. "Reconstructing the past: Some cognitive consequences of person perception." *Journal of Personality and Social Psychology* 36 (9): 941.

Sumner, W. G. 1906. *Folkways: A study of mores, manners, customs and morals.* Ginn.

Tajfel, H., and J. C. Turner. 1982. "Social psychology of intergroup relations." *Annual Review of Psychology* 33 (1): 1–39.

White, H. 1982. "Maximum Likelihood Estimation of Misspecified Models." *Econometrica* 50 (1): 1–25. Accessed February 9, 2023. http://www.jstor.org/stable/1912526.

Word, C. O., M. P. Zanna, and J. Cooper. 1974. "The nonverbal mediation of self-fulfilling prophecies in interracial interaction." *Journal of Experimental Social Psychology* 10 (2): 109–120.

Chapter 2

Meritocracy in Hierarchical Organizations

Timm Behler and Patrik Reichert

Abstract

Competitive promotions are often perceived as meritocratic because they typically select talented players with a high probability. We show that this is not necessarily the case in hierarchical organizations with more than two layers. Competition backfires by inducing middle managers to block talented subordinates' promotions out of fear that they could negatively affect their own career prospects. Seniority-based promotions can mitigate this. However, a new trade-off arises: maximizing middle-manager ability may not maximize top-tier ability. This depends on bottom-tier ability and how well middle managers can infer subordinates' abilities. We discuss implications for wage inequality, effort incentives, and promotion rates.

2.1 Introduction

Organizations—such as firms, political parties, or governmental institutions—are often hierarchical in nature. Their design reflects an attempt to help the most talented of its members to rise through the ranks and take on leading roles. In the economics literature, the allocation of talent within organizations has often been modeled with contests. In contests, several competitors exert costly effort to produce some observable output, based on which the principal makes a promotion decision. Such competitive promotion schemes are typically regarded as meritocratic, because—under the standard modeling paradigm—they help talented members of the organization to get a promotion.

In this paper, we challenge this view by embedding competitive promotion schemes into hierarchical organizations with more than two layers. In our model, competitive promotion schemes on the organization's middle tier can create a negative spillover to the tiers below: If the organization's middle tier can influence the career advancement of subordinates, they may block the promotion of talented individuals out of fear that it could negatively affect their own career prospects. We show that this moral hazard problem can be alleviated by basing promotions on seniority instead. However, depending on how well the middle tier can infer the abilities of individuals on the lowest tier, there may be a trade-off between the selection of talent onto the middle and highest tier of the organization. While the seniority scheme always maximizes the expected talent on the middle tier, it can both increase or decrease the expected talent on the highest tier. Perhaps surprisingly, depending on the distribution of talent on the bottom tier, more noise in the middle tier's evaluation of their subordinates can both favor or disfavor the seniority scheme when the principal wants to maximize expected ability on the highest tier. Beyond this, our paper also discusses implications for effort provision on the bottom and middle tiers, within-firm wage distributions, and promotion rates.

The extent of the moral hazard problem we describe above is illustrated by recent empirical evidence on "talent hoarding" in organizations. Haegele (2022), for example, provides empirical evidence that managers indeed attempt to block the promotion of high ability subordinates. She argues that managers would weaken team performance by promoting talented workers away from the team, thereby decreasing their own payoffs and, hence, resulting in incentives to promote untalented workers (see also Friebel and Raith 2022). Even beyond the academic literature, industry surveys suggest that many firms view the immobility of talent across their hierarchies as a central threat to the efficiency of their organizations (i4cp 2016), and management magazines often publish advice for workers who feel that they are held back by their superiors, further highlighting the scope of the problem (see, for example, Business Insider (Smith 2015) or Forbes (Kurter 2021)).

In section 3.2 of this paper, we formally introduce the model. It consists of an organization that employs one manager (the middle tier) and a large pool of workers (the bottom tier).[1] All members of the organization can have high or low ability. The game proceeds as follows: The manager is tasked with selecting one of the workers to be promoted to management level. Before doing so, she receives, for each worker, a binary ability signal. Thereafter, she makes the promotion decision. Some time after her decision has been implemented she learns about the entrant's true ability and the two managers compete over a promotion to a higher level of the organization. We model this competition through a (biased) all-pay auction: Both managers exert costly effort. Effort, together with the bias, yields the managers' scores and the manager with the higher score is promoted. This biased all-pay contest also captures the seniority scheme when the bias goes towards infinity in favor of the incumbent manager.

We start our analysis with some preliminaries in section 2.3. Our main result is in section 2.4. We focus on how decreasing the competitiveness of the promotion scheme can alleviate the moral hazard problem. Explicitly, we compare the competitive promotion scheme to a seniority-based promotion scheme where the more senior manager is always promoted at the end of the period. Our first result is that the seniority scheme always maximizes the expected ability of the worker who is promoted to the vacant management position because it fully alleviates the moral hazard problem. This intuitive result is reminiscent of Carmichael (1988), who shows that employment-guarantees are a necessary condition to incentivize the current members of an organization to make high-quality recruitments.

Our main result, however, goes beyond this: We consider the expected ability of the manager who is promoted to the subsequent tier of the organization. Whether expected ability is maximized by choosing the seniority scheme depends crucially on noise. Perhaps surprisingly, noise has an ambiguous effect: When the share of high-ability players among the workers is low, little noise favors the competitive scheme. When the share of high-ability players among the workers is high instead, little noise favors the seniority scheme. The intuition is, roughly, as follows: The incumbent manager can have either high or low ability. As our analysis reveals, under

[1]Throughout this paper, we use the manager-worker framework, but the logic applies also to other organizations—e.g., political parties, governmental institutions, or universities.

the competitive scheme, any high ability incumbent wants to promote a low ability worker, whereas a low ability incumbent wants to promote a high ability worker. They will attempt to achieve this by picking a worker who produced a corresponding signal. However, signals will sometimes not match a workers true ability. In this case, the incumbent has made a "mistake" in her promotion decision. From the perspective of the principal there are "good" and "bad" mistakes: If a high ability incumbent mistakenly promotes a high ability worker, this is good for the principal. If, instead, a low ability manager mistakenly promotes a low ability worker, this is bad for the principal. Both types of mistakes become more likely as noise increases. However, if the distribution of ability at the bottom tier is not symmetric, the likelihoods of both mistakes will also increase asymmetrically. This means that one mistake will become *relatively* more likely than the other. This also implies that an increase in noise can both increase or decrease the expected ability of a promoted worker, and therefore the expected ability of a promoted manager under the competitive scheme. As a result, more noise may favor either the seniority or the competitive scheme under appropriate conditions.

In section 2.5, we analyze several trade-offs that the organization faces. First, we consider an effort trade-off: Because expected ability on the middle tier is always maximized by the seniority scheme, whereas effort exerted by the middle tier is maximized by the competitive scheme, these two objectives are always in conflict. Second, because expected ability on the top level is sometimes maximized by the seniority scheme and sometimes by the competitive scheme, it will always be in conflict with one of the other two objectives and aligned with the other.

We discuss our results and their implications for other design choices the principal might have in section 2.6. For example, our results might have an impact on within-firm wage distributions. When the principal decides on the seniority scheme to maximize the expected ability on the middle tier (and potentially on the top tier), then there are no effort incentives anymore since the promotion decision cannot be affected by exerting effort. In this case there is no need for the principal to offer a high wage at the top tier of the hierarchy. This suggests a less steep wage increase between the middle and the top tier. At the same time, because the seniority scheme maximizes expected ability on the middle tier, being a middle manager might become a stronger ability signal to outside organizations. Therefore, the organization may have to increase the wages of the middle tier to retain its managers. This suggest a steeper wage increase between the bottom and the middle tier. We also discuss implications for effort provision on the bottom tier and promotion rates. Regarding the latter, our model suggests a novel rationale for up-or-out rules because they plausibly affect the signal precision of the pool of workers eligible for promotion.

section 2.7 concludes. The remainder of the current section summarizes the most relevant literature. Additional literature will be discussed throughout the article where useful. All omitted proofs are in Appendix D.

Related Literature. Our article is related to two strands of literature, which we discuss successively. We first focus on the literature on perverse incentives of principals to block the progress of talented subordinates, before we move on to a discussion of the the literature on talent identification in contests.

To the best of our knowledge, Haegele (2022) is the first to document the perverse incentives we study using internal firm data. She provides evidence that managers hoard talent by discouraging talented workers from applying for a promotion. She shows that a temporary reduction in managers' incentives to hoard talent increases promotion applications by 123%. Friebel and Raith (2022) provide a theoretical analysis of this phenomenon. Beyond this, while not providing evidence on hoarding directly, management scholars (Hu, Wang, and Huang 2019) also show that managers act less favorably towards talented proteges if they feel threatened by their ability.

Although it is challenging to record these patterns in firms, a closely related research agenda is present in the political science and the political economics literature. Bai and Zhou (2019) documents the purge of competent party cadres by Mao and his faction during the years of the Cultural Revolution. They argue that as external threats subsided after the Chinese Civil War, the relative value of competence vis-à-vis loyalty decreased, which in turn lead to the demotion of competent central committee members (measured as education and military rank percentile by birth cohort). The emerging importance of political connections in the Chinese Communist Party may be considered as a response to this "loyalty-competence trade-off." Strong political connections are considered as insurance against disloyalty by party leaders. The party elite is, therefore, more likely to promote more competent provincial leaders conditional on them having strong connections. Otherwise competence remains to be construed as a threat to power (Jia, Kudamatsu, and Seim 2015).

Considering talented subordinates as threats in the eyes of the elite is not surprising given the numerous examples in history (Egorov and Sonin 2011). Recent empirical evidence further validates this assertion. Using rich register data to measure abilities, Besley et al. (2017) shows that local party leaders indeed have a lower probability of holding on to power when the share of high ability elected representatives on the party list increases. The key assumption in related theoretical work that generates similar results is the positive relationship between productive ability and the ability to wrong the principal (Glazer 2002; Wagner 2006; Egorov and Sonin 2011; Zakharov 2016). We also make this assumption implicitly in our model, given the effect of ability difference on winning probabilities in contests.

An early theoretical contribution focusing on negative selection induced by competitive incentives is by Carmichael (1988). He presents a model of hiring decisions in university departments. Current professors are best equipped to evaluate possible recruits in their fields. They are, however, reluctant to select the most talented applicants if they fear that they may be replaced by them. As a result, an employment-guarantee ("tenure" in the academic context) is a necessary condition to incentivize professors to make favorable recruitments. We obtain a similar result in Proposition 2.1. Sengupta (2004) builds on the same logic but focuses on a different type of moral hazard. He assumes that after filling a vacancy, the firm observes the joint output of the recruiter and her recruit, and interprets this as a signal of their abilities. If the recruiter has recruited a talented worker, she fears to be replaced by her recruit and therefore reduces her own output to make the signal less informative about both players' abilities. Again, an employment-guarantee can mitigate this problem. Other articles consider similar problems and suggest different solutions: Friebel and Raith (2004) show that the moral hazard problem can be alleviated by establishing clear chains of communication within the organization. Ekinci (2016) shows that the problem may not exist if employees want to build a reputation for referring talented workers.

The theoretical work belonging to this strand of literature typically focuses only on the promotion from one tier of the organization to the next (or, almost equivalently, on the recruitment of new employees).[2] We go beyond this by considering not only the promotion from the bottom to the middle tier, but also from the middle to the top tier. This view of hierarchical organizations is similar to Tirole (1986) or Kofman and Lawarrée (1993). As we will show in this article, this extended hierarchy is important as it highlights that the optimal policies to allocate talent to different tiers of the organization are not necessarily aligned.

[2]A notable exception being Ekinci (2016), who considers promotion contests on an organizations middle tier as a microfoundation for why managers might nor refer talented workers. He does, however, not analyze the effects that this microfoundation has on the higher tiers of the organization, and in this sense does not consider the full hierarchy of the organization. Providing this analysis is the purpose of our paper.

We model the promotion from the middle to the top tier as an all-pay contest (Hillman and Riley 1989; Baye, Kovenock, and De Vries 1996; Siegel 2009), connecting our article to the respective literature. Within this literature, our article is mainly related to (mostly) recent research that focuses on how seemingly meritocratic design choices fail to generate meritocratic allocations. Meyer (1991) shows that in a multiple-period selection contest, the designer maximizes the expected ability of the winner if she biases the later stages of the contest in favor of the leader of the previous stages. Perhaps surprisingly, Kawamura and Barreda (2014) show that even in a static setting where the designer's prior regarding the participants' abilities is flat, she maximizes the expected ability of the winner by randomly biasing the contest in favor of one of the participants. Drugov and Ryvkin (2017) derive a similar result in a more general framework. Finally, both Hvide and Kristiansen (2003) and Fang and Noe (2022) consider risk-taking contests in which the expected abilities of the winners can be increased by making the competition less competitive, and thereby *seemingly* less meritocratic. The seniority-based promotion mechanism we consider in this article completely detaches promotions from performance and, hence, can be seen as making the promotion mechanism less competitive, too.

2.2 Model

In this section, we introduce a model of competition in hierarchical organizations. The organization consists of $N > 3$ players, who can be either workers or managers. Beyond the workers and managers, there is a principal whom we do not model as an active player. An individual player is denoted by $i \in \{1, ..., N\}$. Players are heterogeneous in their abilities. That is, players differ in how costly it is for them to exert effort $x_i \geq 0$ to produce output $y_i(x_i) = x_i$. We allow for two ability types: a low ability and high ability player with their respective ability denoted by $a_i \in \{l, h\}$ with both l and h strictly positive and $l < h$. Let the effort cost for Player i be given by $c_{a_i}(x_i)$, where c_{a_i} is an increasing, continuously differentiable, and convex function with $c_{a_i}(0) = 0$. Low ability players' cost of effort is strictly greater for any strictly positive effort level: $c_l(x) > c_h(x), \ \forall x > 0$. The prior probability of any worker to have high ability is given by $\mu \in (0, 1)$.

The game consists of four stages that proceed as follows: In **Stage 0**, there is one incumbent manager and a vacant management position. The prior probability of the incumbent manager having high ability is $q \in (0, 1)$. The incumbent manager receives a binary signal, $\tilde{a}_i \in \{l, h\}$, of each worker's ability. For all workers, the probability that their signal matches their type is given by $p := \Pr[\tilde{a}_i = a_i] > 1/2$. For simplicity, we assume that N is infinitely large, such that the manager will always observe at least one signal of each kind. This simplifies the analysis significantly and avoids uninteresting special cases. We can then use Bayes' rule to pin down the expected abilities of all workers conditional on their signal,

$$\begin{aligned} \Pr[a_i = h|\tilde{a}_i = h] &= \frac{\mu p}{\mu p + (1-\mu)(1-p)}, \quad \text{and} \\ \Pr[a_i = l|\tilde{a}_i = l] &= \frac{(1-\mu)p}{(1-\mu)p + \mu(1-p)}. \end{aligned} \tag{2.1}$$

In **Stage 1**, the incumbent manager decides on one of the workers to be promoted to the vacant management position. Thereafter, both managers, the incumbent and the entrant, learn of their abilities.[3]

In **Stage 2**, the two managers compete in a (seniority-biased) all-pay contest over a prize $v > 0$ provided by the principal. We can think of this prize as the wage increase due to a promotion to the next higher level of the organizational hierarchy. The rules of the (seniority-biased) all-pay contest are as follows: The two competing managers will exert effort, which will then be combined with a bias—set by the principal—in favor of the more senior manager. More specifically, the winning probabilities of the contest are determined as follows

$$w_m(x_m, x_p; b) = \begin{cases} 1 & \text{if } x_m + b > x_p \\ \frac{1}{2} & \text{if } x_m + b = x_p \\ 0 & \text{if } x_m + b < x_p, \end{cases}$$

where x_m is the effort submission of the incumbent manager, x_p is the effort submission of the entrant, and $b \geq 0$ is the bias that the principal can set in favor of the more senior contestant—i.e., the incumbent manager. In other words, the entrant has to outperform the incumbent manager by a margin b to win the contest. Notice that the setup also captures two important special cases: If $b = 0$, the contest is a standard all-pay contest between the incumbent manager and the entrant. If $b \to \infty$ instead, the incumbent manager will win the contest with probability one. Promotion in this case is not competitive, but is solely based on seniority. Generally, we can say that as b increases, the contest becomes less competitive in the sense that the winning probabilities are less sensitive to the output difference between the two contestants.

[3]In reality of course, they might not learn their abilities immediately, but rather by working together for a while.

Finally, in **Stage 3**, the principal selects the winner of the contest and promotes her to the next higher organizational tier. The timing of the game is summarized in Figure 2.1.

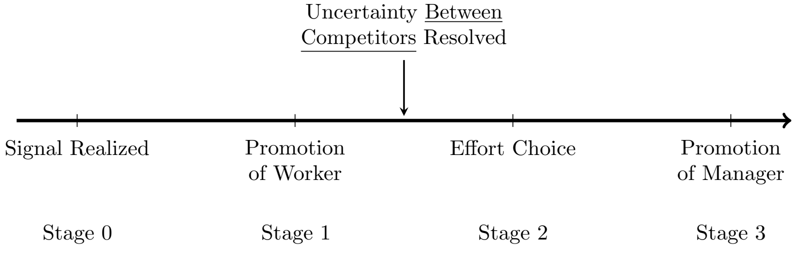

Figure 2.1: Timing of the game.

Throughout the paper, we assume that both the incumbent manager and the entrant are risk-neutral expected utility maximizers. The incumbent manager's decision problem consists of first choosing her opponent and then competing against her by making an effort choice. The entrants's decision problem, instead, consists only of her effort choice in the competition against the incumbent manager. We look for subgame-perfect Nash equilibria (SPNE, henceforth) in the game between the two managers.

Discussion. A discussion of our assumptions is in order. First, we want to discuss how we interpret our model of the promotion process from the worker-level to the management-level. It may not seem plausible that a manager has the power to promote one of her subordinates to a position that is directly equivalent to her own. However, we do not think of the vacant management position as an exact copy of the incumbent manager's position, but rather as any position within the organization that is of a rank such that the incumbent manager and the entrant could, in principle, be considered for the same position later on. In practice, such a position could be below the incumbent manager's position in rank. Another way to think of the delegation of the promotion decision is by envisioning a manager who cannot unilaterally make promotion decisions and who can only refer one of her subordinates to her own superior instead. Since the manager has an informational advantage over her superior, the superior simply trusts the manager and promotes her referral.[4]

[4]Abstracting away, of course, from the question whether this trust is indeed warranted.

Second, we want to discuss whether the incumbent manager should expect punishment for promoting a low-signal worker. Our model implicitly assumes that there is no such punishment. Our assumption stems from the idea that punishment may not be credible given the information asymmetry between manager and principal. For example, the incumbent manager may promote a worker who is performing well in her worker role, yet produces low ability signal along the ability dimension that is relevant for the manager role—the dimension which is relevant for determining the winner of the promotion contest between managers. Hiding the true intention behind the promotion decision is then relatively easy, which in turn make it less likely that the manager will be punished. Relatedly, any punishment should be based on the principal being convinced that the incumbent promoted a low *signal* worker. This implies that the principal must know not the realized ability of the promoted worker, but must be able to verify the signal the incumbent manager received. This is unlikely. Nonetheless, in subsection 2.A.2, we allow for a cost reflecting that the incumbent manager may face punishment from promoting a low-signal worker and show that we approximate the case when the principal does not have the ability to punish.

Third, we assume that the promotion scheme from the middle to the top tier uses a contest and is therefore distinct from the promotion scheme from the bottom to the middle tier. This modeling choice relates to the argument in the previous paragraph. It reflects that in the bottom tier of most organizations, there are no management skills necessary. Hence, performance at the bottom tier may not be a good signal for performance at the middle tier. We therefore believe that it is plausible that the incumbent manager would rely on other signals, such as education, to make her promotion decision. This is very different when considering middle to top promotions. Performance at the middle tier depends on management skills, which are precisely those skills necessary to succeed at the next higher level. Therefore, basing promotions on performance is plausible in this case.

Fourth, we consider an organization which exists for only one period,[5] whereas real-world organizations often exist for longer. This static model is rich enough to provide interesting insights, however, if we maintain the assumption that a principal is short-lived, will be replaced after she retires, and each replacement can decide again independently on the promotion policy. We believe this to be realistic, since in real-world organizations, there is typically some fluctuation between principals and it is seldom the

[5]subsection 2.A.3 extends the analysis to an infinitely-lived organization and shows that some of our insights carry over.

case that one principal can decide on a promotion scheme implemented by an organization from the ground up. We analyze how the choice between a competitive and a seniority based promotion mechanism may influence (i) the expected ability of the worker promoted to the vacant management position, and (ii) the expected ability of the promoted manager.[6]

Finally, we want to touch on the question why the principal delegates the promotion decision from the bottom to the middle tier to the managers.[7] After all, depending on the type of the incumbent manager, the expected ability of the promoted worker can be lower than the expected ability of a randomly drawn worker. Hence, the principal may be better of by randomly selecting a worker to fill the vacant management position. Still, in a complex multi-level organization, implementing even a random draw from a pool of workers may be too costly for the principal and, therefore, the execution of this selection mechanism would need to be delegated nonetheless. The principal may directly order managers to randomly select a worker, but the manager could still select a worker who produced a low signal and *claim* that the selection process was random. As such, our delegation assumption is not a significant departure from the way promotion mechanisms are implemented in most organizations.[8]

[6]We restrict our analysis to considering full seniority promotion schemes to keep our problem tractable. However, as Lemma 2.6 shows, the seniority-based promotion scheme ($b \to \infty$) is simply the limiting case of the promotion scheme that uses some "intermediate" bias $b < c_h^{-1}(v)$ that realigns incentives between the principal and the incumbent manager, when expected punishment costs go to zero.

[7]Delegation is shown to be optimal in the canonical principal-agent model (Strausz 1997), however, the source of moral hazard in our model is managerial discretion in promotion decisions, rather than the possibility of shirking.

[8]Even if we assume that randomly selecting a worker would be feasible for the principal, there may be other reasons for why delegation may be an equilibrium outcome. Consider the following extension of our model: Instead of a two types, workers can be of three types: high ability, low ability, and very low ability. Assume that the workers with very low ability produce a very precise signal, such that any incumbent manager can, essentially, perfectly identify them. If an incumbent manager would promote a worker with very low ability to the vacant management position, this would affect performance so severely that the promotee's type would be perfectly revealed to the principal, too. In turn, the principal would punish the incumbent manager who made the promotion decision. As a result, the incumbent manager will never promote a worker with very low ability. If instead, the principal would decide to randomly select a worker, there is a positive probability that the worker has very low ability. In this case, delegation may be beneficial to the principal.

2.3 Equilibrium in the Promotion Game

In this section, we derive some preliminary results. Specifically, we solve the game by backwards induction, starting with the promotion contest before analyzing the incumbent manager's promotion decision.

2.3.1 The Second Stage

The fully competitive version of the second stage of the game is modeled as an all-pay contest. It is well known that equilibrium in this contest exists only in mixed strategies—although this statement is true only when head starts are not too large as we will show later.[9] The logic why this game has no equilibrium in pure strategy is as follows: For any pure strategy effort submission $\hat{x}$ by some player, the other player has a profitable deviation by bidding $\hat{x} + \varepsilon$ and win the prize with probability one. The best response of the first player is to outbid by a small amount again. This cycle of outbidding continues until one player's maximum valuation is reached. The player with the lower valuation has then an incentive to deviate to exert zero effort, as otherwise she receives a negative expected payoff. But then the bid that prompted the low valuation player to exert zero effort is no dominated by a bid of ε as this would guarantee a win at lower effort cost. At such a low bid, however, the player with the lower valuation becomes competitive again and the cycle of outbidding restarts.

The mixed-strategy equilibrium is relatively straight forward to characterize. When we do not allow for head starts ($b = 0$), players randomize their effort submission on the same support: $[0, c_l^{-1}(v)]$ (Baye, Kovenock, and De Vries 1996; Siegel 2009). The upper bound of the support is given by the low ability player's maximum willingness to pay for the prize.[10] This is equivalent to the effort level x_l that makes the player's expected payoff zero assuming he wins the prize with certainty $v - c_l(x_l) = 0$. Otherwise, the high ability player would bid $c_l^{-1}(v) + \varepsilon$ and win with probability one. Given that the high ability player can outbid the low ability player (the less "powerful" player) and still earn a strictly positive payoff, the low ability player must earn zero in equilibrium. Furthermore, the high ability player must be indifferent between deviating to $c_l^{-1}(v) + \varepsilon$ and playing the equilibrium mixed strategy. This implies that his expected payoff must be

[9] Such that any positive effort submission from the disadvantaged player would be outranked by the favored player's output even if she submits zero effort.

[10] Commonly referred to as the low ability player's "reach" in the literature after Siegel (2009).

equal in both cases, which implies

$$\Pi_h = v - c_h(c_l^{-1}(v)) \quad \text{and} \quad \Pi_l = 0.$$

Using these expected payoffs, we can also back out the equilibrium cumulative distribution functions (CDFs) that players use in the mixed strategy equilibrium; however, we do not use the explicit CDFs for our results in this paper, and hence omit stating them for brevity.

When we allow for positive head starts, players may randomize over different supports. The upper bound of the distribution remains the same for both players and it is still given by the maximum willingness to pay of the less powerful player (Siegel 2009; Kawamura and Barreda 2014). However, when determining which of the two players is more powerful, we must take into account the size of the head start. If head start is large enough, the maximum willingness to pay of a low ability player may exceed that of a high ability player. It is then the low ability player who will earn strictly positive expected payoff in equilibrium. The steps to calculate these expected payoffs (and equilibrium CDFs) remain the same as we saw in the case with no head starts.

In this article we will also consider very large head starts, such that the disadvantaged player can never outbid the advantaged player and still earn a positive expected payoff. In such cases, the equilibrium of the game is in pure strategies: both players submit zero effort in equilibrium and the player who receives the head start wins with probability one (we show this formally in Appendix A (Lemma 2.5)). Using the results in Kawamura and Barreda (2014) and our results from Appendix A, we can specify the expected payoffs for arbitrarily large head starts, $b \in [0, \infty)$. We do this in the following lemma:

Lemma 2.1 (Expected Payoffs)**.** *Players' expected payoffs are characterized as follows:*

1. ***When bias is in favor of the more talented player.***

$$\Pi_h = \begin{cases} v - c_h\left(c_l^{-1}(v) - b\right), & b < c_l^{-1}(v) \\ v, & c_l^{-1}(v) \leq b \end{cases} \qquad \Pi_l = 0$$

2. ***When bias is in favor of the less talented player.***

$$\Pi_h = \begin{cases} v - c_h\left(c_l^{-1}(v) + b\right), & b \le c_h^{-1}(v) - c_l^{-1}(v) \\ 0, & c_h^{-1}(v) - c_l^{-1}(v) < b \end{cases}$$

$$\Pi_l = \begin{cases} 0, & b \le c_h^{-1}(v) - c_l^{-1}(v) \\ v - c_l\left(c_h^{-1}(v) - b\right), & c_h^{-1}(v) - c_l^{-1}(v) < b < c_h^{-1}(v) \\ v, & c_h^{-1}(v) \le b \end{cases}$$

3. ***When both players are high ability.*** *Let the player who receives preferential treatment (positive bias) be denoted h'.*

$$\Pi_{h'} = \begin{cases} v - c_h\left(c_h^{-1}(v) - b\right), & b < c_h^{-1}(v) \\ v, & c_h^{-1}(v) \le b \end{cases} \qquad \Pi_h = 0.$$

4. ***When both players are low ability.*** *Let the player who receives preferential treatment (positive bias) be denoted l'.*

$$\Pi_{l'} = \begin{cases} v - c_l\left(c_l^{-1}(v) - b\right), & b < c_l^{-1}(v) \\ v, & c_l^{-1}(v) \le b \end{cases} \qquad \Pi_l = 0.$$

2.3.2 The First Stage

We proceed with our backwards induction approach and look at the first stage. We can use the expected payoffs in Lemma 2.1 to set up the sufficient condition for the incumbent to promote a worker with a high ability signal:

$$\begin{aligned} &\Pr[a_{-i} = h \mid \tilde{a}_{-i} = h]\left(\Pi_{a_i} \mid b, a_{-i} = h\right) \\ &\qquad + \left(1 - \Pr[a_{-i} = h \mid \tilde{a}_{-i} = h]\right)\left(\Pi_{a_i} \mid b, a_{-i} = l\right) \\ &\qquad \ge \\ &\Pr[a_{-i} = l \mid \tilde{a}_{-i} = l]\left(\Pi_{a_i} \mid b, a_{-i} = l\right) \\ &\qquad + \left(1 - \Pr[a_{-i} = l \mid \tilde{a}_{-i} = l]\right)\left(\Pi_{a_i} \mid b, a_{-i} = h\right). \end{aligned}$$

We can simplify to get

$$\left(\Pi_{a_i} \mid b, a_{-i} = h\right) \ge \left(\Pi_{a_i} \mid b, a_{-i} = l\right) \tag{2.2}$$

Throughout the paper, we assume that when an incumbent is indifferent between promoting a high signal worker and a low signal worker, she promotes the high signal worker.[11]

[11]For several reasons, we believe that this assumption is not particularly strong. First, this assumption makes the contest more favorable from the viewpoint of the principal.

Inspecting the above condition, we can show that the principal must set b so high that she guarantees a win for the incumbent to incentivize the promotion of a high signal worker. This in fact implies that the bias is so strong that the equilibrium in the two-player all-pay auction for the promotion has the same equilibrium effort and win probability as the case when we set $b \to \infty$. We summarize this strategic equivalence result in Lemma 2.2.

Lemma 2.2. *The principal must set bias*

$$c_h^{-1}(v) \leq b,$$

to guarantee that the incumbent will promote a high signal worker. Bias $c_h^{-1}(v) \leq b$ leads to an equilibrium that is strategically equivalent to a full seniority promotion scheme where $b \to \infty$.

The strategic equivalence result in Lemma 2.2 allows us to draw further parallels between our model and how seniority-based promotion schemes are implemented in organizations. An outright guarantee of promotion based on relative tenure is rarely observed. However, if the importance attached to tenure length is substantial enough, this may translate to a head start so large that the resulting strategic interaction is equivalent to what we would expect when promotions are guaranteed for the incumbent.

Now we can consider the equilibrium promotion decision of the incumbent under the competitive scheme and seniority scheme. Generally, in the competitive scheme, a high-ability incumbent manager benefits from trying to promote a low-ability worker through higher expected payoffs in the contest stage. A low-ability incumbent, on the other hand, earns an expected payoff of zero regardless who she competes against. When the sole determinant of promotion to the highest tier is seniority, the ability of the "competitor" does not change the incumbent's expected payoff. These observations lead us to the equilibrium promotion decisions under the two promotion schemes, stated in Lemma 2.3.

In the sense that one of the insights of our paper is that contests may not necessarily br favorable for the principal, the assumption only highlights the severity of our results. Beyond this, allowing for an infinitissimally small cost associated with promoting a low-signal worker (e.g., because of the fear of punishment, reputational concerns, or prosocial preferences in favor of the principal or the organization in genereal) would endogenize this assumption.

Lemma 2.3.

1. *Suppose that the principal has chosen the fully competitive promotion scheme. In the SPNE, a high type incumbent manager promotes a low signal worker, while a low type incumbent promotes a high signal worker.*

2. *Suppose that the principal has chosen the seniority-based promotion scheme. In the SPNE, both types of incumbent managers promote a worker who produced a high signal.*

With these results in hand we are ready to study how the choice of promotion scheme may effect the distribution of talent across the hierarchy.

2.4 The Distribution of Talent Across the Hierarchy

In this section, we consider how the principal of the organization can design the promotion mechanism as to affect the distribution of talent across the organization's hierarchy.

2.4.1 Expected Ability of the Promoted Worker

We start by analyzing which promotion scheme maximizes the expected ability of the worker who is promoted to the vacant management position, summarized in the following proposition.

Proposition 2.1. *For any $p > 1/2$, the expected ability of the worker who fills the vacant management position is higher under the seniority-based promotion scheme than under the fully competitive promotion scheme.*

The intuition is straightforward: When removing the competitive aspect of the promotion mechanism, incumbent managers face no cost associated with promoting a high-ability worker to the vacant position under the seniority scheme. As a result, the seniority scheme will always incentivize them to promote a worker who produced a high signal—i.e., it *completely* alleviates the moral hazard problem.

From Proposition 2.1, we know that the expected ability of the promoted worker is always maximized under the seniority scheme. While signal precision plays no role for this result, it affects the within performance of the different schemes. In the next proposition, which will be helpful to follow the remainder of the analysis, we establish how the expected ability of a promoted worker is affected by a change of precision under both the seniority and the competitive scheme.

Proposition 2.2.

1. *Under the seniority scheme, the expected ability of the promoted worker is increasing in both p and μ.*

2. *Under the competitive scheme, for any $q \in (0,1)$, there exists a threshold value $\mu^t \in (0,1)$ which partitions the interval $(0,1)$ into two sub-intervals, $(0,\mu^t]$ and $[\mu^t,1)$, such that in one sub-interval, expected ability of the promoted worker is increasing in p, whereas it is decreasing in p in the other sub-interval.*

Part (1) of Proposition 2.2 states that under the seniority scheme, expected ability of the promoted worker is increasing in the signal precision and the share of high ability workers. This is intuitive, because both parameters increase the probability of promoting a high ability worker.

Part (2) states that under the competitive scheme, the expected ability of promoted workers can either decrease or increase when signal precision increases. More precisely, for any value of q, there exists a range of μ for which expected ability increases, as well as a range of μ for which expected ability decreases. The intuition is roughly as follows: The incumbent manager can have either high ability or low ability. A high ability incumbent manager promotes a worker with a low ability signal. A low ability incumbent manager promotes a worker with a high ability signal. Noise, however, will imply that sometimes the entrant's ability will not correspond to her signal. In this case, the incumbent has made a mistake. Conditional on having a high ability incumbent, mistakes are good from the perspective of the principal, in the sense that a mistake implies that a high ability worker has been promoted—i.e., more noise increases the expected ability of the promoted worker conditional on the incumbent having high ability. If instead, the incumbent has low ability, mistakes are bad from the perspective of the principal, in the sense that a low ability worker has been promoted—i.e., more noise decreases the expected ability of the promoted worker conditional on the incumbent having low ability. How much noise affects the mistake probabilities depends on the share of high ability workers on the bottom tier, μ. Therefore, whether more precision increases or decreases the expected ability of the promoted worker depends on both q and μ, such that for any q it can either increase or decrease conditional on picking an appropriate μ.

2.4.2 Expected Ability of the Promoted Manager

We now consider which promotion scheme maximizes the expected ability of the manager promoted to the highest tier of the organization. Under the seniority scheme, this expected ability is uniquely pinned down by q. Under the competitive scheme, however, the expected ability of the promoted manager depends on the expected ability of the worker who was previously promoted and on the winning probability of a high ability player in the all-pay contest. This implies that the seniority scheme is maximizing the expected ability of the promoted manager if q is sufficiently large. The next proposition states this.

Proposition 2.3. *The seniority scheme maximizes the expected ability of the promoted manager if*

$$q > \hat{q}(w, p, \mu).$$

The threshold $\hat{q}$ is increasing in ability difference $(h-l)$ and μ. The threshold can either increase and decrease in p depending on appropriate parameter choices.

The threshold level $\hat{q}(w, p, \mu)$ is given in the proof of the proposition, which is in Appendix D. The main insight of the result is intuitive: When q is higher, the expected ability of the promoted manager increases under the seniority scheme and, hence, q must be beyond some threshold for the seniority scheme to be optimal. It is similarly intuitive that the threshold is increasing in ability difference: An increase in the ability difference leads to an increase in w. An increase in w makes the competitive scheme relatively stronger, and, hence, the competitive scheme dominates for lower values of q. The intuition for the effect of a change in μ is more nuanced: Fixing q, good mistakes become more likely while bad mistakes become less likely when we increase μ. This makes the moral hazard induced by the contest less of a concern. The selection advantage of contests then starts to dominate for lower q, even when accounting for the moral hazard in the worker promotion stage.

Even more nuanced is the effect of a change in the precision p. From Proposition 2.2, we know that the expected ability of the promoted worker depends crucially on q, p, and μ. Naturally, this dependency therefore carries over to the expected ability of the promoted manager. The following lemma provides a necessary intermediate step to characterize this relation more carefully.

Lemma 2.4. *Define the threshold noise level*

$$p^t := \frac{q\left((\mu-1)^2 + (2\mu-1)w\right) - \mu^2 w}{(2\mu-1)(q(\mu+w-1)-\mu w)}. \tag{2.3}$$

If and only if $p^t \in (\frac{1}{2}, 1)$, p^t partitions the open interval $(\frac{1}{2}, 1)$ into two sub-intervals, whereby the expected ability of the promoted manager is maximized by the seniority promotion scheme when p is an element of one sub-interval, and by the competitive promotion scheme if p is an element of the other sub-interval. We have that $p^t \in (\frac{1}{2}, 1)$ if and only if either of the following conditions hold

$$\begin{array}{llllll} q \leq \frac{1}{2} & \textit{and} & \mu < \frac{qw-q}{2qw-q-w} & & \\ \frac{1}{2} < q & \textit{and} & w < q & \textit{and} & \frac{qw-q}{2qw-q-w} < \mu \\ \frac{1}{2} < q & \textit{and} & q < w & \textit{and} & \mu < \frac{qw-q}{2qw-q-w}. \end{array}$$

If, instead, $p^t \notin (\frac{1}{2}, 1)$, we have that either the competitive promotion scheme, or the seniority promotion scheme maximizes the expected ability of the promoted manager for all levels of noise $p \in (\frac{1}{2}, 1)$.

Lemma 2.4 states that if a specific threshold value, p^t, of precision exists, then for any precision below that threshold level one of the two schemes maximizes the expected ability of the promoted manager, whereas for any precision above the threshold, the other promotion schemes maximizes this expected ability. The threshold level exists if one of three conditions holds.

Let us briefly interpret these conditions: The first condition states that when $q \leq 1/2$—i.e., when q is relatively small—then μ must be sufficiently small, too, for the threshold value p^t to exist. If μ would be very high instead, then the seniority scheme is always relatively unattractive—i.e., competition is always preferred and hence no p^t exists.

The second condition states that when $q > 1/2$—i.e., q is relatively large—and $w < q$—i.e., ability difference is small relative to q—then we require μ to be sufficiently large for the threshold value p^t to exist. A large q in principle favors the seniority scheme. A small w further favors the seniority scheme. Therefore, if μ would be very small instead, the seniority scheme would be favored even more and, hence, the competitive scheme would never be optimal—i.e., no p^t would exist. The third condition is exactly reverse to the second, and so is its intuition.

Notice that the interpretation of the three conditions has a close connection to Proposition 2.2: First, whether the threshold level exists depends on the prior probability of the incumbent manager having high ability, q. Second, changes in q can, at least to some extend, be offset by appropriate changes in μ to maintain the existence of the described threshold, p^t. The result essentially maintains these ingredients while adding the additional parameter of w.

Although Lemma 2.4 is somewhat technical, it is crucial for our main result. When a partitioning value p^t exists, Lemma 2.4 states that on the interval $p \in (\frac{1}{2}, 1)$ there are two sub-intervals, one in which competition is preferred and one in which seniority is preferred. Whether high noise favors competition or seniority, however, is not immediately obvious. Indeed, the effect of noise is ambiguous and can favor either scheme, depending on other parameter values. The next proposition summarizes this.

Proposition 2.4. *Assume that the threshold noise level $p^t \in (\frac{1}{2}, 1)$ exists.*

1. *If $\mu < 1/2$, the seniority scheme maximizes the expected ability of the promoted manager if and only if signals are sufficiently noisy: $p < p^t$.*
2. *If $\mu > 1/2$, the seniority scheme maximizes the expected ability of the promoted manager if and only if signals are sufficiently precise: $p > p^t$.*
3. *If $\mu = 1/2$, noise does not affect which promotion scheme maximizes the expected ability of the promoted manager.*

We can see from Proposition 2.4 that noise plays an important role in the choice of the optimal promotion mechanism. The intuition, which we will unpack further below, relies heavily on Proposition 2.2, which has shown that the expected ability of the promoted worker can both decrease or increase in signal precision, and the direction depends on the share of high ability workers. Similarly, exactly how precision affects the optimal choice when the goal is to maximize the expected ability of the promoted manager depends crucially on the share of high ability workers. Specifically, if the share of high ability workers is small, then more noise favors the seniority scheme; if the share is large instead, more noise favors the competitive scheme. Figure 2.2 below illustrates this with two numerical examples. It shows the expected ability of the contest winner under seniority (blue line) and competition (red line) for $\mu < 1/2$ (Figure 2.2a) and for $\mu > 1/2$ (Figure 2.2b) for the stated parameter combinations. The blue shaded area indicates precision levels under which the seniority scheme maximizes the expected ability of the contest winner. The red shaded area corresponds to precision levels under which competitive promotions are preferred.

(a) $\mu = \frac{1}{4} < \frac{1}{2}$, $h = \frac{3}{2}$, $l = 1$, $q = \frac{1}{2}$

(b) $\mu = \frac{3}{4} > \frac{1}{2}$, $h = \frac{3}{2}$, $l = 1$, $q = \frac{3}{4}$

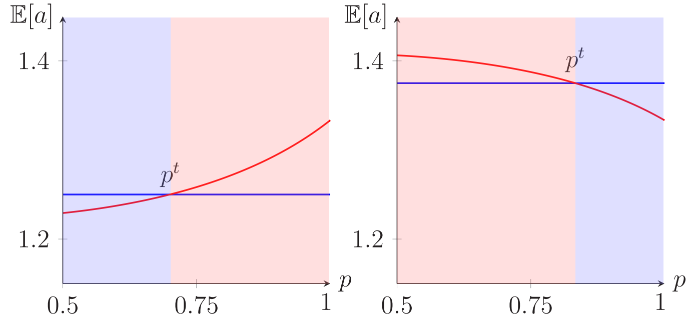

Figure 2.2: Expected ability of the contest winner under seniority (blue line) and competition (red line) for $\mu < 1/2$ (Figure 2.2a) and for $\mu > 1/2$ (Figure 2.2b). The cost function of high ability agents are $c_{a_i}(x_i) = x_i/a_i$ with $h = \frac{3}{2}$ for high ability players and $l = 1$ for low ability players. The blue shaded area indicates precision levels under which the seniority scheme maximizes the expected ability of the contest winner. The red shaded area corresponds to precision levels under which competitive promotions are preferred.

To further unpack the intuition, let us hark back to Proposition 2.3: First note that the seniority scheme is generally superior if q is sufficiently high. Because signal precision p partially determines the expected ability of the promoted worker in the previous stage, what constitutes a q that is "sufficiently high" depends on p as well. The reasoning of *how* this result depends on p is therefore very similar to the role that p plays in Proposition 2.2: A high ability incumbent manager will sometimes mistakenly promote a high ability worker (a "good" mistake), whereas a low ability incumbent manager will sometimes mistakenly promote a low ability worker (a "bad" mistake). The probability of making these mistakes depends on the share of high ability workers. And, even more crucially, the change in the mistake probability with a change in precision depends on the share of high ability workers. This means that, unless the ability distribution among the workers is symmetric, the change in the mistake probability that comes with a change in precision will be asymmetric. Precisely because the change in mistake probability is asymmetric in these cases, depending on an appropriate q, increasing precision can either increase or decrease the expected ability of the promoted manager under the competitive scheme. Because the expected ability of the promoted manager is fixed under the seniority scheme, this implies that noise can sometimes favor and sometimes disfavor the competitive scheme.

We can also look it this more formally by investigating the probabilities that a worker who produced a high signal has low ability and that a worker who produced a low signal has high ability. The corresponding expressions are

$$\Pr[a_i = l|\tilde{a}_i = h] = 1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}, \quad \text{and} \tag{2.4}$$

$$\Pr[a_i = h|\tilde{a}_i = l] = 1 - \frac{(1-\mu)p}{(1-\mu)p + \mu(1-p)}, \tag{2.5}$$

respectively, following from Bayes' rule as stated in expression (2.1). Differentiating (2.4) and (2.5) with respect to p allows us to compare how the mistake probabilities for both types of incumbent managers change when increasing precision.[12] For $\mu < 1/2$, we get

$$\frac{\partial \Pr[a_i = l|\tilde{a}_i = h]}{\partial p} < \frac{\partial \Pr[a_i = h|\tilde{a}_i = l]}{\partial p}, \tag{2.6}$$

In words, when the share of high-ability workers is low, the probability that a low-ability worker produces a high signal decreases more than the probability that a high-ability worker produces a low signal when precision increases. This result directly implies that, as *noise* increases, mistakes in the promotion of workers that are harmful to the organization become relatively more likely and, hence, more noise favors the seniority scheme. This is precisely the intuition behind the result stated in case 1 of Proposition 2.4.

The converse holds true when the share of high-ability workers is high ($\mu > 1/2$). As noise increases, it becomes relatively more likely for an incumbent manager to unintentionally promote a high-ability worker. Hence, less noise favors the seniority scheme in this case. Formally, this is implied by the observation that

$$\frac{\partial \Pr[a_i = l|\tilde{a}_i = h]}{\partial p} > \frac{\partial \Pr[a_i = h|\tilde{a}_i = l]}{\partial p}, \tag{2.7}$$

[12] The explicit derivatives are:

$$\frac{\partial \Pr[a_i = l|\tilde{a}_i = h]}{\partial p} = \frac{p\mu(2\mu-1)}{((1-p)(1-\mu)+p\mu))^2} - \frac{\mu}{(1-p)(1-\mu)+p\mu} < 0,$$

$$\frac{\partial \Pr[a_i = h|\tilde{a}_i = l]}{\partial p} = \frac{p(1-2\mu)(1-\mu)}{(p(1-\mu)+(1-p)\mu)^2} - \frac{1-\mu}{p(1-\mu)+(1-p)\mu} < 0.$$

when $\mu > 1/2$.

Finally, when the shares of low- and high-ability workers are the same ($\mu = 1/2$), both types of signals are equally informative and, hence, noise enters the expected ability of the contest-winner symmetrically. As a result, noise plays no role in this case. Formally,

$$\frac{\partial \Pr[a_i = l|\tilde{a}_i = h]}{\partial p} = \frac{\partial \Pr[a_i = h|\tilde{a}_i = l]}{\partial p}, \tag{2.8}$$

when $\mu = 1/2$.

Proposition 2.4, therefore, implies that whether an organization should use a competitive or a seniority-based promotion scheme—when its goal is to maximize the expected ability of the promoted manager—depends crucially on the interplay between how talent is distributed at the organization's bottom tier and how noisy the information transmission about abilities from the bottom to the middle tier is.

There are reasons for why one would either expect $\mu < 1/2$ or $\mu > 1/2$ in different scenarios. In one interpretation of our model, for example, we can think of the workers not as the lowest tier of the organization, but rather as a pool of outsiders from which hiring takes place. Then, this pool consists essentially of the entire population of job-seekers and therefore it is plausible to assume that ability—regarding the task at hand—is more likely to be low—i.e., $\mu < 1/2$. Instead, we can go with our initial assumption where workers were indeed thought of as the lowest tier within our organization. In many firms—especially in management consultancy, law, finance and "Big Tech" (Popomaronis 2019)—, it is common practice to undertake intense ability screenings even for the lowest tier of hires. It is plausible to assume that ability is likely to be high—i.e., $\mu > 1/2$—in such cases.

2.5 Effort and Ability Across the Hierarchy: a Dilemma

In this section, we highlight several trade-offs that the principal faces when choosing a promotion scheme, suggesting that the optimal choice depends on the particular organizational characteristics at hand.

So far, we have discussed that the principal might employ the different promotion schemes to maximize either talent on the middle tier or the highest tier of the organization. We have not commented on effort incentives yet. One additional benefit of contests that goes beyond selection power, however, is precisely that they incentivize effort.

Indeed, in the real world, organizations may not only care about an efficient allocation of talent, but also about incentivizing its members. From our previous analysis, however, it is easy to see that there may exist trade-offs between these differing objectives. Let us first consider a trade-off between incntivizing effort and the expected ability on the middle tier. The competitive promotion scheme always incentivizes the managers to exert effort (beyond some assumed base level of effort that one could justify through various assumptions—e.g., an intrinsic passion for the job or the fear of being fired if certain performance goals are not realized). The seniority scheme fails to do so. Hence, the competitive promotion scheme will always maximize effort on the management level of the organization. At the same time, Proposition 2.1 shows that expected abilities on the management level will always be maximized by the seniority scheme. This highlights a crucial dichotomy between selection efficiency and extrinsic motivation induced by competition.

Now let us consider the possible trade-offs the principal faces when she wants to maximize the expected ability on the top tier. From Proposition 2.4 we know that when $\mu < 1/2$ and $p < p^t$, this objective is achieved by employing the seniority scheme. If $p > p^t$ instead, the competitive scheme fares better. This also implies that when $p < p^t$, the objectives of maximizing both the expected ability of the promoted workers *and* the promoted managers are aligned. Instead, if $p > p^t$, the objectives of maximizing effort and maximizing the expected ability of the promoted managers are aligned. In the case where $\mu > 1/2$, the logic is reversed.

The following proposition summarizes these trade-offs.

Proposition 2.5. *Assume that p^t exists. Furthermore, suppose that the share of high-ability workers is low, $\mu < 1/2$.*

1. *If $p < p^t$, the principal maximizes both the expected ability of the promoted workers and the promoted managers by choosing the seniority scheme. She minimizes managers' expected effort.*

2. *If $p^t < p$, the principal maximizes both the expected ability of the promoted managers and the managers' expected effort by choosing the competitive scheme. She minimizes the expected ability of promoted workers.*

Now, suppose that the share of high-ability workers is high, $\mu > 1/2$.

3. *If $p < p^t$, the principal maximizes both the expected ability of the promoted managers and the managers' expected effort by choosing the*

competitive scheme. She minimizes the expected ability of promoted workers.

4. *If $p^t < p$, the principal maximizes both the expected ability of the promoted workers and the promoted managers by choosing the seniority scheme. She minimizes managers' expected effort.*

For the case where $\mu < 1/2$, part 1 of Proposition 2.5 points out that if the incumbent manager receives sufficiently noisy signals regarding the workers' abilities, there is a trade-off between the expected abilities of the promoted managers and managers' incentives to exert effort. Part 3 provides the analogous result for the case where $\mu > 1/2$. This can be problematic since most organizations depend on both talented employees rising to high levels of the organization and employees working hard. The optimal promotion mechanism, therefore, depends crucially on how much value a given organization puts on either of these objectives. In practice, this may be different for different organizations: Some firms may hire from a more homogeneous pool of applicants—for example, because the work requires a very specific education which reduces intrinsic differences in ability. If effective abilities are more homogeneous, it appears plausible that promoting a low-ability employee is less harmful to the firm as the difference between a low-ability and a high-ability employee is not as large. Therefore, such organizations may want to put more importance on incentivizing effort. This favors the competitive promotion scheme. Other organizations, however, may hire employees with strong intrinsic motivation—for example, NGOs (Leete 2000) and public sector employers (Finan, Olken, and Pande 2015). Extrinsic rewards may not lead to significant increase in effort.[13] Such organizations may favor seniority-based promotion mechanisms.[14]

[13]It may even crowd out effort as shown in Gneezy and Rustichini (2000) and Gneezy, Meier, and Rey-Biel (2011).

[14]There are other reasons for why organizations might face a trade-off between providing effort incentives and maximizing the expected abilities of the employees they promote. Tsoulouhas, Knoeber, and Agrawal (2007), for example, point out that in a contest between current employees of the organization and outsiders, whether the two objectives can be aligned depends crucially on how insider and outsider abilities differ. Similarly, Stracke et al. (2015) show that the trade-off exists for a principal who can choose between a dynamic and a static version of the same contest. Finally, Denter, Morgan, and Sisak (2022) establish the trade-off in a model where a newcomer can choose to signal their ability to the incumbent against whom they compete.

	p	μ	Expected ability of promoted worker	Expected ability of promoted manager	Managers' effort
Low precision, Low share of talent	$p < p^t$	$\mu < \frac{1}{2}$	seniority	seniority	competitive
Low precision, High share of talent	$p < p^t$	$\frac{1}{2} < \mu$	seniority	competitive	competitive
High precision, Low share of talent	$p^t < p$	$\mu < \frac{1}{2}$	seniority	competitive	competitive
High precision, High share of talent	$p^t < p$	$\frac{1}{2} < \mu$	seniority	seniority	competitive

Table 2.1: Optimal promotion scheme given signal precision and share of talent according to Proposition 2.5. The three objectives of the principal are (1) the expected ability of the promoted *worker*, (2) the expected ability of the promoted *manager* and (3) managers' effort.

For the case where $\mu < 1/2$, part 2 of Proposition 2.5 states that if the incumbent manager receives signals that are too precise regarding the workers' abilities, there is a trade-off between the expected abilities of the promoted workers and the promoted managers. Part 4 provides the analogous result for the case where $\mu > 1/2$. The result may seem counterintuitive at first. After all, the promoted workers constitute the pool of managers from which managers will be selected to rise to higher levels of the organization. It would seem logical that the expected ability of a promoted manager should increase with the expected ability of the pool of managers from which the promotion happens. To understand the logic that drives the surprising result, we want to point out that the competitive promotion scheme always maximizes the expected ability of the manager who is promoted when keeping the pool of managers fixed. In other words, and as pointed out in the introduction of this article, contests have value as talent-identification device. While the seniority scheme always maximizes the expected ability of promoted workers—i.e., the pool of contestants—if the incumbent manager receives too imprecise signals of the workers' abilities, she can make less strategic decisions, and, hence, the expected ability of promoted workers is less sensitive to the chosen promotion scheme. In such a case, the benefits of the competitive scheme as a talent-identification device outweigh the negative effect of creating a moral hazard problem.

This is a pessimistic result for organizations that are interested in allocating talented employees to both the top and the middle tiers of the hierarchy. Whereas it is not obvious how an organization should deal with this potential trade-off, recent evidence stresses the importance of good middle managers for firm performance (Mollick 2012; Giardili, Ramdas, and Williams 2022). In our framework, placing great importance on the quality of middle managers would generally favor the seniority scheme.

2.6 Discussion

In this section we want to discuss some implications of our paper that go beyond what we have analyzed formally. We will touch on three points, (i) within firm wage inequality, (ii) promotion rates, and (iii) incentives on the bottom tier.

Wage Inequality. Our results may have implications for the wage structure within the organization. Suppose, for example, that a principal has selected the seniority scheme. In this case, the selected promotion scheme will not incentivize effort beyond some baseline level on the middle tier of the organization. As a result, the principal is not benefiting from a steep wage increase from the middle to the top tier. This suggests that the seniority scheme will result in a relatively low wage inequality between the middle and the top tier compared to the competitive scheme.

Additionally, the seniority scheme may affect the wage inequality between the bottom and the middle tier. Suppose that competing organizations are able to observe the promotion scheme chosen by our organization, or, almost equivalently, receive an informative signal about it. Then, competing organizations will reason that the expected ability of the middle tier of our organization is relatively high. This will yield higher outside options for agents employed on the middle tier of our organization. As a result, if we maintain the assumption that our organization attempts to retain its members, it will have to increase the wage on the middle tier. Because the seniority scheme has no obvious implication for the optimal wage on the bottom tier, this implies that wage inequality between the bottom and the middle tier will be relatively large compared to the competitive scheme.

Promotion Rates. Our analysis has often highlighted the importance of precision for the choice of the optimal promotion scheme. This importance suggests an obvious question: Can the principal sometimes affect this precision? In the real world, the signal is likely to arise due to the middle manager observing her subordinates over time. Plausibly, therefore, we would expect the precision of the signal to increase if the middle manager observes her subordinates for a longer time. One way for the principal to control this would be for her to implement rules on minimum or maximum tenures at the worker level before a promotion to the next tier is possible—for example by imposing an up-or-out rule. By introducing a minimum tenure, the principal could increase precision, and conversely, by introducing a maximum tenure, she could decrease precision. As we have shown, the principal might sometimes benefit, and sometimes suffer from an increase in precision, and as such, she may indeed have an interest to implement such rules.

Bottom Tier Incentives. So far, we have only considered incentives on the organizations middle tier and have stayed agnostic about how these affect the behavior of workers on the bottom tier. This is because we have assumed their signal to arrive exogenously. The main intuition behind this assumption is that the work on the bottom tier of the hierarchy often requires a fundamentally different set of skills compared to that of work on the management level. Specifically, management skills that contribute to success at higher levels of the organization matters little in producing output as a worker. It follows that organizing a promotion contest on the bottom tier would most likely not generate much useful information about workers' abilities in the dimension relevant on the middle tier of the organization.

Nevertheless, even performance on the bottom tier might sometimes serve as a weak signal for ability on the middle tier and, hence, the middle manager might still want to organize a promotion contest on the bottom tier. Suppose that she indeed organizes such a contest in addition to the ability signal we have considered throughout the paper. How the workers respond to this depends crucially on their beliefs about the middle manager's ability. If they believe the middle manager to have low ability, they would infer that she is trying to promote a high ability worker, and therefore the workers would have an incentive to win the contest. In turn, workers will work. The final signal the middle manager receives about workers' ability is then a combination of their contest performance and the exogenous signals. Suppose, instead, that the workers believe the middle manager to have high ability. In this case, they would infer that the middle manager wants to promote a low ability worker. This provides perverse incentives to *lose* the contest and, as a result, no worker would work. The final ability signal is then solely provided by the exogenous signal.

Overall, this argument shows even more the severity of the moral hazard problem we consider. Not only do competitive promotion schemes induce a middle manager to promote untalented workers; additionally, workers now also want to appear untalented. As a result, the spillovers of the competitive promotion scheme extend to the bottom tier, where effort provision will now be inefficiently low.

2.7 Conclusion

Contests are often regarded as meritocratic because they enable talented players to rise through the ranks. In this paper, we have incorporated a standard all-pay contest into a hierarchical organization. We have shown that the meritocratic value of contests may be obscured: Implementing promotion competitions on the middle level of an organization can create negative spillovers to the management of lower levels, ultimately preventing the promotions of talented employees. Indeed, in some settings contests may help mediocre players to reach the top of the organizational hierarchy. Recent industry surveys highlight the extent of this moral hazard problem, implying the need for research on how organizations can deal with this issue that could negatively impact long-term productivity. We have shown that reducing the competitiveness of the promotion mechanism and basing it on seniority instead can fully alleviate any perverse incentives to promote mediocre employees.

Somewhat more pessimistically, however, we have also identified a novel trade-off that occurs in some settings: The policy which maximizes the expected ability of players allocated to the organization's middle tier can be misaligned with the policy that maximizes the expected ability of players allocated to the top tier. This trade-off hints at a deeper question regarding our understanding of meritocracy. Indeed, whether a policy that blocks talented players from reaching the middle tiers of an organization—even if, conditional on reaching the middle tier, it helps them to reach the top—can be regarded as meritocratic is perhaps more philosophical in nature.

Beyond this, our results make us wonder whether organizations in practice employ seniority-based promotion schemes as a means of alleviating moral hazard problems. Whereas this motive is hard to identify empirically, certain types of organizations where the moral hazard problem is likely to occur do indeed seem to be structured based on seniority—e.g., political parties (Besley et al. 2017; Cirone, Cox, and Fiva 2021). Further (empirical) research in this direction can yield important insights for the design of efficient organizations.

2.A Appendix

2.A.1 Equilibrium in Two-Player All-Pay Auctions with Very Large Head Start

If bias is large enough, the equilibrium derived in Kawamura and Barreda (2014) does not hold any longer. Note that for at least one of the players, the upper bound of the equilibrium effort distribution is strictly decreasing in b for the case when:

1. bias favors the high ability player: $c_l^{-1}(v) - b$ for $F_h^*(x_h)$,

2. large bias favoring the low ability player: $c_h^{-1}(v) - b$ for $F_l^*(x_l)$,

3. and bias is favoring player $i \neq j$ when $a_i = a_j$: $c_i^{-1}(v) - b$ for $F_i^*(x_i)$.[15]

Consequently, the equilibrium prescribes randomization over some strictly negative effort levels when bias is large enough. Explicitly, we need to extend the solution to cover $b > c_l^{-1}(v)$ when the bias favors the high ability player, and $b > c_h^{-1}(v)$ when the bias favors the low ability player (same thresholds apply for homogeneous ability pairs). We do this in the following lemma:

Lemma 2.5. *In case of heterogeneous abilities, (1) if bias favors the high ability player, for very large bias $b \geq \bar{b} \equiv c_l^{-1}(v)$, both players submit zero effort in equilibrium and the high ability player wins with probability one. (2) If bias favors the low ability player, for very large bias $b \geq \bar{b} \equiv c_h^{-1}(v)$, both players submit zero effort in equilibrium and the low ability player wins with probability one.*[16]

[15]In Kawamura and Barreda (2014), there is no separation of this case since the initial assumption covers the homogeneous cases (results from bias favoring the high ability player can be used): $c_h(\cdot) \leq c_l(\cdot)$. Here I present the homogeneous cases separately for clarity.

[16]Note, this corresponds to the degenerate equilibrium CDF in Kawamura and Barreda (2014) when $b = \bar{b}$.

If abilities are homogeneous, then the bias threshold above which both players submit zero effort and the favored player wins with probability one becomes $\bar{b} \equiv c_i^{-1}(v)$*, with* $c_i(\cdot) = c_{-i}(\cdot)$ *for players* i *and* $-i$*.*

Proof. In all cases, bias above $\bar{b}$ is so large that for any feasible effort level of the player with lower "reach", the favored player will strictly outscore her by submitting an effort of zero. It is then straightforward to see that the player with lower "reach" will submit an effort of zero to minimize her effort costs, while the favored player best-responds by also submitting zero effort. The score of the favored player is then strictly greater than her opponent's, and therefore she wins the contest with probability one in equilibrium. □

2.A.2 Extension to Positive Expected Penalties

We may assume, similarly to Haegele (2022), that the incumbent manager faces a cost $k > 0$ from choosing a worker who produced a low signal. This cost may reflect difficulty for the incumbent in presenting a low signal candidate to other managers involved in the promotion decision (if this decision is not made alone by the incumbent). If the principal has the power to punish the manager who made the decision, k reflects the expected punishment associated with being found to act against the principal's best interest.

Although we allow for $k \neq 0$, we restrict our attention to cases where expected penalty is small. We argue in the main text that we can consider the expected penalty to be zero. Even if we allow for this term to be nonzero, our arguments remain relevant, hence we may consider the size of the expected penalty to be vanishingly small. We show that when this is the case, we approximate the main results derive in the main text, even though this assumption makes our problem intractable. To do so, we first define the set of "intermediate" biases—biases that do not guarantee a win for either players—that eliminates the moral hazard. We then show that the lower bound of this set of intermediate biases equals $c_h^{-1}(v)$ in the limit when $k \to 0$. The limit equals the smallest bias that *guarantees* a win for the incumbent. We also show that the expected ability of the promoted manager equals that of the full seniority scheme in the limit when $k \to 0$.

Intermediate Biases. When punishment is possible, such that $k > 0$, the principal may not be restricted to a promotion scheme that is strategically equivalent to the full seniority promotion scheme if she wants to eliminate moral hazard. She may set bias which leads to strictly positive win probability for both competing players, but still guarantees that the incumbent will promote a high signal worker. Let the lowest bias in the set of biases that eliminates moral hazard but induces effort provision[17] be defined

$$b^{\min} \equiv \min\{b : \tilde{a}_{-i} = h, b < c_h^{-1}(v)\}.$$

Note, setting bias to $b^{\min} \leq b < c_h^{-1}(v)$ may or may not be preferred to the full seniority scheme by the principal when she maximizes the expected ability of the promoted manager.[18]

We show that the problem of finding $b^{\min}$ boils down finding the lowest b, such that a low ability incumbent will want to promote a high signal worker. First, note that if a fixed bias b' makes a low ability incumbent promote a high signal worker, then the same bias b' will also guarantee that a high ability incumbent makes the same promotion decision. This is, however, no longer holds in the reverse case: if b'' guarantees that a high ability incumbent selects a high signal worker, it may not guarantee that a low ability incumbent will make the same promotion decision.

The intuition is simple: a high ability player always has weakly greater win probability against *any* player than a low ability player, conditional on holding bias fixed for both types of incumbent players. This result is of course elementary when $b = 0$, but remains true with $b > 0$. To see this, recall that win probability is weakly increasing in a player's reach $c_{a_i}^{-1}(v + b)$ (the maximum willingness to pay for the prize). Strict ability difference requires $c_h(x) < c_l(x) \ \ \forall \ \ x > 0$, which implies $c_h^{-1}(x) > c_l^{-1}(x) \ \ \forall \ \ x > 0$. From here we see that the reach must also be greater for the high ability player as $c_h^{-1}(v + b) > c_l^{-1}(v + b)$, and therefore the win probability must also be weakly greater for him compared to that of a low ability player when b is held constant across the two players.

Second, also note that guaranteeing a win against a high ability player requires higher bias compared to guaranteeing a win against a low ability player. We then have that finding $b^{\min}$ is equivalent to finding the lowest b, such that a low ability incumbent will want to promote a high signal worker. This restricts our candidate biases to $b \in \Big[c_h^{-1}(v) - c_l^{-1}(v), c_h^{-1}(v)\Big)$. The upper bound is given by the lowest bias the guarantees a win for the low ability incumbent. We are interested in biases that are "intermediate" in the

[17]Bias is such that both players have a strictly positive probability to win in equilibrium, therefore their respective equilibrium CDF does not put all mass to zero.

[18]Consider the case when the probability that incumbent manager is high ability is close to (or equals) one. Guaranteeing the promotion of this incumbent manager will then lead to a higher expected ability of a promoted manager than another promotion scheme which leads to a strictly positive probability that a promotee (who may be of low ability) wins the biased promotion contest.

sense that they imply strictly positive win probability for both competing players in equilibrium, therefore $b = c_h^{-1}(v)$ is not in the set of candidate biases.

We can rewrite the sufficient condition for the incumbent to promote a high signal worker, such that now we account for a positive expected penalty:

$$\begin{aligned}
&\overbrace{\frac{\mu p}{\mu p+(1-\mu)(1-p)}}^{\Pr[a_{-i}=h \mid \tilde{a}_{-i}=h]}(\Pi_{a_i} \mid b, a_{-i}=h) \\
&\quad+\overbrace{\left(1-\frac{\mu p}{\mu p+(1-\mu)(1-p)}\right)}^{\Pr[a_{-i}=l \mid \tilde{a}_{-i}=h]}(\Pi_{a_i} \mid b, a_{-i}=l) \\
&\geq \underbrace{\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)}}_{\Pr[a_{-i}=l \mid \tilde{a}_{-i}=l]}(\Pi_{a_i} \mid b, a_{-i}=l) \\
&\quad+\underbrace{\left(1-\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)}\right)}_{\Pr[a_{-i}=h \mid \tilde{a}_{-i}=l]}(\Pi_{a_i} \mid b, a_{-i}=h)-k.
\end{aligned}$$

We can simplify this expression to get a more general version of Equation 2.2—we get back Equation 2.2 by setting $k = 0$. Replacing the weak inequality sign with a strict equality provides us with the equation that implicitly defines $b^{\min}$

$$\frac{k}{\beta} = \left(\Pi_l \mid b^{\min}, a_{-i}=l\right) - \left(\Pi_l \mid b^{\min}, a_{-i}=h\right),$$

where

$$\beta \equiv \underbrace{\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)}}_{=\Pr[a_{-i}=l \mid \tilde{a}_{-i}=l]} + \underbrace{\frac{\mu p}{\mu p+(1-\mu)(1-p)}}_{=\Pr[a_{-i}=h \mid \tilde{a}_{-i}=h]} -1 > 0.$$

When $b^{\min} \in \left[c_h^{-1}(v) - c_l^{-1}(v), c_l^{-1}(v)\right]$, from Lemma 2.1 we have that

$$\begin{aligned}
\frac{k}{\beta} &= v - c_l\left(c_l^{-1}(v) - b^{\min}\right) - v + c_l\left(c_h^{-1}(v) - b^{\min}\right) \\
\frac{k}{\beta} &= c_l\left(c_h^{-1}(v) - b^{\min}\right) - c_l\left(c_l^{-1}(v) - b^{\min}\right) \qquad \Big(> 0\Big).
\end{aligned}$$

Note that

$$\frac{\partial\left[c_l\left(c_h^{-1}(v)-b\right)\right]}{\partial b} = -c_l'\left(c_h^{-1}(v)-b\right) < \frac{\partial\left[c_l\left(c_l^{-1}(v)-b\right)\right]}{\partial b} = -c_l'\left(c_l^{-1}(v)-b\right),$$

since $c_l^{-1}(v) < c_h^{-1}(v)$ and $c_l'' > 0$. This implies that the first term on the right hand side (RHS) decreases faster, when increasing bias, hence the RHS decreases as we increase bias. Using this observation, we also have that when k is larger, the difference on the RHS needs to be larger as well. This implies that $b^{\min}$ is lower: $\partial b^{\min}/\partial k < 0$.

If $b^{\min} \in \left[c_l^{-1}(v), c_h^{-1}(v)\right)$, then $b^{\min}$ is given by

$$\begin{aligned}
\frac{k}{\beta} &= v - v + c_l\left(c_h^{-1}(v) - b^{\min}\right)\\
\frac{k}{\beta} &= c_l\left(c_h^{-1}(v) - b^{\min}\right)\\
c_l^{-1}\left(\frac{k}{\beta}\right) &= c_h^{-1}(v) - b^{\min}\\
b^{\min} &= c_h^{-1}(v) - c_l^{-1}\left(\frac{k}{\beta}\right).
\end{aligned}$$

Having an explicit expression for $b^{\min}$ allows us to check some comparative statics. First, we can show that $b^{\min}$ is increasing with prize v, given that $c_h' > 0$ and the inverse of a continuous monotone increasing function is also increasing:

$$\frac{\partial b^{\min}}{\partial v} = \frac{\partial c_h^{-1}(v)}{\partial v} > 0.$$

This is quite intuitive; for fixed expected punishment, increasing v increases returns from promoting a low signal worker. The converse holds true when increasing punishment k, $\partial b^{\min}/\partial k < 0$.

$$\frac{\partial b^{\min}}{\partial k} = -\left(c_l^{-1}\right)'\left(\frac{k}{\beta}\right) < 0 \ \ \forall\, k > 0.$$

Given that we have shown already that $\partial b^{\min}/\partial k < 0$ when $b^{\min} \in \left[c_h^{-1}(v) - c_l^{-1}(v), c_l^{-1}(v)\right]$, the same result for $b^{\min} \in \left[c_l^{-1}(v), c_h^{-1}(v)\right)$ implies that $b^{\min}$ decreases in k for all values of $b^{\min}$.

Furthermore, we have that $b^{\min}$ is increasing in β, the signal precision accounting for the share of high ability players in the population.

$$\frac{\partial b^{\min}}{\partial \beta} = \frac{k}{\beta^2}\left(c_l^{-1}\right)'\left(\frac{k}{\beta}\right) > 0.$$

The principal has to bias the contest more in favor of the incumbent to make the prospect of promoting a low signal worker less profitable when the incumbent is more certain that choosing a low signal worker will indeed lead to competing against a low ability worker.

Intermediate Bias in the Limit. We show that if the principal wants to eliminate the moral hazard, biases that allow her to do so will get arbitrarily close to $c_h^{-1}(v)$ as expected penalty k goes to zero. In other words, when we compare the competitive promotion scheme that induces moral hazard with those promotion schemes that eliminate moral hazard, we can focus on the full seniority promotion scheme knowing that all alternatives will approximate this case as expected penalty goes to zero. This is precisely the comparison with consider in the main text, hence this result allows us to argue that our analysis is relevant in this more general framework.

Lemma 2.6. *When expected penalty is low such that $k \to 0$, the lowest bias $b^{\min}$ that guarantees the incumbent's promotion of a high signal worker approaches $c_h^{-1}(v)$.*

Proof. We know that $\partial b^{\min}/\partial k < 0$, hence when k is small, the minimum required bias to remove the moral hazard is large.

When k is so small that we need to set a bias $c_l^{-1}(v) < b$, we know that $b^{\min}$ with $k \neq 0$ is in the interval $\left[c_l^{-1}(v), c_h^{-1}(v)\right)$. In this case we have an explicit expression for $b^{\min}$,

$$b^{\min} = c_h^{-1}(v) - c_l^{-1}\left(\frac{k}{\beta}\right).$$

We can easily take the limit as $k \to 0$,

$$\lim_{k\to 0} b^{\min} = \lim_{k\to 0} \quad c_h^{-1}(v) - c_l^{-1}\left(\frac{k}{\beta}\right) = c_h^{-1}(v).$$

□

Expected Ability in the Limit. Note that when $b^{\min}$ does not guarantee a win for the incumbent, the expected ability of the promoted manager is given by

$$\mathbb{E}\left[a_i \mid x_i > x_{-i}, b = b^{\min}\right] = q\Gamma + (1-q)\Xi,$$

with

$$\Gamma := \frac{\mu p}{\mu p + (1-\mu)(1-p)} h + \left(1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}\right)\left[w_h(b^{\min})h + (1 - w_h(b^{\min}))l\right]$$

and

$$\Xi := \frac{\mu p}{\mu p + (1-\mu)(1-p)}\left[w_l(b^{\min})l + (1 - w_l(b^{\min}))h\right] + \left(1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}\right) l,$$

where we introduce new notation $w_h(b)$ and $w_l(b)$, with $w_h(b) > \frac{1}{2}$ the win probability of a high ability incumbent with bias b against a low ability player. The win probability of a *low* ability incumbent with bias b against a *high* ability player is given by $w_l(b)$.

When expected penalty $k \to 0$, we know that $b^{\min} \to c_h^{-1}(v)$ from Lemma 2.2, and therefore, both $w_h(b^{\min}) \to 1$ and $w_l(b^{\min}) \to 1$. We then have that

$$\lim_{k \to 0} \mathbb{E}\left[a_i \mid x_i > x_{-i}, b = b^{\min}\right] = \mathbb{E}\left[a_i \mid x_i > x_{-i}, b \geq c_h^{-1}(v)\right] = qh + (1-q)l.$$

In other words, the expected ability of the promoted manager with bias $b^{\min} < c_h^{-1}(v)$ approaches the expected ability of the promoted manager under the full seniority scheme when expected penalty goes to zero. Even though we cannot analyse when full competition dominates the promotion scheme that solves the moral hazard problem if we allow for $b^{\min}$ that does not guarantee the incumbent's win, we can still conjecture that our results from Lemma 2.4 onward hold for small enough expected penalty k.

2.A.3 Extension to Infinitely-Lived Organizations

We now consider a version of our model in which the organization expects to be infinitely-lived and there is no discounting. The main purpose of this exercise is to show that the intuition behind the main result derived in section 2.4 carries over to a dynamic setting: Perhaps surprisingly, there exist (reasonable) parameter combinations for which an uncompetitive seniority-based promotion scheme helps to allocate talented employees to the top of an organization. This result is not immediately obvious as, in the dynamic model, noise does not only affect the expected ability of the manager promoted under the competitive scheme, but also the expected ability of the manager promoted under the seniority scheme.

Before we proceed, we impose some additional assumptions: We assume that each player lives *at most* two active periods. As active periods, we count all periods in which a player is in the role of manager and therefore has to make strategic decisions—specifically whom to promote and/or how much effort to invest in the competition against this person. We assume that if a player wins the competition between the two managers in the first active period of her life, she collects the prize and leaves the game. Hence, such a player has only lived for one period. Instead, if she does not win the competition in her first active period, she receives nothing. Thereafter, she chooses whom to promote to her level and has a second attempt at collecting the prize—this time by competing against her own promotee. If she wins in the second period, she collects the prize and leaves the game. If she loses in the second period as well, she collects nothing and leaves the game. If at some point during the game both managers leave at the same time, we assume that two workers are randomly selected for promotion to fill the two vacant positions. Allowing players to live either one or two periods helps us to abstract away from the uninteresting case in which a player promotes a worker, knowing that she will not compete against her because she will retire before the competition stage. The moral hazard problem we are interested in is clearly not a concern in this scenario.

Before we state the formal result, we consider the logic of the perfectly competitive case. Like in the previous subsection, we assume that we start with an incumbent manager who has expected ability $qa_h + (1 - q)l$. For clarity, we also assume that the incumbent manager is about to start the second period of her active life, although this does not affect our results. As a benchmark, we consider a scenario without noise—i.e., $p = 1$. It is useful to briefly think about how the game could proceed over the following periods. In the first period, a low ability incumbent manager promotes a high-ability worker. If instead, the incumbent manager has high ability, she promotes a low-ability worker. In the next period, if the winner has high ability, she will promote a low-ability worker. If the winner has low ability, she will promote a high-ability worker. In the case in which both managers have left the game, of course, the next two workers' abilities will be determined by a random draw. The main take away from this consideration is that, in all cases except the last, the contest will be balanced in the sense that there will always be one high-ability manager competing against one low-ability manager. Hence, in these cases, the expected ability of the winner will be the same as the expected ability from the contest winner as considered in Section 3,

$$wa_h + (1 - w)l. \tag{2.9}$$

The remaining question is what happens when signals are noisy instead—i.e., $p \in (1/2, 1)$. Deriving an explicit expression for the expected ability of any period winner in this case is an intractable task. We will, instead, define the average expected ability of all promoted managers from the perspective of the infinitely lived firm as

$$\Psi(p) := \lim_{T \to \infty} \left\{ \frac{\sum_{t=1}^{T} \mathbb{E}\left[a_{i,t} \mid x_{i,t} > x_{-i,t}\right]}{T} \right\} \tag{2.10}$$

where T denotes the number of periods and $\mathbb{E}\left[a_{i,t} \mid x_{i,t} > x_{-i,t}\right]$ is a function of p. We are interested in how this expression changes with p. Let us first consider the case in which the share of high-ability workers is small—i.e., $\mu < 1/2$. In this case, $\Psi(p)$ is increasing in p. The reason is as follows: Both types of incumbent managers want to promote a worker of the opposite type. Incorporating noise implies that they will sometimes accidentally promote a worker of their own type. This is more harmful to the organization in the case where a low-ability incumbent manager makes the mistake. However, since the incumbent manager who makes the promotion decision is the one who previously lost the contest, and since a low-ability manager will, on expectation, lose the contest more often, low-ability managers will be in the position of making promotion decisions more often. Therefore, of the two possible types of mistakes, a low-ability incumbent manager unintentionally promoting a low-ability worker will happen more often, which reduces the expected ability of the period winner compared to **??**. The expected ability of the period winner is unaffected by noise in any period in which the competitors are a random draw from the pool of workers.

In the case in which the share of high-ability workers is high, $\mu > 1/2$, the effect of p on Ψ is ambiguous. This is because the effect described for the opposite case still holds. Additionally, however, we have to consider the change in relative signal informativeness described in the previous section: when μ is high, high signals provide relatively more information then low signals and this effect is exacerbated in noise. This means that with higher noise, low-ability incumbent managers will become *relatively* more likely to succeed in their attempt to promote a high-ability worker.

We proceed by discussing the expected ability of the promoted manager when the seniority scheme is used. Again, we assume that the expected ability of the incumbent manager is given by $qa_h+(1-q)l$, but the importance of q vanishes because the organization is assumed to be infinitely-lived. In the first period, the incumbent manager, independent of her own ability, intends to promote a high ability worker. She succeeds with probability

$$\frac{\mu p}{\mu p + (1-\mu)(1-p)}. \tag{2.11}$$

Thereafter, she will be promoted to a higher position given her seniority. Now her promotee is tasked with promoting a worker to the vacant management position. Since she has no cost of promoting a high-ability worker, but faces costs k for promoting a worker who produced a low signal, she, again, intends to promote a high-ability worker and succeeds with the same probability as her predecessor. Thereafter, she is promoted herself to the higher position. By repeating this argument indefinitely, we can establish that the average expected ability as $T \to \infty$ is given by

$$\frac{\mu p}{\mu p + (1-\mu)(1-p)} a_h + \left[1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}\right] l \tag{2.12}$$

Using these observations, we are ready to state the following proposition.

Proposition 2.6.

1. *For any $p > 1/2$, the expected ability of the worker who fills the vacant management position is higher under the seniority scheme compared to the fully competitive promotion scheme.*

2. *The expected ability of the promoted manager is higher under the seniority scheme compared to the fully competitive scheme for any $p \in \mathcal{P} \subseteq [1/2, 1]$ such that*

$$\Psi(p) < \frac{\mu p}{\mu p + (1-\mu)(1-p)} h + \left[1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}\right] l. \tag{2.13}$$

3. *We have that $\mathcal{P} \neq \emptyset$ and $\mathcal{P} \neq \{1\}$.*

Proof: Parts 1 and 2 follow from the arguments made in the main text. To see that part 3 is true, note the following: When signals are noiseless ($p = 1$), the expected ability of any manager who is promoted under the seniority scheme is given by h. When p decreases, so does expected ability—continuously. Instead, let us look at the expected ability of the contest winner from the fully competitive scheme, again when signals are noiseless. In any period in which the contest takes place after only one worker was promoted, the expected ability of the period winner is given by

$$wh + (1 - w)l. \tag{2.14}$$

In any period in which the contest takes place after two workers have been promoted, the expected ability of the contest winner depends also on μ. In this case, explicitly, expected ability of the winner is:

$$\mu^2 h + (1 - \mu)^2 l + 2\left(\mu - \mu^2\right)[wh + (1 - w)l],$$

where μ^2 is the probability that both workers who are promoted at the same time are of high ability, $(1 - \mu)^2$ is that they are both of low ability and $2(\mu - \mu^2)$ is the probability that they are of different ability, in which case the expected ability of the winner is given, again, by expression (2.14). The crucial part, however, is that there is always a discrete difference between the expected ability of the winner of the contest in any given period. Now, consider what happens when we add noise. In the case where $\mu < 1/2$, noise decreases the average expected ability of the contest winner. This follows from arguments made in the main text. Similarly, it follows from the main text that, when $\mu > 1/2$, the direction of the effect of noise is ambiguous. Crucially, however, in both cases, the average expected winner is continuous in noise. This is because the average expected ability is continuous in the expected abilities of both competitors of the contest and the expected abilities of the competitors are continuous in noise. This follows from the expressions derived using Bayes' rule. By continuity of the expected abilities of the promoted managers under both the contest and the seniority scheme follows that there exists some non-empty set of values p for which (2.13) holds (i.e., $\mathcal{P} \neq \emptyset$) and that this set contains cases with strictly positive noise (i.e., $\mathcal{P} \neq \{1\}$). □

Proposition 2.6 states that the seniority scheme is always preferred by the principal if her objective is to maximize the expected ability of the worker selected to fill the vacant management position (Part 1)—i.e., it implies that Proposition 2.1 carries over to the game with infinitely-lived firms. Parts 2 and 3 state that the seniority scheme is preferred if the signals about workers' abilities received by the incumbent managers are sufficiently precise and her objective is to maximize the expected ability of the promoted manager —i.e., the main insight from Proposition 2.4 carries over to the game with infinitely lived firms: Making promotions from the middle to the top tier uncompetitive can increase the ability of those managers who are promoted. The intuition is as follows: The seniority scheme guarantees that the incumbent managers always *want* to promote a high-ability worker. Therefore, if the signals are noiseless, it is obvious that the seniority scheme is superior at identifying talent. When the signals become less precise, however, incumbent managers will sometimes promote low-ability workers even in the seniority scheme. Given that sometimes low-ability workers manage to be promoted to a management position, the competitive scheme becomes relatively more valuable as a mechanism to identify talent.

One difference compared to Proposition 2.4 is that in the one-period model, there exist parameter combinations such that the threshold above which the seniority scheme is optimal is above one. For such parameter combinations, hence, the seniority scheme can never be optimal. This is different in the dynamic model with infinite periods. This difference is driven by the optimal policy's dependence on the principal's beliefs regarding the first period incumbent manager's ability in the one-period model. For example, when the principal believes with probability one that the incumbent manager has low ability, the seniority scheme clearly cannot be optimal. In the model with an infinitely-lived firm, however, the importance of these beliefs vanishes.

Furthermore, for other parameter combinations, Proposition 2.4 yields the result that the seniority scheme is always optimal. Due to the mathematical intractability of the problem, we are unable to demonstrate whether this is the case for infinitely-lived firms too. It is important to note, however, that the version of the model with infinitely-lived firms we consider is without discounting. The one-period model instead, can be interpreted as a special case of the model with an infinitely-lived firm that cares only for the outcome of the first period. Any other scenario is an intermediate case between these two cases. As such, we expect that the behavior of the threshold value is intermediate too. The logic is that the importance of the first-period beliefs vanishes in the model without discounting. With sufficient discounting, however, the importance of these beliefs increases and, hence, they play a role in the design of the optimal policy.

2.A.4 Proofs

Proof of Lemma 2.1

Follows from Kawamura and Barreda (2014) and Lemma 2.5 in subsection 2.A.1. □

Proof of Lemma 2.2

First, take the sufficient condition to promote a high signal worker:

$$(\Pi_{a_i} \mid b, a_{-i} = h) \geq (\Pi_{a_i} \mid b, a_{-i} = l)$$

Now note that the expected payoff from competing against a player of low ability is always weakly greater than the expected payoff from competing against a high ability opponent,

$$(\Pi_{a_i} \mid b, a_{-i} = l) \geq (\Pi_{a_i} \mid b, a_{-i} = h)\,.$$

These weak inequalities can only hold at the same time if

$$(\Pi_{a_i} \mid b, a_{-i} = l) = (\Pi_{a_i} \mid b, a_{-i} = h)\,.$$

From Lemma 2.1, we can see that these expected payoffs equal if and only if the incumbent is guaranteed to win against a player of any ability type, i.e. bias is such that

$$c_h^{-1}(v) \leq b.$$

The strategic equivalence between $c_h^{-1}(v) \leq b$ and $b \to \infty$ follows from Lemma 2.5. This result holds for both types of incumbents. □

Proof of Lemma 2.3.

Equilibrium Promotion Decision under Full Seniority ($b \to \infty$). When promotion is guaranteed for the incumbent manager, her payoff is unaffected by her promotion decision. Since we assume that the incumbent promotes a high signal worker if she is indifferent between promoting a high signal worker and promoting a low signal worker, it follows that she promotes a high signal worker in equilibrium.

Equilibrium Promotion Decision under Full Competition ($b = 0$). We first look at the high ability incumbent managers' incentives. The sufficient condition for a high ability incumbent to promote a high signal worker is given by

$$(\Pi_h \mid b = 0, a_{-i} = h) \geq (\Pi_h \mid b = 0, a_{-i} = l)\,.$$

Note that when abilities are homogeneous ($a_{-i} = h$) and $b = 0$, the incumbent's expected ability will be zero. Using this observation and replacing $(\Pi_h \mid b = 0, a_{-i} = l)$ with the appropriate expression from Lemma 2.1, we have

$$\begin{aligned} 0 &\geq v - c_h\left(c_l^{-1}(v)\right) \\ c_h\left(c_l^{-1}(v)\right) &\geq v. \end{aligned}$$

This inequality is never satisfied given strict ability difference, hence we know that a high ability incumbent will promote a low signal worker when $b = 0$.

It is clear from Lemma 2.1 that a low-ability incumbent manager earns an expected payoff of zero regardless of her opponent's ability in an unbiased contest ($b = 0$), hence she is indifferent between promoting a high signal or a low signal worker. It follows then that she promotes the worker who produced the high signal. □

Proof of Proposition 2.1

Follows immediately from Lemma 2.3. □

Proof of Proposition 2.2

To prove part (1) differentiate the expected ability under the seniority scheme

$$\frac{\mu p}{\mu p + (1-\mu)(1-p)} \cdot h + \left(1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}\right) \cdot l$$

with respect to p and μ to get

$$\begin{aligned} \frac{\partial \mathbb{E}[a_i]}{\partial p} &= (h-l)\frac{(1-p)p}{((2\mu-1)p+1-\mu)^2} > 0, \\ \frac{\partial \mathbb{E}[a_i]}{\partial \mu} &= (h-l)\frac{(1-\mu)\mu}{((2\mu-1)p+1-\mu)^2} > 0. \end{aligned}$$

To prove part (2), consider the following: Under the competitive scheme, the expected ability of the promoted worker from the perspective of the principal is given by

$$\begin{aligned} &q\left[\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)} \cdot l + \left(1 - \frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)}\right) \cdot h\right] \\ &+ (1-q)\left[\frac{\mu p}{\mu p + (1-\mu)(1-p)} \cdot h + \left(1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}\right) \cdot l\right], \end{aligned}$$

taking the derivative with respect to p yields

$$q\left[\left(\frac{\partial \Pr[a_i = h|\tilde{a}_i = l]}{\partial p}\right) h - \left(\frac{\partial \Pr[a_i = h|\tilde{a}_i = l]}{\partial p}\right) l\right]$$
$$+ (1-q)\left[\left(\frac{\partial \Pr[a_i = l|\tilde{a}_i = h]}{\partial p}\right) l - \left(\frac{\partial \Pr[a_i = l|\tilde{a}_i = h]}{\partial p}\right) h\right],$$

where

$$\frac{\partial \Pr[a_i = l|\tilde{a}_i = h]}{\partial p} = \frac{p\mu(2\mu-1)}{((1-p)(1-\mu)+p\mu))^2} - \frac{\mu}{(1-p)(1-\mu)+p\mu} < 0$$

and

$$\frac{\partial \Pr[a_i = h|\tilde{a}_i = l]}{\partial p} = \frac{p(1-2\mu)(1-\mu)}{(p(1-\mu)+(1-p)\mu)^2} - \frac{1-\mu}{p(1-\mu)+(1-p)\mu} < 0$$

Simplifying gives us

$$q\left[\left(\frac{p(1-2\mu)(1-\mu)}{(p(1-\mu)+(1-p)\mu)^2} - \frac{1-\mu}{p(1-\mu)+(1-p)\mu}\right)(h-l)\right]$$
$$+ (1-q)\left[\left(\frac{p\mu(2\mu-1)}{((1-p)(1-\mu)+p\mu))^2} - \frac{\mu}{(1-p)(1-\mu)+p\mu}\right)(l-h)\right].$$

By flipping the signs in the second line, it becomes easier for us to investigate the sign of the overall expression. Let us check when the expression is positive—i.e., when noise reduces the expected ability of promoted workers. We can write

$$q\left[\left(\frac{p(1-2\mu)(1-\mu)}{(p(1-\mu)+(1-p)\mu)^2} - \frac{1-\mu}{p(1-\mu)+(1-p)\mu}\right)(h-l)\right]$$
$$- (1-q)\left[\left(\frac{p\mu(2\mu-1)}{((1-p)(1-\mu)+p\mu))^2} - \frac{\mu}{(1-p)(1-\mu)+p\mu}\right)(h-l)\right]$$
$$\geq 0$$

Rewriting gives us

$$q\left[\left(\frac{p(1-2\mu)(1-\mu)}{(p(1-\mu)+(1-p)\mu)^2} - \frac{1-\mu}{p(1-\mu)+(1-p)\mu}\right)\right]$$
$$\geq (1-q)\left[\left(\frac{p\mu(2\mu-1)}{((1-p)(1-\mu)+p\mu))^2} - \frac{\mu}{(1-p)(1-\mu)+p\mu}\right)\right]$$

It is then useful to express this in ratios. We can write it as follows:

$$\frac{q}{1-q} \leq \frac{\left(\frac{p\mu(2\mu-1)}{((1-p)(1-\mu)+p\mu))^2} - \frac{\mu}{(1-p)(1-\mu)+p\mu}\right)}{\left(\frac{p(1-2\mu)(1-\mu)}{(p(1-\mu)+(1-p)\mu)^2} - \frac{1-\mu}{p(1-\mu)+(1-p)\mu}\right)}.$$

Notice the flip of the inequality sign in the last step because we are dividing by a negative number. Both sides of the above expression have a clear interpretation. The left hand side (LHS) is the likelihood ratio of the incumbent manager being high vs. low ability. The right hand side (RHS) is the relative change of mistake probability for different incumbent abilities when precision changes. Regarding the RHS of this expression we can say that $RHS = 1$ whenever $\mu = 1/2$. We have $RHS < 1$ if $\mu < 1/2$. Conversely, $RHS > 1$ if $\mu > 1/2$.

This observation is helpful for us to interpret the above condition. As a baseline, suppose that $q = 1/2$. Then $LHS = 1$. If additionally $\mu = 1/2$. Then $RHS = 1$ too. The above expression then holds with equality, implying that expected abilities are not affected by noise. However, once $\mu > 1/2$—even if only marginally—then the RHS is always < 1 and, hence, the above inequality never holds. If, instead, $\mu < 1/2$, then the RHS is always > 1 and, hence, the inequality always holds strictly.

Let us now look at the cases where $q \neq 1$. First, suppose that $q < 1/2$. Then the LHS is always < 1. If we again assume that $\mu = 1/2$ and hence the $RHS = 1$, then clearly the above inequality holds strictly. If $\mu < 1/2$ but only marginally, the inequality still holds. If μ is sufficiently smaller than 1/2 the inequality begins to not hold anymore. Hence, a lower q implies that we can get away with a lower μ for the above inequality to still be satisfied. The reverse is also true (a higher q implies that we can get away with a higher μ).

In this sense, for any $q \neq 1/2$, we can compensate by choosing a $\mu \neq 1/2$. This works even in the extreme cases. To see this, first look at the limits of the LHS. We have that

$$\lim_{q\to 0} LHS = 0 \quad \text{and} \quad \lim_{q\to 1} LHS = \infty \quad \Rightarrow \quad LHS \in (0,\infty)$$

We now need to investigate the limits of the RHS. We have

$$\lim_{\mu\to 0} RHS = 0 \quad \text{and} \quad \lim_{\mu\to 1} RHS = \infty \quad \Rightarrow \quad RHS \in (0,\infty)$$

From this we can see that both the LHS and the RHS have the same range. Because both sides are also continuous (the LHS in q and the RHS in μ)

and the RHS is monotonic in p,

$$\frac{\partial RHS}{\partial \mu} = -\frac{2(2p-1)((2\mu-1)p-\mu)}{(\mu-2\mu p+p-1)^3} < 0$$

this immediately implies that for any $q \in (0,1)$ there will exist values of μ for which the inequality above either (i) holds strictly, (ii) does not hold, or (iii) holds with equality. This third case implicitly pins down the threshold value μ^t. □

2.A.5 Proof of Proposition 2.3

Our approach to this proof is similar to the steps we have taken when deriving the condition under which the seniority scheme leads to higher expected ability, but assuming that the expected ability of the *promoted worker* is fixed.

From the arguments made in the main text, it follows that the expected ability of the promoted manager is maximized by choosing the seniority scheme if

$$\begin{aligned}
&\mathbb{E}\left[a_i \mid x_i > x_{-i}, b=0\right] = \\
&= [wh + (1-w)l] \cdot \left[q\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)} + (1-q)\frac{\mu p}{\mu p + (1-\mu)(1-p)} \right] \\
&\quad + q\left[1 - \frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)}\right]h + (1-q)\left[1-\frac{\mu p}{\mu p+(1-\mu)(1-p)}\right]l \\
&< \; qh + (1-q)l = \mathbb{E}\left[a_i \mid x_i > x_{-i}, b \to \infty\right],
\end{aligned}$$

where $w := \Pr[x_i > x_{-i} \mid a_i = h, a_{-i} = l, b = 0]$ is the probability that a high ability player wins against a low ability player. Note, w is not a function of μ, p or q. We can rearrange the LHS of this expression by factoring out h and l such that we have

$$\begin{aligned}
&\overbrace{\Omega}^{\text{“weight” on outcome } h} \cdot h \\
&+ \left[(1-w)\left(q\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)} + (1-q)\frac{\mu p}{\mu p+(1-\mu)(1-p)} \right) \right. \\
&\left. + (1-q)\left(1 - \frac{\mu p}{\mu p + (1-\mu)(1-p)}\right) \right] l < \underbrace{q}_{\text{“weight” on outcome } h} h + (1-q)l,
\end{aligned}$$

for the necessary and sufficient condition for the seniority scheme to lead to higher expected ability of the promoted manager. Hereby

$$\Omega := w\left(q\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)} + (1-q)\frac{\mu p}{\mu p+(1-\mu)(1-p)}\right) + q\left(1-\frac{(1-\mu)p}{(1-\mu)p+\mu(1-p)}\right)$$

Given that the factors of h and l are probabilities and add up to one, the initial inequality holds if and only if the factor of h on the RHS is greater than the factor of h on the LHS. (We are calculating a weighted average where the outcomes are fixed and the weights add up to one on both sides. Then it is sufficient to look at the weight on the higher outcome to determine which weighted average is greater.) We then have that $\mathbb{E}[a_i \mid x_i > x_{-i}] < \mathbb{E}[a_i \mid i = \text{incumbent}]$ if and only if

$$\Omega < q,$$

which we can rearrange to have q on the RHS such that

$$\frac{2pw\mu^2 - w\mu^2 - pw\mu}{p+w-pw+2\mu-3p\mu-2w\mu+2pw\mu-\mu^2+2p\mu^2-1} = \frac{\mu w((2\mu-1)p-\mu)}{w+(2\mu-1)p(\mu+w-1)-2\mu w-(\mu-1)^2} \equiv \hat{q}(w,p,\mu) < q.$$

One can check that $\hat{q} \in (0,1)$ for any parameter combination of p, w and μ. It follows then that there always exists a $q \in (0,1)$ for which the seniority rule will lead to a higher expected ability for the promoted manager than the competitive promotion scheme. Furthermore,

$$\frac{\partial \hat{q}}{\partial w} > 0, \qquad \frac{\partial \hat{q}}{\partial \mu} > 0 \qquad \frac{\partial \hat{q}}{\partial p} \in (-\infty, \infty),$$

implying that the sufficient condition is easier to satisfy—a lower q is sufficient to satisfy—, when the share of high ability workers are low and when ability difference is small (which implies that the win probability of the high ability player is lower when competing against a low ability player). □

2.A.6 Proof of Lemma 2.4

We build on the proof of Proposition 2.3. We can proceed with finding the threshold noise level $p^t \in \left(\frac{1}{2}, 1\right)$, whereby the promotion scheme that leads to a greater expected ability of the promoted manager is determined by which side of the threshold parameter p is, holding all else constant.

When the threshold noise level exists, it is implicitly defined by setting the intial inequality to a strict equality, such that the expected ability is the same for both promotion schemes. After rearrangement, the explicit expression (assuming $\mu \neq \frac{1}{2}$) is given by:

$$\begin{aligned} p^t &= \frac{-q\mu^2 + 2\mu q - 2\mu q w + qw - q + \mu^2 w}{-2\mu^2 q + 3\mu q - 2\mu q w + qw - q + 2\mu^2 w - \mu w} \\ &= \frac{q\left((\mu-1)^2 + (2\mu-1)w\right) - \mu^2 w}{(2\mu-1)(q(\mu+w-1) - \mu w)}, \end{aligned}$$

if either of the following conditions hold:

$$\begin{array}{lllll} 0 < q \leq \frac{1}{2} & \text{and} & \frac{1}{2} < w < 1 & \text{and} & 0 < \mu < \frac{qw-q}{2qw-q-w} \\ \frac{1}{2} < q < 1 & \text{and} & \frac{1}{2} < w < q & \text{and} & \frac{qw-q}{2qw-q-w} < \mu < 1 \\ \frac{1}{2} < q < 1 & \text{and} & q < w < 1 & \text{and} & 0 < \mu < \frac{qw-q}{2qw-q-w}. \end{array}$$

These are the conditions given in Lemma 2.4. □

Proof of Proposition 2.4.

We proceed from the proof of Lemma 2.4 above by checking how the threshold $\hat{q}$ changes when we change noise:

$$\begin{aligned} \frac{\partial \hat{q}}{\partial p} =& \frac{2w\mu^2 - w\mu}{p + w - pw + 2\mu - 3p\mu - 2w\mu + 2pw\mu - \mu^2 + 2p\mu^2 - 1} \\ &- \frac{(1 - w - 3\mu + 2w\mu + 2\mu^2)\,(2pw\mu^2 - pw\mu - w\mu^2)}{(p + w - pw + 2\mu - 3p\mu - 2w\mu + 2pw\mu - \mu^2 + 2p\mu^2 - 1)^2}. \end{aligned}$$

From here we can check that

$$\left.\frac{\partial \hat{q}}{\partial p}\right|_{\mu<1/2} > 0, \quad \left.\frac{\partial \hat{q}}{\partial p}\right|_{\mu>1/2} < 0, \quad \text{and} \quad \left.\frac{\partial \hat{q}}{\partial p}\right|_{\mu=1/2} = 0, \qquad (2.15)$$

which corresponds to the three cases described in the proposition.

Now assume that p^t exists and fix w, p and $\mu < 1/2$, and take q $s.t.$ $\hat{q} < q$. Then from $\left.\frac{\partial \hat{q}}{\partial p}\right|_{\mu<1/2} > 0$ we have that if we increase noise sufficiently, then $\hat{q}$

will increase above q, implying that expected ability is maximized under a seniority scheme. For this we need initial $\hat{q}$ to be in the neighbourhood of q such that the largest admissible $\bar{p} < 1$ will indeed increase $\hat{q}$ above q: we need q to be sufficiently low.

We replicate this argument for the case when $1/2 < \mu$. Take again a q $s.t.$ $\hat{q} < q$. Then from $\frac{\partial \hat{q}}{\partial p}\big|_{\mu > 1/2} < 0$ we have that decreasing noise increases threshold $\hat{q}$. We just need to make sure q is again low enough, such that the lowest admissible bias $\underline{p} \in (\frac{1}{2}, 1)$ can still increase $\hat{q}$ above q.

Last, when $\mu = 1/2$, changing signal precision has no effect on $\hat{q}$ given that $\frac{\partial \hat{q}}{\partial p}\big|_{\mu = 1/2} = 0$. This concludes the proof. □

Proof of Proposition 2.5

The first part of the statements follow from Proposition 2.4. The result that seniority scheme minimizes expected effort simply follows from the observation that effort is costly and prize v is deterministically allocated prior to effort choice. □

2.B References

Bai, Y., and T. Zhou. 2019. “”Mao’s last revolution”: a dictator’s loyalty-competence tradeoff.” *Public Choice* 180 (3): 469–500.

Baye, M. R., D. Kovenock, and C. G. De Vries. 1996. “The all-pay auction with complete information.” *Economic Theory* 8 (2): 291–305.

Besley, T., O. Folke, T. Persson, and J. Rickne. 2017. “Gender quotas and the crisis of the mediocre man: Theory and evidence from Sweden.” *American Economic Review* 107 (8): 2204–2242.

Carmichael, H. L. 1988. “Incentives in academics: why is there tenure?” *Journal of Political Economy* 96 (3): 453–472.

Cirone, A., G. W. Cox, and J. H. Fiva. 2021. “Seniority-based nominations and political careers.” *American Political Science Review* 115 (1): 234–251.

Denter, P., J. Morgan, and D. Sisak. 2022. “Showing off or laying low? The economics of psych-outs.” *American Economic Journal: Microeconomics* 14 (1): 529–580.

Drugov, M., and D. Ryvkin. 2017. “Biased contests for symmetric players.” *Games and Economic Behavior* 103:116–144.

Egorov, G., and K. Sonin. 2011. “Dictators and Their Viziers: Endogenizing the Loyalty-Competence Trade-off.” *Journal of the European Economic Association* 9 (5): 903–930.

Ekinci, E. 2016. “Employee referrals as a screening device.” *The RAND Journal of Economics* 47 (3): 688–708.

Fang, D., and T. H. Noe. 2022. “Less competition, more meritocracy?” *Journal of Labor Economics.*

Finan, F., B. A. Olken, and R. Pande. 2015. *The Personnel Economics of the State.* NBER Working Paper.

Friebel, G., and M. Raith. 2004. "Abuse of authority and hierarchical communication." *The RAND Journal of Economics*, 224–244.

———. 2022. "Talent Management: The Role of Bosses." *Available at SSRN 4162631.*

Giardili, S., K. Ramdas, and J. W. Williams. 2022. "Leadership and productivity: a study of US automobile assembly plants." *Management Science.*

Glazer, A. 2002. "Allies as rivals: internal and external rent seeking." *Journal of Economic Behavior & Organization* 48 (2): 155–162.

Gneezy, U., S. Meier, and P. Rey-Biel. 2011. "When and Why Incentives (Don't) Work to Modify Behavior." *Journal of Economic Perspectives* 25 (4): 191–210.

Gneezy, U., and A. Rustichini. 2000. "Pay Enough or Don't Pay at All." *The Quarterly Journal of Economics* 115 (3): 791–810.

Haegele, I. 2022. "Talent Hoarding in Organizations." *arXiv preprint arXiv:2206.15098.*

Hillman, A. L., and J. G. Riley. 1989. "Politically contestable rents and transfers." *Economics & Politics* 1 (1): 17–39.

Hu, C.-Y., S. Wang, and J.-C. Huang. 2019. "Too Good to be Mentored? Testing the Rising Star Hypothesis in Formal Mentoring." In *Academy of Management Proceedings*, 2019:16491. 1. Academy of Management Briarcliff Manor, NY 10510.

Hvide, H. K., and E. G. Kristiansen. 2003. "Risk taking in selection contests." *Games and Economic Behavior* 42 (1): 172–179.

i4cp. 2016. "Talent Mobility Matters." *Industry Report.*

Jia, R., M. Kudamatsu, and D. Seim. 2015. "Political Selection in China: The Complementary Roles of Connections and Performance." *Journal of the European Economic Association* 13 (4): 631–668.

Kawamura, K., and I. M. de Barreda. 2014. "Biasing selection contests with ex-ante identical agents." *Economics Letters* 123 (2): 240–243.

Kofman, F., and J. Lawarrée. 1993. "Collusion in hierarchical agency." *Econometrica*, 629–656.

Kurter, H. L. 2021. *4 Useful Tips For When Your Manager Is Sabotaging Your Career Growth.* https://www.forbes.com/sites/heidilynnekurter/2021/08/19/4-useful-tips-for-when-your-boss-is-sabotaging-your-career-growth/?sh=772285413c86.

Leete, L. 2000. "Wage equity and employee motivation in nonprofit and for-profit organizations." *Journal of Economic Behavior Organization* 43 (4): 423–446.

Meyer, M. A. 1991. "Learning from coarse information: Biased contests and career profiles." *The Review of Economic Studies* 58 (1): 15–41.

Mollick, E. 2012. "People and process, suits and innovators: the role of individuals in firm performance." *Strategic Management Journal* 33 (9): 1001–1015.

Popomaronis, T. 2019. *Here's how many Google interviews it takes to hire a Googler.* https://www.cnbc.com/2019/04/17/heres-how-many-google-job-interviews-it-takes-to-hire-a-googler.html.

Sengupta, S. 2004. "Delegating recruitment under asymmetric information." *International Journal of Industrial Organization* 22 (8-9): 1327–1347.

Siegel, R. 2009. "All-Pay Contests." *Econometrica* 77 (1): 71–92.

Smith, J. 2015. *Here's a good sign your boss feels threatened by you.* https://www.businessinsider.com/sign-your-boss-feels-threatened-by-you-2015-2?r=US&IR=T.

Stracke, R., W. Höchtl, R. Kerschbamer, and U. Sunde. 2015. "Incentives and selection in promotion contests: Is it possible to kill two birds with one stone?" *Managerial and Decision Economics* 36 (5): 275–285.

Strausz, R. 1997. "Delegation of Monitoring in a Principal-Agent Relationship." *The Review of Economic Studies* 64 (3): 337–357.

Tirole, J. 1986. "Hierarchies and bureaucracies: On the role of collusion in organizations." *Journal of Law, Economics & Organization* 2:181.

Tsoulouhas, T., C. R. Knoeber, and A. Agrawal. 2007. "Contests to become CEO: incentives, selection and handicaps." *Economic Theory* 30 (2): 195–221.

Wagner, A. F. 2006. "A model of loyalty and competence." *mimeo*, 59.

Zakharov, A. V. 2016. "The Loyalty-Competence Trade-Off in Dictatorships and Outside Options for Subordinates." *The Journal of Politics* 78 (2): 457–466.

Chapter 3

Promotions Under Excessive Workplace Surveillance

Timm Behler

Abstract

When designing promotion contests, managers might not only be interested in providing effort incentives, but also in selecting the employee with the highest ability. In this paper, I show that noisy contests can perform better regarding this objective than noiseless contests. While noise generally distorts the correlation between output and ability, noiseless contests induce mixed strategies which can distort this correlation even more. I show that there always exists a uniform noise level that improves both (expected) aggregate effort and selection efficiency over the noiseless contest. Furthermore, this noise level improves employees' welfare. My main interpretation is that excessive workplace-surveillance can backfire (i) by hampering meritocratic selection, and (ii) by reducing productivity.

3.1 Introduction

Consider the manager of a firm who is designing a promotion contest. In the contest, the firm's employees compete against each other over a promotion by submitting costly effort. The goal of such a contest is twofold: First, it can be used to provide effort incentives for the employees, thereby increasing the firm's productivity. Second, if the employees are heterogeneous in abilities, it can also provide information about the employees' abilities; such information can, of course, be relevant for the promotion, for example if an employee of high ability would be best suited to successfully complete the work at a higher level of the organizational hierarchy.

One important question regarding the design of such promotion contests is how the manager can monitor employees' performances: Getting precise information about how hard individuals work is not easy. In recent years, however, new technologies have emerged that can help the manager. New software, for example, enables her to track how much time employees spend working on which tasks, GPS systems can help the manager to track employees' movements and whereabouts throughout the workday, and facial recognition can, for example, precisely determine employees' arrival time. More generally, these new technologies enable the manager to identify better who is shirking and who is working hard—in other words, they help to generate a less noisy signal of each employee's effort submission.

These new technologies have sparked much interest, particularly as more and more people work from home and managers' desire to monitor employees remotely grows larger.[1] At the same time, many people are concerned about the increasing workplace surveillance and its potentially harmful effects. For example, constant monitoring might come at a negative psychological cost for the employees and, hence, might impact their overall well-being. A further concern is the invasion of employees' privacy.[2] While such concerns are often publicly debated, many of these technologies—for example, facial recognition—remain largely legal in the USA.[3] The European Parliament, instead, has succesfully voted on a ban of such recognition software in 2023.[4]

In this paper, I argue that even from an organization's perspective, less noise need not be beneficial. Particularly, I look at how the organization's manager can strategically employ noise to maximize both selection efficiency—i.e., the probability of selecting an employee of high ability—and (expected) aggregate effort. I consider a simple contest with two employees who differ

[1]See, for example, this New York Times article.
[2]See, for example, this New York Times article.
[3]See, for example, this Bloomberg article.
[4]See, for example, this official statement.

in their ability. In the contest, the principal only receives a noisy signal of each agent's individual effort. The agent with the highest signal is selected as the winner of the contest. This setup is roughly in line with the rank-order tournament originally considered by Lazear and Rosen (1981). My setup differs in the assumption that I make about the noise distribution. Particularly, I make no explicit assumption on individual noise terms and, instead, assume that the *difference* of the individual noise terms is uniformly distributed. This also implies that, unlike in Lazear and Rosen (1981), individual noise is not necessarily iid. As I argue throughout this paper, this type of noise is plausible in the settings that I consider because it implies that the manager can observe the organization's total output—her uncertainty only corresponds to individuals' *contribution* to total output.

I allow for the principal to decide on how strict she monitors the agents. A stricter monitoring leads to a less noisy signal of each agent's effort. As monitoring becomes perfect, the setup converges to the (noiseless) all-pay contest (Siegel 2009).[5]

It may appear intuitive that noise distorts the correlation between ability and output and, hence, that selection efficiency should be higher in the noiseless setting. The first part of this intuition is correct: Noise generally distorts the correlation between ability and output. The second part, however, is less obvious. The intuition misses that, in contrast to the noisy setting, the noiseless contest has no equilibrium in pure strategies. Instead, agents will choose a continuous distribution over some effort interval—a well-known result in the contest literature. Effectively, this makes output random, too, and can be interpreted as *endogenous* noise that is added by the contestants. As a result, the answer to the question which monitoring scheme selects the high-ability agent with greater probability becomes non-apparent and warrants an investigation. As we shall see, the question is resolved unambiguously: The principal can generally improve upon the selection efficiency of the noiseless framework by introducing some noise. Furthermore, she can also improve upon the selection efficiency of contests with *small* noise such that equilibrium is fully mixed by choosing a contest with high enough noise such that a pure strategy equilibrium exists. However, there may be strictly positive levels of noise that allow for semi-mixed equilibria whose analysis I omit in this paper. Nevertheless, while I cannot provide a solution to determine the "optimal" noise level, my results serve as an important demonstration that introducing some noise is, in general, good.

After introducing the model in section 3.2, and deriving the probabilities

[5]Throughout this paper, I will use the expressions "noiseless contest" and "noisy contest" interchangeably with "all-pay contest" and "rank-order tournament", respectively.

of selecting the high-ability agent under both promotion paradigms in section 3.3, I move on to a comparison in section 3.4. I derive a sufficient condition for selection efficiency to be higher in the noisy contest. The permissible noise level is restricted from above and from below. The upper bound is determined by the value that is attached to the promotion. The intuition is that, in the noisy contest, selection efficiency decreases in the noise level, but increases in the value attached to the promotion because the value motivates agents of differing ability asymmetrically. A higher value, hence, permits a higher noise level, too. Perhaps interestingly, the permissible noise level does not depend on how asymmetric the agents are. This is because the effect of the high-ability agent's relative advantage on selection efficiency exactly offsets itself when comparing the different contests.

The lower bound of the permissible noise level is determined by the noise level that is sufficiently high to guarantee the existence of a pure-strategy equilibrium. The noise level that guarantees the existence of a pure-strategy equilibrium depends positively on both the value attached to the promotion and on how asymmetric the agents are. This is because this noise level is given by the noise level that is high enough to guarantee the low-ability agent a positive expected rent from playing the pure strategy. Since a higher value motivates the agents asymmetrically, it lowers the low-ability agent's winning probability and, hence, her expected rent. Similarly, an increase in the high-ability agent's advantage demotivates the low-ability agent and, as a result, lowers her chance of winning and her expected rent. The decrease of the low-ability agent's expected rent due to an increase in either the value attached to the promotion or the agent's heterogeneity must be compensated by a higher noise level to break even and ensure the existence of a pure-strategy equilibrium.

As my main result, I show that, for any positive value that is attached to the promotion and any level of heterogeneity, there exists a noise level such that a pure-strategy equilibrium exists and which is sufficiently low such that selection efficiency in the noisy contest is higher than in the noiseless contest. Finally, I show that, among the noise levels I consider, the optimal noise level regarding a maximization of selection efficiency coincides with the optimal noise level that maximizes (expected) aggregate effort of the agents. This compatibility result is in contrast to some of the earlier literature (e.g., Stracke et al. 2015).

After deriving the main results, in section 3.5, I consider employee welfare and show that it is, perhaps surprisingly—after all employees work on expectation more, while the prize is kept constant—, maximized by the noisy contest as well. In section 3.6, I first discuss formally what kind of noise the principal can employ strategically. The type of noise the principal can employ strategically is indeed the type of noise we would expect to be induced through less stringent monitoring. I highlight this through an example. Finally, section 3.7 concludes. Most of the proofs follow immediately from the analyses presented in the main text right before the corresponding results. In these cases, I omit the proof.

Related Literature. This paper is related to several strands of literature. Conceptually, it is related to the literature on selection efficiency in contests. This literature, like much of the contest literature in general, differentiates between either contests with- or without exogenous noise. Nevertheless, some well known results hold in both types of contests. In one of the literature's seminal papers, Meyer (1991) considers a multi-stage contest with exogenous noise. She shows that, when the goal is to identify the most talented agent, the principal should bias the later stages of the contest in favor of the "leader" of the previous stages. In a static model with endogenous noise, Kawamura and Barreda (2014) derive a similar result. They show that the principal should bias the contest in favor of the agent who has the higher prior probability of having a high ability. If this prior probability is the same for the agents, the principal should randomly assign headstarts to either agent. A paper by Drugov and Ryvkin (2017) yields a similar result. This last result has a particular connection to my paper. Since the optimal bias must be random, it can also be interpreted as some form of noise which, like in my paper, increases selection efficiency.

My paper differs from much of the previous literature in that I do not focus on one of the two types of contests—i.e., contest with exogenous vs. endogenous noise. Instead, I assume that it is the manager who can endogenously choose an exogenous noise level. As pointed out by Morgan, Tumlinson, and Vardy (2018), noise is seldom viewed as something that can be employed strategically by the principal. This is perhaps surprising since, in reality, the designer arguably has some influence on noise, most obviously through the decision on how rigorously to monitor the agents. Nevertheless, some papers assume that the designer can choose noise. Morgan, Tumlinson, and Vardy (2018) and Letina, Liu, and Netzer (2020) both focus on the optimality of noise when the objective of the principal is to maximize (expected) aggregate effort. Both find that effort is maximized at the lowest noise level that guarantees the existence a of pure-strategy equilibrium. Besides these papers, the idea that a principal should deter agents from adding endogenous noise to their output is also found outside of the contest literature.Barron, Georgiadis, and Swinkels (2020), for example, show that when agents are strategic risk-takers, the principal optimally chooses a contract that concavifies an agent's utility function in order to deter her from risk-taking. Generally, the logic behind these results is similar to my paper. Endogenous noise leads to different inefficiencies and, hence, exogenous noise should be added by the designer to deter the agents from adding noise endogenously.

The interpretation of this result is very closely related to the results by Meyer (1991) and Kawamura and Barreda (2014), but also to those by Hvide and Kristiansen (2003) andFang and Noe (2022), who both focus on a specific form of endogenous noise—strategic risk-taking—and how it affects selection efficiency. Both find, somewhat counterintuitively, that when contestants are strategic risk-takers, more competition can make selection less meritocratic—i.e., decrease the probability of selecting a high-ability type. More generally, all of these papers show how seemingly meritocratic design choices can actually hamper meritocratic selection. This is a central observation in my model, too: At first sight, making a contest more noisy is clearly unmeritocratic because it distorts the correlation between ability and output. At the same time, the outcome turns out to be more meritocratic since a reduction in noise would induce the agents to strategically distort the ability-output correlation even more than a sufficiently low level of exogenous noise does.

Importantly, this also implies that an excessive monitoring of employees can actually work against the firm's goals. Other papers have previously derived similar results, particularly regarding the effect of excessive monitoring on effort. Zhao (2008) provides conditions under which it is optimal for the firm to rely only on noisy output measures when compensating employees, even when it can perfectly monitor employee's effort submissions. Similarly, Bar-Isaac and Deb (2021) show that infrequent monitoring of an agent can increase effort compared to the constant-monitoring benchmark. In their model, this is driven by the agent's reputational concerns. Roughly speaking, if the agent anticipates not being monitored in the next period, then exerting more effort in this period is worth more, since she can coast on her high-effort reputation in the next period. While these papers hint at the adverse effects of excessive monitoring (although not in a contest setting), there is, to the best of my knowledge, none that focuses on the effects on meritocratic selection.

3.2 Model

Consider two agents, denoted by i, who compete for a promotion. Agents can be one of two types, *high ability* or *low ability*. I denote the types by $\theta_i \in \Theta \equiv \{h, l\}$, respectively. In a slight abuse of notation, I will sometimes index an h-type agent by h and an l-type agent by l. The agents' types are drawn randomly but are perfectly correlated in the sense that there will always be one h- and one l-type. From the perspective of the principal, each agent has the same ex-ante probability to be either type. Both agents can observe their realized type and, hence, know the type of the other agent, too. The principal, however, can observe neither agent's type.

Each agent produces an observable signal of output

$$q_i = x_i + \epsilon_i, \tag{3.1}$$

where ϵ_i is a random noise term, which I assume to be such that the difference of the noise terms $\epsilon_i - \epsilon_{j\neq i}$ is uniformly distributed,[6] specifically

$$\Delta\epsilon \overset{\mathrm{d}}{\sim} \mathcal{U}\left[-\frac{1}{2}\xi, +\frac{1}{2}\xi\right];$$

x_i is non-negative effort that i can exert at cost $c_i(x_i)$. Specifically, I assume that cost is given as follows:

$$c_i(x_i) = \begin{cases} x_i^2 & \text{if } \theta_i = h \\ \alpha x_i^2 & \text{if } \theta_i = l, \end{cases} \tag{3.2}$$

This implies that for any x, we have that $c_l(x) > c_h(x)$.

The principal promotes the agent with the highest output q_i. I denote the value attached to the promotion by v. The agent who is not promoted gets, without loss of generalization, a prize of value zero. Hence, agents' expected payoff is

$$\Pi_i = \Pr\left[q_{j\neq i} < q_i\right] v - c_i(x_i), \tag{3.3}$$

where $\Pr\left[q_{j\neq} < q_i\right]$ is i's probability of being promoted. The agents' winning probabilities will be specified further in the next sections. Generally, in

[6] This specification was previously used in similar settings by Konrad (2009), Ederer (2010), and Brown and Minor (2014). This distributional assumption is plausible, for example, when the principal can exert aggregate output but not each agent's individual contribution. Instead the principal randomly overestimates the contribution by some amount while underestimating the contribution of the other agent by the same amount. In my setting, this is a particularly plausible assumption, which I will argue in more detail in Section 5.

the all-pay contest, the winning probability depends only on the agents' (random) effort choices, whereas in the rank-order tournament, it depends on the distribution of the random noise term.

The principal's objective is to maximize selection efficiency:

Definition 3.1 (Selection Efficiency)**.** *Selection efficiency is the probability that the high-ability agent wins the contest,* $\Pr[q_h > q_l]$

That is, selection efficiency is defined by the probability of promoting the high-ability agent which is equivalent to the probability that the high-ability agent produces the higher output. The principal has discretion over choosing the noise level of the random shocks ϵ_i. I will later specify in more detail what this means formally. As suggested by my motivating example, in practice, a principal could vary noise by monitoring agents more or less rigorously—although, of course, other interpretations are possible.

Intuition and Discussion. Without noise, the framework is equivalent to the standard all-pay contest and, as discussed by Siegel (2009), there exists no pure-strategy equilibrium:[7] Assume otherwise that the agents choose pure strategies. Then, the high-ability agent could always profitably outperform the low-ability agent since her effort-cost is lower. Given that the high-ability agent would always outperform the low-ability agent and, hence, always win, however, the low-ability agent would have no incentive to exert effort at all. But if the low-ability agent exerts no effort, the high-ability agent's effort incentives would break down, too. Hence, in equilibrium we must allow for some probability that the low-ability type will be selected. This is the case when agents play mixed strategies.

To the contrary, when noise is sufficiently high, there generally exists a pure-strategy equilibrium. This is the case because noise already guarantees the low-ability agent some positive winning probability. The conditions under which selection efficiency is higher in either of the contests will be derived in the next sections.

There is also the case of intermediate noise—i.e., some positive noise that is too small to ensure (fully) pure strategy equilibria. These intermediate cases are very difficult to analyze, however, and even though some breakthroughs have been made over the years (see, for example, Baye, Kovenock,

[7]Notice, however, that the general logic goes back further than Siegel (2009). For example to Hillman and Riley (1989), Baye, Kovenock, and De Vries (1993) and Baye, Kovenock, and De Vries (1996), who provide two applied and one very comprehensive analysis of the the all-pay auction, respectively.

and De Vries (1994), Ewerhart (2015), and Ewerhart (2017)), there is no comprehensive analysis of these cases available yet. Because the aim of this paper is to provide an important application of two canonical contest models—which can be demonstrated even through a comparison of the extreme cases—I do not consider the intermediate cases in this paper. It is important to stress, however, that Ewerhart (2017) has shown that strictly positive noise levels that are, in a specific sense, "small" are equivalent to the noiseless setting. In this sense, my results demonstrate clearly that noise is, quite generally, good as long as it is not too high. The only cases that I omit are the cases where noise is too high for fully mixed equilibria to be unique but too low for a pure strategy equilibrium to exist.

3.3 Equilibrium Analysis

In this section, I characterize the equilibrium and the resulting probabilities of selecting the high-ability agent in both cases—i.e., with- and without exogenous noise. I start with the (noiseless) all-pay contest in section 3.1, before moving on to the (noisy) rank-order tournament in section 3.2.

3.3.1 The All-Pay Contest

I start by analyzing the case without noise—i.e., I am assuming that the distribution of ϵ_i is degenerate at $\epsilon_i = 0$. The contest is then equivalent to the all-pay contest with two heterogeneous agents. I am looking for an equilibrium in which both types of agents randomize their effort submissions according to the cumulative distribution functions (CDF, henceforth) F_l and F_h, respectively. The expected payoff of the l-type from playing her maximum bid must be zero. Otherwise, she could always marginally increase her maximum bid and thereby strictly increase both her winning probability and, hence, her expected rent. This directly implies that her maximum bid must be $c_l^{-1}(v)$. The expected payoff from any other effort submission $x_l \in [0, c_l^{-1}(v)]$ must, hence, be zero, too. This reasoning yields the h-type's mixed equilibrium strategy.

Similarly, the l-type chooses F_l such that the h-type's expected profit from submitting any effort $x_h \in [0, c_l^{-1}(v)]$ is equal to the payoff she receives from submitting $x_l = c_l^{-1}(v)$ and winning with probability one (her best deviation strategy). Again, this reasong yields the l-types mixed equilibrium strategy.

The equilibrium can more completely be characterized by the following Lemma which follows from Hillman and Riley (1989).

Lemma 3.1 (All-Pay Contest (Hillman and Riley 1989)). *Suppose that $\alpha > 1$. We can summarize the unique equilibrium of the all-pay contest as follows:*

1. *The agents' equilibrium strategies are given by*

$$F_h^*(x_h) \; = \frac{1}{v} c_l(x_h) \quad \text{for}\, x_h \in [0, c_l^{-1}(v)] \tag{3.4}$$

and

$$F_l^*(x_l) := 1 + \frac{1}{v} c_h(x_l) - \frac{1}{v} c_h(c_l^{-1}(v)) \quad \text{for}\, x_l \in [0, c_l^{-1}(v)], \tag{3.5}$$

respectively

2. *The probability that the h-type wins is given by*

$$\Pr^{\text{AP}}\left[q_l < q_h\right] = 1 - \frac{1}{2\alpha}. \tag{3.6}$$

3. *The agents' expected effort submissions are given by*

$$\mathbb{E}^{\text{AP}}\left[x_h\right] = \frac{2}{3}\sqrt{\frac{v}{\alpha}} \tag{3.7}$$

and

$$\mathbb{E}^{\text{AP}}\left[x_l\right] = \frac{1}{\alpha}\mathbb{E}^{\text{AP}}\left[x_h\right], \tag{3.8}$$

respectively.

Proof: Follows from Hillman and Riley (1989). □

3.3.2 The Rank-Order Tournament

I will proceed with the analysis of the contest with noise high enough to ensure the existence of a pure-strategy equilibrium—i.e., the rank-order tournament à la Lazear and Rosen (1981)—again, however, with a modification to the distributional assumptions regarding the error term. In this case, we can write the h-type's winning probability as

$$\begin{aligned} \Pr^{\text{RO}}\left[q_l < q_h\right] &= \Pr\left[x_h - x_l > \epsilon_l - \epsilon_h\right] \\ &= \Pr\left[x_h - x_l > \Delta\epsilon\right] \\ &= \Phi(x_h - x_l) \end{aligned}$$

where $\Phi(\cdot)$ is the CDF of the random variable $\Delta\epsilon \equiv \epsilon_l - \epsilon_h$. Its PDF will be denoted by $\varphi(\cdot)$. Using this, the expected payoffs of the h-type and l-type become

$$\Pi_h(x_h) = \Phi(x_h - x_l)v - c_h(x_h), \quad \text{and} \quad \Pi_l(x_l) = [1 - \Phi(x_h - x_l)]\, v - c_l(x_l), \tag{3.9}$$

respectively. The corresponding FOCs are then

$$c_h'(x_h^*) = \varphi(x_h^* - x_l^*)v, \quad \text{and} \quad c_l'(x_l^*) = \varphi(x_h^* - x_l^*)v, \tag{3.10}$$

and implicitly specify both agents' equilibrium effort submissions, x_h^* and x_l^*. Notice that from condition (3.10), we can immediately see that

$$c_h'(x_h^*) = c_l'(x_l^*),$$

which, using the cost function specified in (3.2), implies

$$x_h^* = \alpha x_l^*.$$

Notice that the equilibrium effort submissions of the h-type are always (strictly) higher than those of the l-type. This is not surprising given that, while the marginal winning probabilities of the two types are the same, the l-type faces higher marginal cost.

Given distributional properties of $\Delta\epsilon$, we know that $\varphi(\cdot) = 1/\xi$ and, hence, we can rewrite condition (3.10) as

$$2x_h^* = \frac{1}{\xi}v, \quad \text{and} \quad 2\alpha x_l^* = \frac{1}{\xi}v$$

to finally find the equilibrium effort submissions of the h- and the l-type,

$$x_h^* = \frac{v}{2\xi}, \quad \text{and} \quad x_l^* = \frac{v}{2\xi\alpha}, \tag{3.11}$$

respectively. Using these terms, we can rewrite the winning probability of the h-type as

$$\Pr^{\mathrm{RO}}[q_l < q_h] := \Phi(x_h^* - x_l^*) = \frac{1}{\xi}\left[\frac{v}{2\xi} - \frac{v}{2\xi\alpha} + \frac{1}{2}\xi\right] = \frac{v}{2\xi^2}\left(1 - \frac{1}{\alpha}\right) + \frac{1}{2}. \tag{3.12}$$

Notice that, unlike the h-type's winning probability in the noiseless contest, the winning probability now depends not only on the h-type's cost benefit, α, but also on the value attached to the promotion, and on how noisy the contest is, described by ξ.

Lemma 3.2 (The Rank-Order Tournament)**.** *Suppose that* $\alpha > 1$. *We can summarize the equilibrium of the rank-order tournament as follows:*

1. *Both agents' equilibrium strategies are given in* (3.11).
2. *The probability that the h-type wins is given by* (3.12) *and depends positively on her cost advantage and the value of the promotion, and negatively on the noise level.*

Proof: Part (1) follows from the arguments made in the text. Part (2) follows from the arguments in the text together with an inspection of (3.12): $\partial \mathrm{Pr}^{\mathrm{RO}} [q_l < q_h] / \partial\alpha > 0$; $\partial \mathrm{Pr}^{\mathrm{RO}} [q_l < q_h] / \partial v > 0$, and $\partial \mathrm{Pr}^{\mathrm{RO}} [q_l < q_h] / \partial\xi < 0$. □

3.4 Should the Principal Choose a Noisy Contest?

In this section, I compare the two contests to see whether and when the principal should choose a noisy over a noiseless setting. I first show that the set of noise levels that ensure a higher selection efficiency than the no-noise benchmark is always non-empty. I further show that, if the principal can choose the noise level freely (ignoring conceptually difficult cases with intermediate noise levels), then she should choose a contest with the lowest noise level that still guarantees the existence of a pure-strategy equilibrium to maximize selection efficiency. Afterwards, I show that the noise level that maximizes selection efficiency in this sense coincides with the noise level that maximizes (expected) aggregate effort.

3.4.1 Selection Efficiency

We need to compare the probability of selecting the h-type in the noisy and the noiseless contest. The probability of selecting the h-type is higher in the noisy contest if

$$\mathrm{Pr}^{\mathrm{RO}} [q_l < q_h] = \frac{v}{2\xi^2}\left(1 - \frac{1}{\alpha}\right) + \frac{1}{2} \geq 1 - \frac{1}{2\alpha} = \mathrm{Pr}^{\mathrm{AP}} [q_l < q_h] .$$

Rearranging this gives us the critical value of ξ^2, $\overline{\xi}^2$, such that the rank-order tournament *just* dominates the all-pay contest in terms of selection efficiency:

$$\xi^2 \leq \overline{\xi}^2 := v. \tag{3.13}$$

Condition (3.13) puts an upper bound on the permissible noise level such that selection efficiency is higher in the rank-order tournament. It is intuitive that this upper bound depends positively on the value attached to the promotion: The value motivates the agents asymmetrically and, hence, a higher value increases selection efficiency. This also implies that a higher-valued promotion admits a higher noise level while still obtaining a higher selection efficiency than the all-pay contest. We must note, however, that it is unclear whether such a critical value of ξ^2 actually exists: If ξ^2 is too low, a pure-strategy equilibrium fails to exist. This implies that if the noise level that is still high enough to ensure the existence of a pure-strategy equilibrium lies above $\overline{\xi}^2$, then it is impossible to find a noise level that yields a higher selection efficiency than the noiseless contest.

To check this, I proceed by deriving the critical level of ξ^2 such that a pure-strategy equilibrium exists. As pointed out by Baye, Kovenock, and De Vries (1994), the existence problem regarding a pure-strategy equilibrium for insufficient noise levels lies in the observation that expected rent may be negative and, hence, zero effort would be a profitable deviation from the proposed equilibrium strategies. To eliminate this possibility, I derive the noise level sufficiently high such that both types of agents earn positive expected rent given the proposed equilibrium effort submissions. From expression (3.9) follows that the expected rent of the l-type will always be lower than that of the h-type and, hence, it imposes the stronger restriction; it suffices to investigate the expected rent of the l-type: In the proposed equilibrium, she earns

$$\Pi_l^* \equiv \left[1 - \frac{1}{\xi}\left(\frac{v}{2\xi} - \frac{v}{2\xi\alpha} + \frac{1}{2}\xi\right)\right] v - \alpha\left(\frac{v}{2\xi\alpha}\right)^2,$$

which can be simplified to

$$\Pi_l^* = \frac{1}{2v} - \frac{1}{\xi}\left(\frac{1}{2} - \frac{1}{4\alpha}\right). \tag{3.14}$$

This expression needs to be larger than zero for a pure-strategy equilibrium to exist. This yields our critical value, a lower bound, of ξ^2, $\underline{\xi}^2$, such that

$$\xi^2 \geq \underline{\xi}^2 := \frac{v\,(2\alpha - 1)}{2\alpha}. \tag{3.15}$$

Lemma 3.3 (Existence of Pure Strategy Equilibrium)**.** *A pure-strategy equilibrium exists if and only if condition* (3.15) *holds.*

Together, conditions (3.13) and (3.15) imply that the principal can set $\xi^2 \in \left[\underline{\xi}^2, \overline{\xi}^2\right]$ to end up with a promotion mechanism that outperforms the all-pay contest in terms of selection efficiency. This is only feasible if the interval is indeed non-empty. This is the case if

$$v \geq \frac{v\,(2\alpha - 1)}{2\alpha} \Leftrightarrow 2\alpha \geq 2\alpha - 1,$$

which always holds. Consequently there always exists a value $\xi^2 \in \left[\underline{\xi}^2, \overline{\xi}^2\right]$ such that a rank-order tournament with that noise level has a higher selection efficiency than the noiseless all-pay contest. We can state this result as follows:

Proposition 3.1 (Selection Efficiency)**.** *For any $v > 0$, $\alpha > 1$, there exists a noise level $\xi^2 \in \left[\underline{\xi}^2, \overline{\xi}^2\right]$ such that a noisy contest with noise ξ^2 has a higher selection efficiency than the noiseless contest or a contest with small noise in the sense of Ewerhart (2017)*

Proposition 3.1 implies that, when the principal can coarsely choose between different noise levels, the principal should always choose a noise level in $\left[\underline{\xi}^2, \overline{\xi}^2\right]$ over a contest with no or small noise in the sense of Ewerhart (2017).

3.4.2 Aggregate Effort

So far, I have shown that for any values v¿0 and $\alpha > 1$, there exists a noise level $\xi^2 \in \left[\underline{\xi}^2, \overline{\xi}^2\right]$ such that the rank-order tournament with that noise level has a higher selection efficiency than the corresponding all-pay contest. Since many contests are used not only to select the best agent, but also to motivate the agents, it is interesting to see to what extend these two aims are compatible. In this subsection I compare the (expected) aggregate effort submissions of the two agents in the two contests.

I start with a comparison of the h-type's equilibrium effort submissions. We have that effort in the rank-order tournament is higher if

$$x_h^* = \frac{v}{2\xi} \geq \frac{2}{3}\sqrt{\frac{v}{\alpha}} = \mathbb{E}^{\mathrm{AP}}\left[x_h\right] \Leftrightarrow \xi^2 \leq \tilde{\xi}^2 := \frac{9}{16}v\alpha.$$

This puts another upper bound on the noise level. I proceed by checking whether this is compatible with the previously derived lower bound—i.e., the optimal noise level in terms of selection efficiency—that permits a pure-strategy equilibrium in the rank-order tournament. In other words, I check whether the interval $\xi^2 \in \left[\underline{\xi}^2, \tilde{\xi}^2\right]$ is non-empty. This is the case if

$9/16 \cdot v\alpha \geq \underline{\xi}^2$. Plugging in an explicit value for $\underline{\xi}^2$ and rearranging yields the quadratic inequality

$$\alpha^2 - \frac{16}{9}\alpha + \frac{8}{9} \geq 0,$$

which, upon closer inspection, always holds. This means that the equilibrium effort submission of the h-type in the rank-order tournament with sufficiently small noise is higher than the expected equilibrium effort submission of the h-type in the all-pay contest.

I now move on to look at the equilibrium effort submissions of the l-type. Remember that that in the all-pay auction we had

$$\mathbb{E}^{\text{AP}}\left[x_l\right] = \frac{1}{\alpha}\mathbb{E}^{\text{AP}}\left[x_h\right],$$

whereas in the rank-order tournament, we had

$$x_l^* = \frac{1}{\alpha}x_h^*.$$

Comparing the last two expressions, we can state the following:

Lemma 3.4 (Relative Difference in Effort submissions)**.** *The relative difference between the (expected) effort submissions of both agents is the same in the noisy and in the noiseless setting.*

While this observation might be interesting in its own right, for us it is important because it implies that the (expected) aggregate effort of the l-type is higher in the noisy contest than in the noiseless contest whenever the same is true for the h-type. Since we have previously shown that there always exists a noise level such that this is true for the h-type, Lemma 3.4 immediately yields the following proposition.

Proposition 3.2 (Aggregate Effort)**.** *For any $v > 0$, $\alpha > 1$, there exists a noise level $\xi^2 \in \left[\underline{\xi}^2, \overline{\xi}^2\right]$ such that a noisy contest with noise ξ^2 induces both a higher selection efficiency and higher aggregate effort than a noiseless contest or a contest with small noise in the sense of Ewerhart (2017).*

Proof: Follows from Proposition 3.1 and Lemma 3.4. □

We can further say the following:

Corollary 3.1 (Aligned Objectives)**.** *Out of the noise levels considered here, the selection-efficiency maximizing and the effort maximizing noise levels coincide.*

Proof: Follows from Proposition 3.1 and Proposition 3.2 together with the observations that $\partial x_h^*/\partial\xi < 0$, $\partial x_l^*/\partial\xi < 0$, and $\partial \Pr^{\text{RO}}[\cdot]/\partial\xi < 0$. □

This is an important result, because it shows that selection efficiency and effort maximization are not mutually exclusive. This is not always the case in contest theory and, indeed, there are other scenarios in which these two objectives cannot be achieved at the same time Stracke et al. (2015). Thereby it also highlights the extent to which excessive workplace-surveillance can be harmful in the promotion process—it can adversely affect both effort incentives and meritocratic selection.

Corollary 1 has an underlying intuition which is, perhaps, not immediately obvious: Remember hat the l-type's equilibrium effort submission is a fixed share, $1/\alpha$, of the h-type's equilibrium effort submission. This implies that the absolute difference in their effort submissions increases in aggregate effort. One could say that, as effort incentives increase, the agent's equilibrium behavior grows more and more asymmetric in absolute terms. More asymmetric behavior, in turn, increases the informativeness of the agent's strategies and, hence, the probability of selecting the high-ability agent.

3.5 Employee Welfare

So far we have seen that a less precise monitoring scheme can be beneficial from the manager's point of view. Precisely, the noisy contest maximizes manager welfare if her objective is to maximize (i) selection efficiency, (ii) aggregate output, or (iii) a convex combination thereof. The manager may of course have other objectives—e.g., maximizing winner effort—which I do not consider here.

One question that follows is: If the manager is better off, does this imply that the employees are worse off? As it turns out, this is not the case. First, notice that if we are in the case where the manager sets the noise level such that it maximizes both selection efficiency and aggregate effort among the noise levels that I consider, the welfare of the less talented employee is unaffected by this. In both cases, her expected utility will be zero. Instead, let us look at the expected utility of the more talented agent. In the all-pay auction, it is given by

$$\Pi_h^{\text{AP}} = \left(1 - \frac{1}{2\alpha}\right) v - \int_0^{c_l^{-1}(v)} f_h(x_h) \cdot c_h(x_h) dx_h = \left(1 - \frac{1}{\alpha}\right) v.$$

We can now compare this expected utility with the expected utility than a talented employee would receive in the optimal noisy contest,

$$\Pi_h^{\mathrm{RO}} = \left[\frac{v}{2\xi^2}\left(1-\frac{1}{\alpha}\right)+\frac{1}{2}\right]v-\left(\frac{v}{2\xi}\right)^2.$$

Comparing these two expressions yields

$$\left[\frac{v}{2\xi^2}\left(1-\frac{1}{\alpha}\right)+\frac{1}{2}\right]v-\left(\frac{v}{2\xi}\right)^2 \geq \left(1-\frac{1}{\alpha}\right)v \Leftrightarrow 1 \leq \frac{2\alpha-1}{\alpha},$$

which holds because $1 \leq \alpha$ by definition. These observations allow us to state the following proposition.

Proposition 3.3 (Employee Welfare)**.** *There always exists a noise monitoring scheme which is a Pareto improvement over a monitoring with no noise or small noise in the sense of Ewerhart (2017).*

Notice that at least two parties, the high ability agent and the firm, are better off under the optimal noisy monitoring scheme. The expected utility of the low ability agent is unaffected. In principal it might be possible to achieve strict Pareto improvements over the noiseless contest, in which also the low ability agent earns a positive rent on expectation. No such scheme, however, can be optimal from the firm's perspective.

Overall, Proposition 3.3 is reassuring. At first, one might be inclined to think that the additional welfare generated for the firm must have been taken away from some other party in the game. However, this is not the case; indeed, no player is ever worse off under the optimal noisy monitoring scheme than under a noiseless scheme. This is partly due to our assumption of convex cost. In the mixed-strategy equilibrium of the noiseless case, employees will sometimes exert a very high effort. Due to the exponential increase of effort cost, this, together with an increase in the winning probability of the high ability agent, outweighs the increase of the expected cost due to a higher expected effort submission under the noisy monitoring scheme.

3.6 Discussion: What Kind of Noise Can the Manager Employ Strategically?

In this section, I discuss what kind of noise the manager can employ strategically in the sense described in this paper. I first provide an example that shows how the type of noise considered here is plausible in the relevant settings. I then proceed by showing that other forms of noise—which are distinct to those we would expect due to a lack of monitoring—do not have the same effects.

3.6.1 Example

Consider any type of task whose outcome is dependent on a team effort. The manager can observe the team's output without any noise, but she cannot precisely evaluate the two team-member's individual contributions. For concreteness, consider a store that employs two salespersons. Suppose that there is a baseline number of items that the store sells in a given timeframe. For simplicity, assume that this baseline number is constant and that all the variation in sales beyond that number is due to the salespersons' efforts put into persuading customers to buy an item. The manager of the store can observe the number of items sold beyond the baseline number perfectly. However, she can only imperfectly observe each salesperson's individual contribution to the sales. She will randomly over- or underestimate the number of sales by salesperson i by $\epsilon_i \overset{\mathrm{d}}{\sim} \mathcal{U}[0,1]$ and, hence, over- or underestimate the number of sales by salesperson $j \neq i$ by $\epsilon_j = 1 - \epsilon_i$. Given these distributional assumptions, we have that the difference of the random over- or underestimations of an individual agent's effort, $\Delta\epsilon = \epsilon_i - \epsilon_j = 2\epsilon_i - 1$, is distributed according to

$$\Delta\epsilon \overset{\mathrm{d}}{\sim} \mathcal{U}[-1,1],$$

which corresponds to the type of noise considered in this paper. A possible explanation for this type of noise would be as follows: Suppose that the store has two floors with one of the salespersons working on each floor. The manager can only monitor one floor at the time and, hence, splits her time between monitoring the two floors. Whether a customer enters the floor that is currently monitored by the manager is random. Hence, whether the manager observes a particular salesperson making a particular sale is random, too. As a result, there is a positive probability that the manager will miss at least one sale. Without further modeling, it is unclear to which employee the manager would attribute the sale. However, two heuristics appear plausible. Firstly, the manager could attribute an unobserved sale randomly to either of the salespersons. Secondly, the manager could attribute an unobserved sale to the salesperson of which the manager had previously observed more sales. Importantly, both of these heuristics imply that the manager will (weakly) overestimate the number of sales of the salesperson of which the manager observed more sales previously. Now suppose that that the manager of the store can install video-cameras on both floors that allow her to monitor them at the same time. This allows her to precisely monitor the sales attributed to each salesperson and, hence, their respective efforts. This example highlights two things: First, there are many environments in which the type of noise (correlated) considered here is plausible. Namely, all environments in which a manager can observe total output but not individual contributions perfectly. Second, it shows that the manager often only has a coarse choice between different noise levels—i.e., some noise in the baseline case vs. essentially no noise in the case with cameras.

3.6.2 Ex-Post Noise

I want to point out is that the manager cannot simply choose any kind of noise to increase selection efficiency and effort. Indeed, we can think of two types of noise. The first type of noise affects the relationship between an employee's effort submission and her performance rank. This is the type of noise we have considered so far. It makes sure that the employee with the lower effort submission can sometimes still outperform her opponent and, hence, win. The second type of noise affects the relationship between an employee's performance rank and her probability of being promoted. This type of noise makes sure that the employee with the weaker performance can sometimes win.

While the first type of noise can be beneficial to the firm—this is what I have shown in the previous sections—the second type of noise cannot. Consider the easiest way in which the second type of noise can be implemented: The manager of the firm can commit to a promotion rule which promotes the employee with the higher performance only with probability $\rho \in (1/2, 1)$. We can see easily that unlike the first type of noise, the second type of noise alone will never lead to a pure-strategy equilibrium. Suppose, for example, that both employees would submit an effort of zero and, hence, incur no effort cost. Both employees would receive an expected utility of $1/2v$. If one employee would deviate to an effort level marginally above zero, she would receive ρv, which, since $1/2 < \rho$, would be beneficial. The same exercise can be made for any effort level between zero and $c_l^{-1}\left((2\rho - 1)\, v\right)$—the highest effort level the low-ability employee would be willing to submit given that she believes that she will have the highest performance at that effort level with probability one. Suppose that $\rho v - c_h\left(c_l^{-1}\left((2\rho - 1)\, v\right)\right) > 0$; then, the high-ability employee might, in principal, be willing to exert effort larger or equal to $c_l^{-1}\left((2\rho - 1)\, v\right)$. However, the low-ability agent would not and, hence, would have an incentive to submit no effort at all. This destroys also the high-ability employee's incentive to submit an effort level that high and, hence, there will be no pure strategy equilibrium.

Effectively, in this setting, we an simply rewrite the prize from winning as $\rho \cdot v + (1-\rho) \cdot 0$. The prize from losing instead becomes $\rho \cdot 0 + (1-\rho) \cdot v$. The prize difference, however, is still positive. Since we already know from Lemma 1, however, that the prize difference is not important for selection efficiency as long as it is strictly positive, we also know that selection efficiency is unaffected by this type of noise. The following proposition summarizes this.

Proposition 3.4 (Ex-Post Noise)**.** *Suppose that the winner receives the prize only with probability* $\rho \in (1/2, 1)$*. With probability* $1-\rho$*, the prize is given to the loser instead. Selection efficiency is unaffected compared to the standard all-pay contest described in Lemma 1.*

3.7 Conclusion

In this paper, I have demonstrated that the kind of noise induced through less stringent employee surveillance may be beneficial in three ways. It increases (i) selection efficiency, (ii) aggregate output, and (iii) employee welfare—and as a result, social welfare under the assumption that the principal cares about (i), (ii), or a convex combination thereof.

The result comes with the caveat that I do not provide a solution to the optimal noise level. Providing a more comprehensive analysis of contests with noise too low for the existence of pure strategy equilibria is a long standing issue within the literature. As my practical application has demonstrated, while some things can be said even with current tools, a more comprehensive analysis will likely reveal additional insights on how organizations are optimally designed.

Because the topic considered in this paper has a clear practical application, empirical research in a similar direction would yield crucial additional insights. There are several difficulties associated with testing my results with observational data. First, it is difficult to find a natural variation of surveillance intensities (with some exceptions, like the introduction of Video Assistant Referees (VARs) in football). Secondly, and perhaps even more crucially, it is very difficult to measure whether the winner of a promotion contest was indeed the most talented one. As such an experimental investigation is most likely to provide insights.

3.A References

Bar-Isaac, H., and J. Deb. 2021. “Reputation with opportunities for coasting.” *Journal of the European Economic Association* 19 (1): 200–236.

Barron, D., G. Georgiadis, and J. Swinkels. 2020. “Optimal contracts with a risk-taking agent.” *Theoretical Economics* 15 (2): 715–761.

Baye, M. R., D. Kovenock, and C. G. De Vries. 1993. “Rigging the lobbying process: an application of the all-pay auction.” *The American Economic Review* 83 (1): 289–294.

———. 1994. “The solution to the Tullock rent-seeking game when R¿ 2: Mixed-strategy equilibria and mean dissipation rates.” *Public Choice* 81 (3): 363–380.

———. 1996. “The all-pay auction with complete information.” *Economic Theory* 8 (2): 291–305.

Brown, J., and D. B. Minor. 2014. “Selecting the best? Spillover and shadows in elimination tournaments.” *Management Science* 60 (12): 3087–3102.

Drugov, M., and D. Ryvkin. 2017. “Biased contests for symmetric players.” *Games and Economic Behavior* 103:116–144.

Ederer, F. 2010. “Feedback and motivation in dynamic tournaments.” *Journal of Economics & Management Strategy* 19 (3): 733–769.

Ewerhart, C. 2015. “Mixed equilibria in Tullock contests.” *Economic Theory* 60 (1): 59–71.

———. 2017. “Contests with small noise and the robustness of the all-pay auction.” *Games and Economic Behavior* 105:195–211.

Fang, D., and T. H. Noe. 2022. “Less competition, more meritocracy?” *Journal of Labor Economics.*

Hillman, A. L., and J. G. Riley. 1989. “Politically contestable rents and transfers.” *Economics & Politics* 1 (1): 17–39.

Hvide, H. K., and E. G. Kristiansen. 2003. "Risk taking in selection contests." *Games and Economic Behavior* 42 (1): 172–179.

Kawamura, K., and I. M. de Barreda. 2014. "Biasing selection contests with ex-ante identical agents." *Economics Letters* 123 (2): 240–243.

Konrad, K. A. 2009. *Strategy and dynamics in contests.* Oxford University Press.

Lazear, E. P., and S. Rosen. 1981. "Rank-order tournaments as optimum labor contracts." *Journal of Political Economy* 89 (5): 841–864.

Letina, I., S. Liu, and N. Netzer. 2020. "Optimal Contest Design: A General Approach." *Working Paper.*

Meyer, M. A. 1991. "Learning from coarse information: Biased contests and career profiles." *The Review of Economic Studies* 58 (1): 15–41.

Morgan, J., J. Tumlinson, and F. J. Vardy. 2018. "The limits of meritocracy." *Working Paper.*

Siegel, R. 2009. "All-pay contests." *Econometrica* 77 (1): 71–92.

Stracke, R., W. Höchtl, R. Kerschbamer, and U. Sunde. 2015. "Incentives and selection in promotion contests: Is it possible to kill two birds with one stone?" *Managerial and Decision Economics* 36 (5): 275–285.

Zhao, R. R. 2008. "All-or-nothing monitoring." *American Economic Review* 98 (4): 1619–1628.

Previous doctoral theses in the Department of Economics, Gothenburg

Avhandlingar publicerade innan serien Ekonomiska Studier startades
(Theses published before the series Ekonomiska Studier was started):

Östman, Hugo (1911), Norrlands ekonomiska utveckling
Moritz, Marcus (1911), Den svenska tobaksindustrien
Sundbom, I. (1933), Prisbildning och ändamålsenlighet
Gerhard, I. (1948), Problem rörande Sveriges utrikeshandel 1936/38
Hegeland, Hugo (1951), The Quantity Theory of Money
Mattsson, Bengt (1970), Cost-Benefit analys
Rosengren, Björn (1975), Valutareglering och nationell ekonomisk politik
Hjalmarsson, Lennart (1975), Studies in a Dynamic Theory of Production and its Applications
Örtendahl, Per-Anders (1975), Substitutionsaspekter på produktionsprocessen vid massaframställning
Anderson, Arne M. (1976), Produktion, kapacitet och kostnader vid ett helautomatiskt emballageglasbruk
Ohlsson, Olle (1976), Substitution och odelbarheter i produktionsprocessen vid massaframställning
Gunnarsson, Jan (1976), Produktionssystem och tätortshierarki – om sambandet mellan rumslig och ekonomisk struktur
Köstner, Evert (1976), Optimal allokering av tid mellan utbildning och arbete
Wigren, Rune (1976), Analys av regionala effektivitetsskillnader inom industribranscher
Wästlund, Jan (1976), Skattning och analys av regionala effektivitetsskillnader inom industribranscher
Flöjstad, Gunnar (1976), Studies in Distortions, Trade and Allocation Problems
Sandelin, Bo (1977), Prisutveckling och kapitalvinster på bostadsfastigheter
Dahlberg, Lars (1977), Empirical Studies in Public Planning
Lönnroth, Johan (1977), Marxism som matematisk ekonomi
Johansson, Börje (1978), Contributions to Sequential Analysis of Oligopolistic Competition

Ekonomiska Studier, utgivna av Nationalekonomiska institutionen vid Göteborgs Universitet. Nr 1 och 4 var inte doktorsavhandlingar. (The contributions to the department series 'Ekonomiska Studier' where no. 1 and 4 were no doctoral theses):

2. **Ambjörn, Erik** (1959), Svenskt importberoende 1926-1956: en ekonomisk-statistisk kartläggning med kommentarer
3. **Landgren, K-G.** (1960), Den "Nya ekonomien" i Sverige: J.M. Keynes, E. Wigfors och utecklingen 1927-39
5. **Bigsten, Arne** (1979), Regional Inequality and Development: A Case Study of Kenya
6. **Andersson, Lars** (1979), Statens styrning av de kommunala budgetarnas struktur (Central Government Influence on the Structure of the Municipal Budget)
7. **Gustafsson, Björn** (1979), Inkomst- och uppväxtförhållanden (Income and Family Background)
8. **Granholm, Arne** (1981), Interregional Planning Models for the Allocation of Private and Public Investments
9. **Lundborg, Per** (1982), Trade Policy and Development: Income Distributional Effects in the Less Developed Countries of the US and EEC Policies for Agricultural

Commodities

10. **Juås, Birgitta** (1982), Värdering av risken för personskador. En jämförande studie av implicita och explicita värden. (Valuation of Personal Injuries. A comparison of Explicit and Implicit Values)
11. **Bergendahl, Per-Anders** (1982), Energi och ekonomi - tillämpningar av input-output analys (Energy and the Economy - Applications of Input-Output Analysis)
12. **Blomström, Magnus** (1983), Foreign Investment, Technical Efficiency and Structural Change - Evidence from the Mexican Manufacturing Industry
13. **Larsson, Lars-Göran** (1983), Comparative Statics on the Basis of Optimization Methods
14. **Persson, Håkan** (1983), Theory and Applications of Multisectoral Growth Models
15. **Sterner, Thomas** (1986), Energy Use in Mexican Industry.
16. **Flood, Lennart** (1986), On the Application of Time Use and Expenditure Allocation Models.
17. **Schuller, Bernd-Joachim** (1986), Ekonomi och kriminalitet - en empirisk undersökning av brottsligheten i Sverige (Economics of crime - an empirical analysis of crime in Sweden)
18. **Walfridson, Bo** (1987), Dynamic Models of Factor Demand. An Application to Swedish Industry.
19. **Stålhammar, Nils-Olov** (1987), Strukturomvandling, företagsbeteende och förväntningsbildning inom den svenska tillverkningsindustrin (Structural Change, Firm Behaviour and Expectation Formation in Swedish Manufactury)
20. **Anxo, Dominique** (1988), Sysselsättningseffekter av en allmän arbetstidsförkortning (Employment effects of a general shortage of the working time)
21. **Mbelle, Ammon** (1988), Foreign Exchange and Industrial Development: A Study of Tanzania.
22. **Ongaro, Wilfred** (1988), Adoption of New Farming Technology: A Case Study of Maize Production in Western Kenya.
23. **Zejan, Mario** (1988), Studies in the Behavior of Swedish Multinationals.
24. **Görling, Anders** (1988), Ekonomisk tillväxt och miljö. Förorenings-struktur och ekonomiska effekter av olika miljövårdsprogram. (Economic Growth and Environment. Pollution Structure and Economic Effects of Some Environmental Programs).
25. **Aguilar, Renato** (1988), Efficiency in Production: Theory and an Application on Kenyan Smallholders.
26. **Kayizzi-Mugerwa, Steve** (1988), External Shocks and Adjustment in Zambia.
27. **Bornmalm-Jardelöw, Gunilla** (1988), Högre utbildning och arbetsmarknad (Higher Education and the Labour Market)
28. **Tansini, Ruben** (1989), Technology Transfer: Dairy Industries in Sweden and Uruguay.
29. **Andersson, Irene** (1989), Familjebeskattning, konsumtion och arbetsutbud - En ekonometrisk analys av löne- och inkomstelasticiteter samt policysimuleringar för svenska hushåll (Family Taxation, Consumption and Labour Supply - An Econometric Analysis of Wage and Income Elasticities and Policy Simulations for Swedish Households)
30. **Henrekson, Magnus** (1990), An Economic Analysis of Swedish Government Expenditure
31. **Sjöö, Boo** (1990), Monetary Policy in a Continuous Time Dynamic Model for Sweden

32. **Rosén, Åsa** (1991), Contributions to the Theory of Labour Contracts.
33. **Loureiro, Joao M. de Matos** (1992), Foreign Exchange Intervention, Sterilization and Credibility in the EMS: An Empirical Study
34. **Irandoust, Manuchehr** (1993), Essays on the Behavior and Performance of the Car Industry
35. **Tasiran, Ali Cevat** (1993), Wage and Income Effects on the Timing and Spacing of Births in Sweden and the United States
36. **Milopoulos, Christos** (1993), Investment Behaviour under Uncertainty: An Econometric Analysis of Swedish Panel Data
37. **Andersson, Per-Åke** (1993), Labour Market Structure in a Controlled Economy: The Case of Zambia
38. **Storrie, Donald W.** (1993), The Anatomy of a Large Swedish Plant Closure
39. **Semboja, Haji Hatibu Haji** (1993), Energy and Development in Kenya
40. **Makonnen, Negatu** (1993), Labor Supply and the Distribution of Economic Well-Being: A Case Study of Lesotho
41. **Julin, Eva** (1993), Structural Change in Rural Kenya
42. **Durevall, Dick** (1993), Essays on Chronic Inflation: The Brazilian Experience
43. **Veiderpass, Ann** (1993), Swedish Retail Electricity Distribution: A Non-Parametric Approach to Efficiency and Productivity Change
44. **Odeck, James** (1993), Measuring Productivity Growth and Efficiency with Data Envelopment Analysis: An Application on the Norwegian Road Sector
45. **Mwenda, Abraham** (1993), Credit Rationing and Investment Behaviour under Market Imperfections: Evidence from Commercial Agriculture in Zambia
46. **Mlambo, Kupukile** (1993), Total Factor Productivity Growth: An Empirical Analysis of Zimbabwe's Manufacturing Sector Based on Factor Demand Modelling
47. **Ndung'u, Njuguna** (1993), Dynamics of the Inflationary Process in Kenya
48. **Modén, Karl-Markus** (1993), Tax Incentives of Corporate Mergers and Foreign Direct Investments
49. **Franzén, Mikael** (1994), Gasoline Demand - A Comparison of Models
50. **Heshmati, Almas** (1994), Estimating Technical Efficiency, Productivity Growth And Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture
51. **Salas, Osvaldo** (1994), Efficiency and Productivity Change: A Micro Data Case Study of the Colombian Cement Industry
52. **Bjurek, Hans** (1994), Essays on Efficiency and Productivity Change with Applications to Public Service Production
53. **Cabezas Vega, Luis** (1994), Factor Substitution, Capacity Utilization and Total Factor Productivity Growth in the Peruvian Manufacturing Industry
54. **Katz, Katarina** (1994), Gender Differentiation and Discrimination. A Study of Soviet Wages
55. **Asal, Maher** (1995), Real Exchange Rate Determination and the Adjustment Process: An Empirical Study in the Cases of Sweden and Egypt
56. **Kjulin, Urban** (1995), Economic Perspectives on Child Care
57. **Andersson, Göran** (1995), Volatility Forecasting and Efficiency of the Swedish Call Options Market
58. **Forteza, Alvaro** (1996), Credibility, Inflation and Incentive Distortions in the Welfare State
59. **Locking, Håkan** (1996), Essays on Swedish Wage Formation
60. **Välilä, Timo** (1996), Essays on the Credibility of Central Bank Independence

61. **Yilma, Mulugeta** (1996), Measuring Smallholder Efficiency: Ugandan Coffee and Food-Crop Production
62. **Mabugu, Ramos E.** (1996), Tax Policy Analysis in Zimbabwe Applying General Equilibrium Models
63. **Johansson, Olof** (1996), Welfare, Externalities, and Taxation; Theory and Some Road Transport Applications.
64. **Chitiga, Margaret** (1996), Computable General Equilibrium Analysis of Income Distribution Policies in Zimbabwe
65. **Leander, Per** (1996), Foreign Exchange Market Behavior Expectations and Chaos
66. **Hansen, Jörgen** (1997), Essays on Earnings and Labor Supply
67. **Cotfas, Mihai** (1997), Essays on Productivity and Efficiency in the Romanian Cement Industry
68. **Horgby, Per-Johan** (1997), Essays on Sharing, Management and Evaluation of Health Risks
69. **Nafar, Nosratollah** (1997), Efficiency and Productivity in Iranian Manufacturing Industries
70. **Zheng, Jinghai** (1997), Essays on Industrial Structure, Technical Change, Employment Adjustment, and Technical Efficiency
71. **Isaksson, Anders** (1997), Essays on Financial Liberalisation in Developing Countries: Capital mobility, price stability, and savings
72. **Gerdin, Anders** (1997), On Productivity and Growth in Kenya, 1964-94
73. **Sharifi, Alimorad** (1998), The Electricity Supply Industry in Iran: Organization, performance and future development
74. **Zamanian, Max** (1997), Methods for Mutual Fund Portfolio Evaluation: An application to the Swedish market
75. **Manda, Damiano Kulundu** (1997), Labour Supply, Returns to Education, and the Effect of Firm Size on Wages: The case of Kenya
76. **Holmén, Martin** (1998), Essays on Corporate Acquisitions and Stock Market Introductions
77. **Pan, Kelvin** (1998), Essays on Enforcement in Money and Banking
78. **Rogat, Jorge** (1998), The Value of Improved Air Quality in Santiago de Chile
79. **Peterson, Stefan** (1998), Essays on Large Shareholders and Corporate Control
80. **Belhaj, Mohammed** (1998), Energy, Transportation and Urban Environment in Africa: The Case of Rabat-Salé, Morocco
81. **Mekonnen, Alemu** (1998), Rural Energy and Afforestation: Case Studies from Ethiopia
82. **Johansson, Anders** (1998), Empirical Essays on Financial and Real Investment Behavior
83. **Köhlin, Gunnar** (1998), The Value of Social Forestry in Orissa, India
84. **Levin, Jörgen** (1998), Structural Adjustment and Poverty: The Case of Kenya
85. **Ncube, Mkhululi** (1998), Analysis of Employment Behaviour in Zimbabwe
86. **Mwansa, Ladslous** (1998), Determinants of Inflation in Zambia
87. **Agnarsson, Sveinn** (1998), Of Men and Machines: Essays in Applied Labour and Production Economics
88. **Kadenge, Phineas** (1998), Essays on Macroeconomic Adjustment in Zimbabwe: Inflation, Money Demand, and the Real Exchange Rate
89. **Nyman, Håkan** (1998), An Economic Analysis of Lone Motherhood in Sweden
90. **Carlsson, Fredrik** (1999), Essays on Externalities and Transport
91. **Johansson, Mats** (1999), Empirical Studies of Income Distribution

92. **Alemu, Tekie** (1999), Land Tenure and Soil Conservation: Evidence from Ethiopia
93. **Lundvall, Karl** (1999), Essays on Manufacturing Production in a Developing Economy: Kenya 1992-94
94. **Zhang, Jianhua** (1999), Essays on Emerging Market Finance
95. **Mlima, Aziz Ponary** (1999), Four Essays on Efficiency and Productivity in Swedish Banking
96. **Davidsen, Björn-Ivar** (2000), Bidrag til den økonomisk-metodologiske tenkningen (Contributions to the Economic Methodological Thinking)
97. **Ericson, Peter** (2000), Essays on Labor Supply
98. **Söderbom, Måns** (2000), Investment in African Manufacturing: A Microeconomic Analysis
99. **Höglund, Lena** (2000), Essays on Environmental Regulation with Applications to Sweden
100. **Olsson, Ola** (2000), Perspectives on Knowledge and Growth
101. **Meuller, Lars** (2000), Essays on Money and Credit
102. **Österberg, Torun** (2000), Economic Perspectives on Immigrants and Intergenerational Transmissions
103. **Kalinda Mkenda, Beatrice** (2001), Essays on Purchasing Power Parity, RealExchange Rate, and Optimum Currency Areas
104. **Nerhagen, Lena** (2001), Travel Demand and Value of Time - Towards an Understanding of Individuals Choice Behavior
105. **Mkenda, Adolf** (2001), Fishery Resources and Welfare in Rural Zanzibar
106. **Eggert, Håkan** (2001), Essays on Fisheries Economics
107. **Andrén, Daniela** (2001), Work, Sickness, Earnings, and Early Exits from the Labor Market. An Empirical Analysis Using Swedish Longitudinal Data
108. **Nivorozhkin, Eugene** (2001), Essays on Capital Structure
109. **Hammar, Henrik** (2001), Essays on Policy Instruments: Applications to Smoking and the Environment
110. **Nannyonjo, Justine** (2002), Financial Sector Reforms in Uganda (1990-2000): Interest Rate Spreads, Market Structure, Bank Performance and Monetary Policy
111. **Wu, Hong** (2002), Essays on Insurance Economics
112. **Linde-Rahr, Martin** (2002), Household Economics of Agriculture and Forestry in Rural Vienam
113. **Maneschiöld, Per-Ola** (2002), Essays on Exchange Rates and Central Bank Credibility
114. **Andrén, Thomas** (2002), Essays on Training, Welfare and Labor Supply
115. **Granér, Mats** (2002), Essays on Trade and Productivity: Case Studies of Manufacturing in Chile and Kenya
116. **Jaldell, Henrik** (2002), Essays on the Performance of Fire and Rescue Services
117. **Alpizar, Francisco, R.** (2002), Essays on Environmental Policy-Making in Developing Countries: Applications to Costa Rica
118. **Wahlberg, Roger** (2002), Essays on Discrimination, Welfare and Labor Supply
119. **Piculescu, Violeta** (2002), Studies on the Post-Communist Transition
120. **Pylkkänen, Elina** (2003), Studies on Household Labor Supply and Home Production
121. **Löfgren, Åsa** (2003), Environmental Taxation – Empirical and Theoretical Applications
122. **Ivaschenko, Oleksiy** (2003), Essays on Poverty, Income Inequality and Health in Transition Economies

123. **Lundström, Susanna** (2003), On Institutions, Economic Growth and the Environment
124. **Wambugu, Anthony** (2003), Essays on Earnings and Human Capital in Kenya
125. **Adler, Johan** (2003), Aspects of Macroeconomic Saving
126. **Erlandsson, Mattias** (2003), On Monetary Integration and Macroeconomic Policy
127. **Brink, Anna** (2003), On the Political Economy of Municipality Break-Ups
128. **Ljungwall, Christer** (2003), Essays on China's Economic Performance During the Reform Period
129. **Chifamba, Ronald** (2003), Analysis of Mining Investments in Zimbabwe
130. **Muchapondwa, Edwin** (2003), The Economics of Community-Based Wildlife Conservation in Zimbabwe
131. **Hammes, Klaus** (2003), Essays on Capital Structure and Trade Financing
132. **Abou-Ali, Hala** (2003), Water and Health in Egypt: An Empirical Analysis
133. **Simatele, Munacinga** (2004), Financial Sector Reforms and Monetary Policy in Zambia
134. **Tezic, Kerem** (2004), Essays on Immigrants' Economic Integration
135. INSTÄLLD
136. **Gjirja, Matilda** (2004), Efficiency and Productivity in Swedish Banking
137. **Andersson, Jessica** (2004), Welfare Environment and Tourism in Developing Countries
138. **Chen, Yinghong** (2004), Essays on Voting Power, Corporate Governance and Capital Structure
139. **Yesuf, Mahmud** (2004), Risk, Time and Land Management under Market Imperfections: Applications to Ethiopia
140. **Kateregga, Eseza** (2005), Essays on the Infestation of Lake Victoria by the Water Hyacinth
141. **Edvardsen, Dag Fjeld** (2004), Four Essays on the Measurement of Productive Efficiency
142. **Lidén, Erik** (2005), Essays on Information and Conflicts of Interest in Stock Recommendations
143. **Dieden, Sten** (2005), Income Generation in the African and Coloured Population – Three Essays on the Origins of Household Incomes in South Africa
144. **Eliasson, Marcus** (2005), Individual and Family Consequences of Involuntary Job Loss
145. **Mahmud, Minhaj** (2005), Measuring Trust and the Value of Statistical Lives: Evidence from Bangladesh
146. **Lokina, Razack Bakari** (2005), Efficiency, Risk and Regulation Compliance: Applications to Lake Victoria Fisheries in Tanzania
147. **Jussila Hammes, Johanna** (2005), Essays on the Political Economy of Land Use Change
148. **Nyangena, Wilfred** (2006), Essays on Soil Conservation, Social Capital and Technology Adoption
149. **Nivorozhkin, Anton** (2006), Essays on Unemployment Duration and Programme Evaluation
150. **Sandén, Klas** (2006), Essays on the Skill Premium
151. **Deng, Daniel** (2006), Three Essays on Electricity Spot and Financial Derivative Prices at the Nordic Power Exchange
152. **Gebreeyesus, Mulu** (2006), Essays on Firm Turnover, Growth, and Investment Behavior in Ethiopian Manufacturing

153. **Islam, Nizamul Md.** (2006), Essays on Labor Supply and Poverty: A Microeconometric Application
154. **Kjaer, Mats** (2006), Pricing of Some Path-Dependent Options on Equities and Commodities
155. **Shimeles, Abebe** (2006), Essays on Poverty, Risk and Consumption Dynamics in Ethiopia
156. **Larsson, Jan** (2006), Four Essays on Technology, Productivity and Environment
157. **Congdon Fors, Heather** (2006), Essays in Institutional and Development Economics
158. **Akpalu, Wisdom** (2006), Essays on Economics of Natural Resource Management and Experiments
159. **Daruvala, Dinky** (2006), Experimental Studies on Risk, Inequality and Relative Standing
160. **García, Jorge** (2007), Essays on Asymmetric Information and Environmental Regulation through Disclosure
161. **Bezabih, Mintewab** (2007), Essays on Land Lease Markets, Productivity, Biodiversity, and Environmental Variability
162. **Visser, Martine** (2007), Fairness, Reciprocity and Inequality: Experimental Evidence from South Africa
163. **Holm, Louise** (2007), A Non-Stationary Perspective on the European and Swedish Business Cycle
164. **Herbertsson, Alexander** (2007), Pricing Portfolio Credit Derivatives
165. **Johansson, Anders C.** (2007), Essays in Empirical Finance: Volatility, Interdependencies, and Risk in Emerging Markets
166. **Ibáñez Díaz, Marcela** (2007), Social Dilemmas: The Role of Incentives, Norms and Institutions
167. **Ekbom, Anders** (2007), Economic Analysis of Soil Capital, Land Use and Agricultural Production in Kenya
168. **Sjöberg, Pål** (2007), Essays on Performance and Growth in Swedish Banking
169. **Palma Aguirre, Grisha Alexis** (2008), Explaining Earnings and Income Inequality in Chile
170. **Akay, Alpaslan** (2008), Essays on Microeconometrics and Immigrant Assimilation
171. **Carlsson, Evert** (2008), After Work – Investing for Retirement
172. **Munshi, Farzana** (2008), Essays on Globalization and Occupational Wages
173. **Tsakas, Elias** (2008), Essays on Epistemology and Evolutionary Game Theory
174. **Erlandzon, Karl** (2008), Retirement Planning: Portfolio Choice for Long-Term Investors
175. **Lampi, Elina** (2008), Individual Preferences, Choices, and Risk Perceptions – Survey Based Evidence
176. **Mitrut, Andreea** (2008), Four Essays on Interhousehold Transfers and Institutions in Post-Communist Romania
177. **Hansson, Gustav** (2008), Essays on Social Distance, Institutions, and Economic Growth
178. **Zikhali, Precious** (2008), Land Reform, Trust and Natural Resource Management in Africa
179. **Tengstam, Sven** (2008), Essays on Smallholder Diversification, Industry Location, Debt Relief, and Disability and Utility
180. **Boman, Anders** (2009), Geographic Labour Mobility – Causes and Consequences
181. **Qin, Ping** (2009), Risk, Relative Standing and Property Rights: Rural Household Decision-Making in China

182. **Wei, Jiegen** (2009), Essays in Climate Change and Forest Management
183. **Belu, Constantin** (2009), Essays on Efficiency Measurement and Corporate Social Responsibility
184. **Ahlerup, Pelle** (2009), Essays on Conflict, Institutions, and Ethnic Diversity
185. **Quiroga, Miguel** (2009), Microeconomic Policy for Development: Essays on Trade and Environment, Poverty and Education
186. **Zerfu, Daniel** (2010), Essays on Institutions and Economic Outcomes
187. **Wollbrant, Conny** (2010), Self-Control and Altruism
188. **Villegas Palacio, Clara** (2010), Formal and Informal Regulations: Enforcement and Compliance
189. **Maican, Florin** (2010), Essays in Industry Dynamics on Imperfectly Competitive Markets
190. **Jakobsson, Niklas** (2010), Laws, Attitudes and Public Policy
191. **Manescu, Cristiana** (2010), Economic Implications of Corporate Social Responsibility and Responsible Investments
192. **He, Haoran** (2010), Environmental and Behavioral Economics – Applications to China
193. **Andersson, Fredrik W.** (2011), Essays on Social Comparison
194. **Isaksson, Ann-Sofie** (2011), Essays on Institutions, Inequality and Development
195. **Pham, Khanh Nam** (2011), Prosocial Behavior, Social Interaction and Development: Experimental Evidence from Vietnam
196. **Lindskog, Annika** (2011), Essays on Economic Behaviour: HIV/AIDS, Schooling, and Inequality
197. **Kotsadam, Andreas** (2011), Gender, Work, and Attitudes
198. **Alem, Yonas** (2011), Essays on Shocks, Welfare, and Poverty Dynamics: Microeconometric Evidence from Ethiopia
199. **Köksal-Ayhan, Miyase Yesim** (2011), Parallel Trade, Reference Pricing and Competition in the Pharmaceutical Market: Theory and Evidence
200. **Vondolia, Godwin Kofi** (2011), Essays on Natural Resource Economics
201. **Widerberg, Anna** (2011), Essays on Energy and Climate Policy – Green Certificates, Emissions Trading and Electricity Prices
202. **Siba, Eyerusalem** (2011), Essays on Industrial Development and Political Economy of Africa
203. **Orth, Matilda** (2012), Entry, Competition and Productivity in Retail
204. **Nerman, Måns** (2012), Essays on Development: Household Income, Education, and Female Participation and Representation
205. **Wicks, Rick** (2012), The Place of Conventional Economics in a World with Communities and Social Goods
206. **Sato, Yoshihiro** (2012), Dynamic Investment Models, Employment Generation and Productivity – Evidence from Swedish Data
207. **Valsecchi, Michele** (2012), Essays in Political Economy of Development
208. **Teklewold Belayneh, Hailemariam** (2012), Essays on the Economics of Sustainable Agricultural Technologies in Ethiopia
209. **Wagura Ndiritu, Simon** (2013), Essays on Gender Issues, Food Security, and Technology Adoption in East Africa
210. **Ruist, Joakim** (2013), Immigration, Work, and Welfare
211. **Nordén, Anna** (2013), Essays on Behavioral Economics and Policies for Provision of Ecosystem Services
212. **Yang, Xiaojun** (2013), Household Decision Making, Time Preferences, and

Positional Concern: Experimental Evidence from Rural China

213. **Bonilla Londoño, Jorge Alexander** (2013), Essays on the Economics of Air Quality Control
214. **Mohlin, Kristina** (2013), Essays on Environmental Taxation and Climate Policy
215. **Medhin, Haileselassie** (2013), The Poor and Their Neighbors: Essays on Behavioral and Experimental Economics
216. **Andersson, Lisa** (2013), Essays on Development and Experimental Economics: Migration, Discrimination and Positional Concerns
217. **Weng, Qian** (2014), Essays on Team Cooperation and Firm Performance
218. **Zhang, Xiao-Bing** (2015), Cooperation and Paradoxes in Climate Economics
219. **Jaime Torres, Monica Marcela** (2015) Essays on Behavioral Economics and Policy Design
220. **Bejenariu, Simona** (2015) Determinants of Health Capital at Birth: Evidence from Policy Interventions
221. **Nguyen, Van Diem** (2015) Essays on Takeovers and Executive Compensation
222. **Tolonen, Anja** (2015) Mining Booms in Africa and Local Welfare Effects: Labor Markets, Women's Empowerment and Criminality
223. **Hassen, Sied** (2015) On the Adoption and Dis-adoption of Household Energy and Farm Technologies
224. **Moursli, Mohamed-Reda** (2015) Corporate Governance and the Design of Board of Directors
225. **Borcan, Oana** (2015) Economic Determinants and Consequences of Political Institutions
226. **Ruhinduka, Remidius Denis** (2015) Essays on Field Experiments and Impact Evaluation
227. **Persson, Emil** (2016) Essays on Behavioral and Experimental Economics, Cooperation, Emotions and Health.
228. **Martinangeli, Andrea** (2017) Bitter divisions: inequality, identity and cooperation
229. **Björk, Lisa** (2017) Essays on Behavioral Economics and Fisheries: Coordination and Cooperation
230. **Chegere, Martin Julius** (2017) Post-Harvest Losses, Intimate Partner Violence and Food Security in Tanzania
231. **Villalobos-Fiatt, Laura** (2017) Essays on forest conservation policies, weather and school attendance
232. **Yi, Yuanyuan** (2017) Incentives and Forest Reform: Evidence from China
233. **Kurz, Verena** (2017) Essays on behavioral economics: Nudges, food consumption and procedural fairness
234. **Yashodha, Yashodha** (2017) Contract Choice and Trust in Informal Groundwater Markets
235. **Meles, Tensay Hadus** (2017) Power Outages, Increasing Block Tariffs and Billing Knowledge
236. **Mühlrad, Hanna** (2018) The Impact of Reproductive and Birth Technologies on Health, Fertility and Labor Outcomes
237. **Gakii Gatua, Josephine** (2018) Primary Health Care Interventions and Social Ties in Kenya
238. **Felgendreher, Simon** (2018) Essays on behavioral economics and the effects of the colonial rule on Java
239. **Demeke, Eyoual** (2019) Essays on Environmental and Behavioral Economics
240. **Kiss, Tamas** (2019) Predictability in Equity Markets: Estimation and Inference
241. **Khomenko, Maksym** (2019) Essays on the Design of Public Policies and .

Regulations
242. **Mukanjari, Samson** (2019) Climate Policy and Financial Markets
243. **Schürz, Simon** (2020) Economic and Intergenerational Decision-Making in Families
244. **Lau, Debbie** (2020) Empirical Essays on Education and Health Policy Evaluation
245. **Ho, Hoang-Anh** (2020) Essays on Culture, Institutions, and Economic Development
246. **Rubio Ramos, Melissa** (2020) The Economics of Coercive Institutions, Conflict, and Development
247. **Lindahl, Anna** (2021) Empirical tests of exchange rate and stock return models
248. **Sjöholm, Carolin** (2021) Essays on public policy in the informal sector context
249. **Tian, Ruijie** (2021) Impacts of Climate Policy and Natural Disasters: Evidence from China
250. **Larsson, Sebastian** (2021) Providers and Profiteers: Essays on Profits and Competition in the Provision of Public Services
251. **Norrgren, Lisa** (2022) Carpe Diem or Seize your Health? The Economics of Time Preferences, Health, and Education
252. **Andersson, Lina** (2022) Emotions in Game Theory: Fear, friendliness and hostility
253. **Enlund, Jakob** (2022) Three essays in applied economics
254. **Berggren, Andrea** (2023) Choices among Doctors, Students and Primary Care Providers: Empirical Evidence from Sweden
255. **Hansson, Magnus** (2023) Decentralized Finance and Central Bank Communication
256. **Bilén, David** (2023) Sustainable Consumption and Prosocial Actions
257. **Inkinen, Ville** (2023) Wetland Mitigation Banking in the United States
258. **Tesemma, Tewodros** (2023) Essays on Environmental and Behavioral Economics
259. **Behler, Timm** (2024)Telling Talent:Essays on Discrimination and Promotion Contests