NBER WORKING PAPER SERIES

STATISTICAL DISCRIMINATION AND THE DISTRIBUTION OF WAGES

Prashant Bharadwaj
Rahul Deb
Ludovic Renou

Working Paper 32562
http://www.nber.org/papers/w32562

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2024

This paper builds on and subsumes a previous version titled “Which wage distributions are consistent with statistical discrimination?” by the second and third authors. We are very grateful to Paula Onuchic for engaging in a conversation with us that inspired this paper and to her, Aislinn Bohren, Michael Grubb, Peter Hull, Alex Imas, Namrata Kala, Georgy Kalashnov, Rob McMillan, and Konstantinos Tatsiramos for detailed discussions. We would also like to thank Torsha Chakravorty, Juan Eberhard, Kory Kroft, Aditya Kuvalekar, Marco Manacorda, Mallesh Pai, Debraj Ray, Maher Said and numerous seminar participants for insightful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by Prashant Bharadwaj, Rahul Deb, and Ludovic Renou. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Statistical Discrimination and the Distribution of Wages
Prashant Bharadwaj, Rahul Deb, and Ludovic Renou
NBER Working Paper No. 32562
June 2024
JEL No. J01,J7

ABSTRACT

We characterize the conditions under which the wage distributions for two groups are consistent with a general model of statistical discrimination. We adapt this theoretical characterization to develop a novel empirical test, the rejection of which we interpret as evidence of taste-based discrimination. In doing so, we provide a theoretical foundation via which the wage structure effect in the decomposition of wage distributions can be interpreted as evidence of taste-based discrimination. We provide a proof of concept application using Census and NLSY-79 data, which suggests taste-based discrimination at work against Black male workers in several broad occupation categories.

Prashant Bharadwaj
Department of Economics
University of California, San Diego
9500 Gilman Drive #0508
La Jolla, CA 92093
and NBER
prbharadwaj@ucsd.edu

Rahul Deb
Department of Economics
University of Toronto
150 St George St
Toronto, ON
Canada
rahul.deb@utoronto.ca

Ludovic Renou
School of Economics and Finance
Queen Mary University of London
London E1 4NS
United Kingdom
l.renou@qmul.ac.uk

1. INTRODUCTION

While it is well established that economic outcomes for observationally identical individuals can differ based on their group identity, it is significantly harder to determine the reason for these disparate outcomes. Discrimination is one possible explanation. At a broad level, economists characterize discrimination as either *statistical* (outcomes differ because of differences in information) or *taste-based* (bias or animus towards one group drives outcome differences). While both forms of discrimination are problematic, taste-based discrimination is particularly pernicious because, unlike statistical discrimination, it is unresponsive to information. Establishing the form of discrimination is important both for accountability and to devise corrective policies.

Consider, for instance, the differences in the wage distributions for Black and White workers of identical age and education who work the same jobs in the same location. Suppose that, despite being observationally identical, White workers have higher average wages. Is such a wage gap the result of discrimination? Not necessarily. Even if Black and White workers have the same *average* productivities, their productivity *distributions* need not be identical. Thus, if wages were a nonlinear function of the workers' productivities, wage gaps can arise even without any discrimination. Suppose instead, we assume that the entire productivity distributions of both groups were identical. Then, the wage gap must be the result of some form of discrimination but it is possible to infer the type? Clearly, taste-based discrimination can result in wage gaps. But so can statistical discrimination. If employers receive different signals for each group, the posterior distributions of perceived productivities (via these signals) for the employers can differ across groups. Once again, if wages are nonlinear functions of perceived productivities, wage gaps can arise.

In this paper, we propose a general model of statistical discrimination in the labor market and theoretically characterize the set of wage distributions that are consistent with this model. This is the set of wage distributions that can theoretically be explained by statistical discrimination *alone*. If a pair of observed wage distributions does not lie in this set then, not only can we conclude that discrimination is present but, we have also uncovered evidence of taste-based discrimination. This theoretical characterization in turn yields a nonparametric test for statistical discrimination, rejections of which we interpret as evidence of taste-based discrimination. A strength of this approach is that it can be applied to commonly available cross-sectional data (such as Census data) and it provides a framework for interpreting when the "unexplained" part of a wage distribution decomposition is evidence of discrimination in general, and taste-based discrimination in particular. Our empirical application in this paper shows evidence of taste-based discrimination against Black workers in certain occupations.

Our model is in the spirit of Phelps (1972). There are two groups whose productivity distributions differ. The group identity is observable to employers, but productivities are not. Instead, employers learn about the workers' productivity from signals whose distributions may vary across the groups. For example, these signals could be the information that employers receive from the job screening process that includes interviews, tests, curricula vitae, data-drive software etc. Signal realizations induce posterior productivity distributions (via Bayes' rule) and, in particular, these can be used to compute posterior estimates (the mean of the productivity conditional on the signal realization) of the unobserved productivity. Therefore, each group's signal generates a distribution over posterior productivity estimates. Wages are then determined via a strictly increasing, continuous function of the posterior productivity estimate that, importantly, does not depend on the group. The combination of assumptions make our model more general than others in the literature: we do not require the productivity distributions or statistical experiments to be Gaussian and we allow for nonlinear wage functions to capture imperfectly competitive labor markets. Our theory aims to characterize the pairs of wage distributions that are rationalizable by this model under different assumptions about the set of permissible productivity distributions.

We first consider the baseline case where we assume that both groups have equal mean productivities but, apart from this, the distributions can differ arbitrarily. We show that a necessary and sufficient condition for a pair of wage distributions to be rationalizable under this assumption of equal mean productivities is that neither wage distribution strictly first-order stochastically dominates the other. Thus, if wages are ordered by strict first-order stochastic dominance, then they cannot be explained by statistical discrimination alone. Importantly, this key insight applies to other, non-labor market contexts. It might be particularly useful for audit and correspondence studies (as long as the outcome variable of interest is non-binary) where it is assumed that both groups are identical on average but the type of discrimination is nonetheless typically hard to pin down.[1] Our result says that researchers in this space can uncover taste-based discrimination by simply testing for strict first-order stochastic dominance.

Under non-experimental contexts, holding mean productivity constant is a bigger challenge. We argue that an immediate consequence of our baseline result is a characterization of rationalizable wage distributions assuming that one group has a weakly *higher* mean productivity: the wage distribution for the group with lower mean productivity should not be strictly first-order stochastically dominant (we call this the "ordered means" case). In the empirical application of this paper, we take this result to the data while using education as a proxy for productivity (controlling for other covariates). If we observe that the wage distribution of the group with lower mean productivity (less education) strictly first-order stochastically dominates the wage distribution of the group with higher mean productivity (more education), we interpret this as evidence of taste-based discrimination. We emphasize that, while it is well known that simply observing mean wage differentials is not enough to conclude the *type* of discrimination, we show that we can learn something about the type of discrimination by comparing the entire distributions of

[1]As Bertrand and Duflo (2017) observe: "while field experiments have been overall successful at documenting that discrimination exists, they have (with a few exceptions) struggled with linking the patterns of discrimination to a specific theory." An excellent example of applying our test in audit study settings with continuous outcome variables is Ayalew, Manian, and Sheth (2023) who examine bias in scoring entrepreneurship competitions.

wages.

Our empirical application employs recent econometric advancements in distributional decompositions. Specifically, we use the method of Chernozhukov, Fernández-Val, and Melly (2013) to test whether the wage distribution of Black workers is strictly first-order stochastically dominated by the counterfactual wage distribution generated by assuming that the wages for Black workers are determined by the same wage setting process as those of White workers (in language familiar to labor economists, this is the "wage structure" effect). In words, this compares the observed wage distribution for Black workers with what they would receive if they were treated as White; we conclude taste-based discrimination is present if these distributions are ordered by strict first-order stochastic dominance. We test the ordered means case in two distinct ways using two different data sets. We first implement wage distribution decompositions on publicly available Census data by creating samples where Black workers have *more* years of education than White workers. As mentioned above, the assumption here is that more educated Black workers are more productive on average. We then use NLSY-79 panel data that allows us to decompose wage distributions while conditioning on past wages and occupations. Here, the ordered means assumption is satisfied if there was discrimination in determining last period wages. In other words, ordered means is only violated if, in the previous period, Black workers were paid more than White workers of the same expected productivity (we view such "reverse discrimination" to be unlikely). We find evidence that the wage structure effect exhibits strict first-order stochastic dominance in several occupations (although not always).

One of the main contributions of our paper is that it provides a theoretical lens to interpret the decompositions of wages. Wage decompositions in labor economics have a rich history, starting with the seminal work of Kitagawa (1955), Oaxaca (1973) and Blinder (1973) who developed the framework to understand whether differences in outcomes were the result of differing characteristics or differential returns to characteristics across groups. The "unexplained portion" of the Kitagawa–Oaxaca–Blinder decomposition has long been a North Star for labor economists aiming to quantify the *amount* of discrimination.[2] This strand of the literature has largely evolved in parallel to the work that aims to determine the *type* of discrimination.[3,4] Our novel theory combined with empirical advances made possible by DiNardo,

[2]Early influential papers that use and build on this approach include Juhn, Murphy, and Pierce (1993) and Altonji and Blank (1999). The recent work of Bohren, Hull, and Imas (2022) provides more nuance on how to interpret the unexplained portion and introduces the ideas of "direct" and "systemic" discrimination.

[3]The classic way to test for taste-based discrimination is to use what are known as "outcome tests" (in the spirit of Becker, 1957, 1993). These tests require the researcher to have access to not just to the decision (whether or not a loan is granted, a driver is searched by a police officer, etc) but also the post-decision result (whether or not the loan is repaid, contraband is found on the driver, etc). This requires devising empirical strategies to identify the post-decision results of marginal cases or models that provide a systematic relationship between the average and marginal post-decision result. See, for instance, Knowles, Persico, and Todd (2001), Anwar and Fang (2006), Arnold, Dobbie, and Yang (2018) and Canay, Mogstad, and Mountjoy (2023).

[4]Tests for statistical discrimination are typically based on how decision makers update their behavior in response to information (a classic example is Altonji and Pierret, 2001). Recently, Bohren, Imas, and Rosenberg (2019) conduct an experiment

Fortin, and Lemieux (1996) and Chernozhukov, Fernández-Val, and Melly (2013) help these literatures speak to each other by providing a way of interpreting the unexplained portion of the wage decomposition through the lens of the two dominant models of discrimination in labor economics. Our theoretical insight is that the comparison of wage distributions is informative but that of mean wages is less so.

The penultimate section of the paper extends the theory in several directions. Notably, we examine the case where we do not assume average productivities are ordered. We derive a tight lower bound on the average productivity differences required to rationalize a given pair of wage distributions with statistical discrimination alone. We discuss how this bound can be used to uncover evidence for taste-based discrimination. We then argue that deriving a similar bound for percentage differences in average productivities is not possible unless we make further assumptions about the relationship between productivities and wages.

2. THE MODEL

This section presents our model of statistical discrimination that can be thought of as a non-parametric generalization of the model of Phelps (1972).

There are two groups—1 and 2—of workers; examples include female and male, Black and White, junior and senior, or disabled and able bodied. In the theoretical results, we do not take a stand on which of these two groups is advantaged/disadvantaged, if any.

We observe two *wage distributions* G_1 and G_2, with $G_i(w) \in [0,1]$ being the fraction of workers in group $i \in \{1,2\}$ who are paid a wage of $w \geq 0$ or less.[5] We assume that the wage distributions are bounded, that is, $G_i(\overline{w}) = 1$ for some $\overline{w} > 0$, $i = 1,2$.

The question we address is: under what conditions are the observed wage distributions rationalized by (or consistent with) a general model of statistical discrimination? As we shall see, the answer provides a test of *taste-based* discrimination, in that, whenever the wage distributions are *not* rationalizable, statistical discrimination *alone* cannot explain the data, but taste-based discrimination can. In statistical terms, our null hypothesis is that the data is consistent with statistical discrimination alone (which includes the case of no discrimination), and we are interested in rejecting the null hypothesis. The answer to the above question underpins our statistical test. We interpret a rejection of the null hypothesis as evidence of taste-based discrimination and we demonstrate the validity of this interpretation. We now present the aforementioned model in detail, starting with the productivity distributions.

in which they can control the precision and dynamics of information. This allows them to differentiate between whether disparate outcomes are the result of (correct or incorrect) beliefs or preferences of the decision makers.

[5]Throughout, all distributions are right-continuous and have limits on the left.

Productivity distributions: Workers differ in their (true) productivities, with $\theta_i \in \mathbb{R}_+$ denoting the *productivity* of a worker in group $i \in \{1, 2\}$, and H_i its (cumulative) distribution. The productivity refers to the marginal product of a worker and is thus measured in the same unit as wages (US dollars in our empirical application).

We hypothesize a set $\widehat{\mathcal{H}}$ of pairs of productivity distributions (H_1, H_2) and derive the testable implications of our model assuming that the productivity distributions (H_1, H_2) belong to $\widehat{\mathcal{H}}$. Throughout, we assume that every pair of productivity distributions $(H_1, H_2) \in \widehat{\mathcal{H}}$ are supported on a subset of some bounded interval $[0, \overline{\theta}] =: \Theta$. Different assumptions imposed on the set $\widehat{\mathcal{H}}$ may have different testable implications. As an example, in our baseline model below, we assume that $\widehat{\mathcal{H}}$ is the set of all pairs of productivity distributions that have equal means but can otherwise differ. As another example, we assume that $\widehat{\mathcal{H}}$ contains all pairs of distributions, whose means differ by at most $d > 0$. The validity of the hypothesized set $\widehat{\mathcal{H}}$ must be argued, either empirically or theoretically.

Information: Employers do not directly observe the productivities of workers, but receive informative signals (from CVs, reference letters, interviews, tests, data-driven software etc.). Employers then form an expectation of the productivity of workers and pay them accordingly: wages are strictly increasing in expected productivity. Since wages only depend on the expected productivity, we assume that the signals employers receive are posterior estimates of the productivity. This is without loss of generality.

More precisely, a *signal* (S_i, π_i) for group $i \in \{1, 2\}$ consists of a set of *signal realizations* $S_i = \Theta$ and a joint distribution π_i over $\Theta \times S_i$, whose marginal distribution over Θ is the (prior) productivity distribution H_i. We denote the marginal distribution of π_i over S_i by F_i. We assume that the *posterior estimate* $\mathbb{E}_{\pi_i}[\theta_i \mid s_i]$ of the productivity satisfies

$$s_i = \mathbb{E}_{\pi_i}[\theta_i \mid s_i],$$

for all s_i in the support of F_i. In words, the signal realization s_i is an accurate estimate of the true productivity θ_i. As mentioned, this is without loss of generality, as we can always relabel signals to guarantee that they are accurate in the above sense. In what follows, we slightly abuse notation and write θ_i for the posterior estimate (the signal realization).

It is well known that F_i is a distribution of posterior estimates arising from *some* signal if, and only if, the prior distribution H_i is a *mean-preserving spread* of the posterior distribution F_i, which we denote by $F_i \succcurlyeq_2 H_i$ (where the notation reflects second-order stochastic dominance). Formally, the mean-preserving spread condition requires that

$$\int_0^\theta H_i(\theta_i) d\theta_i \geq \int_0^\theta F_i(\theta_i) d\theta_i \text{ for all } \theta \in [0, \overline{\theta}], \text{ with equality at } \theta = \overline{\theta}.$$

Note that the requirement of equality at $\theta_i = \overline{\theta}$ is the same as ensuring that H_i and F_i have the same mean.[6]

We stress that the above formulation subsumes *all* possible signaling technologies. In particular, this includes the common formulation (as in Phelps, 1972; Aigner and Cain, 1977) of modeling signals as $s_i = \theta_i + \varepsilon_i$, where ε_i is a noise term whose distribution (typically assumed to be normal) can depend on the group i and possibly the productivity θ_i as well.

Moreover, it is worth highlighting that we do not take a stand on the source of these signals. In particular, we can allow for employers to design different screening practices for each group. We simply require employers to be Bayesian with accurate beliefs (but we discuss the possibility of inaccurate beliefs and non-Bayesian updating below).

Wage function: If an employer estimates the productivity of a worker to be θ, they pay the worker $W(\theta)$, where the *wage function* $W : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous and strictly increasing.[7] Let $\mathcal{W}$ be the set of continuous and strictly increasing functions with both domain and range being $\mathbb{R}_+$. As with the set of productivity distributions, we hypothesize a set of wage functions $\widehat{\mathcal{W}}$. As an example, employers may pay workers a fixed fraction of their expected productivities, in which case $\widehat{\mathcal{W}}$ is the set of linear functions.

It is worth making a few comments about the wage functions. First, the wage functions do not depend on the group identity. It is in this sense that our model captures statistical, but not taste-based, discrimination. Second, the labor economics literature frequently assumes perfectly competitive labor markets. In our notation, this amounts to assuming $\widehat{\mathcal{W}} = \{W_{id}\}$ where $W_{id}(\theta) = \theta$. As will become clear from our results below, a strength of our framework is that we require no such restrictive assumptions and we can accommodate imperfectly competitive labor markets (via the reduced form wage function) quite generally.

Induced wage distributions: The distribution F_i over posterior estimates *induces the wage distribution* G_i via the wage function W. Formally, for both $i \in \{1, 2\}$, $G_i(w)$ is the measure of the set $\{\theta : W(\theta) \leq w\}$ according to F_i, that is, $G_i(w) = F_i(W^{-1}(w))$ for $w \in [W(0), W(\overline{\theta})]$, $G_i(w) = 0$ for $w < W(0)$ and $G_i(w) = 1$ for $w > W(\overline{\theta})$.[8] Note that, even though the wage function does not depend on group identity, the wage distributions G_1 and G_2 may differ across groups for the simple reason that the distributions of posterior estimates F_1 and F_2 may differ.

Rationalizability: We say that the observed wage distributions G_1 and G_2 are *rationalizable* (given $\widehat{\mathcal{H}}$, $\widehat{\mathcal{W}}$) if there exist (i) productivity distributions $(H_1, H_2) \in \widehat{\mathcal{H}}$, (ii) distributions of posterior estimates F_i

[6]Integration by parts implies that the mean satisfies $\int_0^{\overline{\theta}} \theta_i dF_i(\theta_i) = \theta_i F_i(\theta_i)|_0^{\overline{\theta}} - \int_0^{\overline{\theta}} F_i(\theta_i)d\theta_i = \overline{\theta} - \int_0^{\overline{\theta}} F_i(\theta_i)d\theta_i$.

[7]We could additionally assume that $W(\theta) \leq \theta$ (workers are paid less than their marginal product) and no result in the paper changes. While this is a natural assumption, we do not require it so choose not to impose it.

[8]We define W^{-1} as the inverse of W on the domain $[W(0), W(\overline{\theta})]$.

that satisfy $F_i \succcurlyeq_2 H_i$ for $i \in \{1,2\}$, and (iii) a wage function $W \in \widehat{\mathcal{W}}$, such that these jointly induce the observed wage distributions.

Before moving on to the analysis, we discuss the model and the question it addresses. Our model is in the spirit of the seminal models of Phelps (1972) and Aigner and Cain (1977). Phelps considers two populations whose productivities are drawn from a normal distribution. Signals are also normally distributed, differ across groups, and the wage function is linear in the posterior estimate. If the means of the productivity distributions for both groups are the same, then the Phelps' model implies that the average wage for both groups is the same (because the posterior distribution must have the same mean as the prior, and the wage function is linear). In this case, there is no discrimination at the group level even though the wage distributions differ (so there is individual level discrimination). Aigner and Cain (1977) observe it is possible to generate discrimination at the group level via more general wage functions even when the productivity distributions for both groups are identical. In their model, wages depend both on the mean and the variance of the posterior belief. In the normal learning environment, the posterior variance is the same for all signal realizations so they model the wage as just the difference between the posterior mean and some multiple of the (signal independent) variance of the posterior belief. Hence, different normally distributed signals can generate distinct mean wages.

Our model is more general than these seminal papers (and most of the literature) in that we do not assume that the productivity distributions are Gaussian and precisely known by the analyst (they instead lie in set $\widehat{\mathcal{H}}$), and we allow for unrestricted (not necessarily Gaussian) signals. Moreover, we do not restrict wages to be linear in the posterior estimate.[9] That said, unlike Aigner and Cain (1977), wages in our model only depend on the mean of the posterior distribution, but not on the variance. This choice is deliberate: our model is very general and our assumptions balance this generality against meaningful testable implications that can be taken to the data. As we discuss in Section 6, allowing wages to depend on both the mean and the variance makes the testable implications of our model vacuous, even when relatively restrictive assumptions are imposed on $\widehat{\mathcal{H}}$ (equal mean productivities) and $\widehat{\mathcal{W}}$ (wages that are affine in the posterior mean and variance).

Similar to Phelps (1972) and Aigner and Cain (1977), we do not model the underlying reason that productivities differ across groups. In this sense, we differ from papers like Coate and Loury (1993) whose primary purpose is to explain how stereotypes (that assume the disadvantaged group has lower mean productivity) can be self-fulfilling. Note that we can accommodate self-fulfilling stereotypes against group 1 in our model by hypothesizing the set $\widehat{\mathcal{H}} = \{(H_1, H_2) \,|\, \mathbb{E}_{H_1}[\theta_1] \leq \mathbb{E}_{H_2}[\theta_2]\}$.

[9]There is one recent paper, Chambers and Echenique (2021), that also uses an information design approach but their model and question are very different from ours. In their model, workers have skills, firms learn about these skills via signals, optimally match workers to tasks and pay the worker the value they generate. They characterize properties of the information structure such that every signal leads to the same expected wages. They also establish a relation with the production technology.

Our main theoretical aim is to fully characterize the set of rationalizable pairs of wage distributions under different assumptions on $\widehat{\mathcal{H}}$ and $\widehat{\mathcal{W}}$. This separates our analysis from most theoretical papers on discrimination that aim to derive the main implications of their models, but do not provide a complete characterization of the testable implications. In this sense, our analysis is closer to the theoretical literature on decision theory and revealed preference.

3. CHARACTERIZING RATIONALIZABLE WAGE DISTRIBUTIONS

In this section, we derive the main theoretical results. We begin with a baseline result that assumes both groups have equal mean productivities and establish its economic implications. We then generalize to a setting where we make the weaker assumption that one group has weakly higher mean productivity than the other: this generalization is the basis of our empirical test.

3.1. BASELINE RESULT: EQUAL MEAN PRODUCTIVITIES

As we have emphasized earlier, the key novelty of our framework is that we allow for (arbitrary) non-Gaussian priors, signals and nonlinear wages. In this subsection, we isolate the implications of these modeling features by assuming that both groups have equal mean productivities but, apart from this, we impose no further restrictions. In other words, we characterize how the wage distributions can differ due to a combination of statistical discrimination and nonlinear wages for two groups that are identical on average. We weaken this assumption in subsequent sections.

We define

$$\mathcal{H}_= := \{(H_1, H_2) \,|\, \mathbb{E}_{H_1}[\theta_1] = \mathbb{E}_{H_2}[\theta_2]\}$$

to be the set of productivity distribution pairs that have equal means. Note that, barring the equal means requirement, the productivity distributions can differ arbitrarily.[10] Formally, we characterize the set of rationalizable wage distributions given that the productivity distributions lie in the set $\mathcal{H}_=$ but the wage function is unrestricted (that is, we take $\widehat{\mathcal{W}} = \mathcal{W}$).

Despite the equal means assumption, the model is very general as the signals and the wage function are unrestricted. Given this generality, the first natural question to ask is: is there *any* pair of wage distributions that is *not* rationalizable (given $\mathcal{H}_=$, $\mathcal{W}$)? To this point, note that our model allows the posterior estimate distribution of group 1 to be a strict mean-preserving spread of group 2, in which case a strictly convex wage function W will generate higher mean wages for group 1.[11] In other words, differences in mean wages (a wage gap) can arise purely via statistical discrimination, even though both groups have equal mean productivities. So, to find inconsistent distributions, we need to consider higher moments

[10] Distributions can be discrete, continuous or a mixture of the two.

[11] The mean-preserving spread is strict when $F_2 \succcurlyeq_2 F_1$ and $F_2 \neq F_1$.

of the wage distribution. In fact, as we now argue, we need to consider *all* moments via the following order.

The wage distribution G_i *strictly first-order stochastically dominates* the wage distribution G_j, which we denote $G_i \succ_1 G_j$, if $G_i(w) \leq G_j(w)$ for all $w \in \mathbb{R}_+$, with the inequality strict for some w.

Now, suppose that the wage distribution of group i strictly first-order stochastically dominates that of group j. We argue that these distributions are *not* rationalizable. By contradiction, assume that these distributions are rationalizable (given $\mathcal{H}_=$, $\mathcal{W}$). This implies that there exist posterior estimate distributions F_i and F_j, and a wage function W, such that

$$F_i(\theta) = G_i(W(\theta)) \leq G_j(W(\theta)) = F_j(\theta) \qquad \text{for all } \theta \in [0, \overline{\theta}],$$

with the inequality strict for some θ, that is, $F_i \succ_1 F_j$. It follows that F_i has a strictly higher mean than F_j, which is a contradiction since F_i and F_j are *mean-preserving* contractions of some productivity distributions $(H_i, H_j) \in \mathcal{H}_=$, which have the same mean.

The above argument shows that a necessary condition for a pair of wage distributions to be rationalizable (given $\mathcal{H}_=$, $\mathcal{W}$) is that neither strictly first-order stochastically dominates the other. Our first result shows that this condition is also sufficient.

THEOREM 1. *Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_=$, $\mathcal{W}$) if, and only if, neither G_1 nor G_2 strictly first-order stochastically dominates the other.*

The remainder of this subsection discusses Theorem 1. We first observe that, in terms of the testable implications, our model has in-build redundant generality. This redundancy has economic implications. We use

$$\widetilde{\mathcal{H}}_\equiv := \{(H_1, H_2) \in \mathcal{H}_= \mid (H_1, H_2) \text{ are both supported on two points } \{0, \tilde{\theta}\} \text{ with } \tilde{\theta} \in (0, \infty)\}$$

to denote the subset of $\mathcal{H}_=$ containing the pairs of distributions with the same binary support. Note that $(H_1, H_2) \in \widetilde{\mathcal{H}}_\equiv$ implies that $H_1 = H_2$, since their means and binary supports are the same.

Take any distribution H_i, and let $\widetilde{H}_i$ be the (discrete) distribution supported on $\{0, \tilde{\theta}\}$ such that the support of H_i is a subset of $[0, \tilde{\theta}]$ and both distributions have equal means: $\mathbb{E}_{H_i}[\theta_i] = \mathbb{E}_{\tilde{H}_i}[\theta_i]$. Now, observe that every mean-preserving contraction F_i of H_i is also a mean-preserving contraction of $\widetilde{H}_i$. This is because $\widetilde{H}_i$ is the distribution that is the "most spread" (in that all the mass is at both end points of the interval 0 and $\tilde{\theta}$) amongst all distributions with mean $\mathbb{E}_{H_i}[\theta_i]$ that are supported on a subset of

$[0, \tilde{\theta}]$. Consequently,

$$\{(F_1, F_2) \mid F_1 \succcurlyeq_2 H_1,\ F_2 \succcurlyeq_2 H_2 \text{ and } (H_1, H_2) \in \widetilde{\mathcal{H}}_{\equiv}\} = \{(F_1, F_2) \mid F_1 \succcurlyeq_2 H_1,\ F_2 \succcurlyeq_2 H_2 \text{ and } (H_1, H_2) \in \mathcal{H}_{=}\}.$$

Denoting $\mathcal{H}_{\equiv} := \{(H_1, H_2) \mid H_1 = H_2\}$ to be the set of all pairs of identical productivity distributions (with not necessarily binary support), the above equality then implies that

$$\{(F_1, F_2) \mid F_1 \succcurlyeq_2 H_1,\ F_2 \succcurlyeq_2 H_2 \text{ and } (H_1, H_2) \in \mathcal{H}_{\equiv}\} = \{(F_1, F_2) \mid F_1 \succcurlyeq_2 H_1,\ F_2 \succcurlyeq_2 H_2 \text{ and } (H_1, H_2) \in \mathcal{H}_{=}\},$$

because $\widetilde{\mathcal{H}}_{\equiv} \subset \mathcal{H}_{\equiv} \subset \mathcal{H}_{=}$. This shows is that the set of rationalizable wage distributions given $(\mathcal{H}_{=}, \mathcal{W})$ is the same as the set of rationalizable wage distributions given $(\mathcal{H}_{\equiv}), \mathcal{W})$. In words, the testable implications of our model are the same whether we assume equal mean productivities or *identical productivity distributions*.

THEOREM 1 (CONTINUED). *The following statements are equivalent.*

(i) Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_{=}$, $\mathcal{W}$).

(ii) Neither G_1 nor G_2 strictly first-order stochastically dominates the other.

(iii) Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_{\equiv}$, $\mathcal{W}$).

It has been argued that, for the distributions of certain traits, men and women have the same mean, but the former have a higher variance. This is sometimes referred to as the "variability hypothesis" and is used to explain differential outcomes. The third statement of Theorem 1 implies that any two wage distributions that are not ordered by strict first-order stochastic dominance, no matter how different, could have resulted from statistical discrimination on *identical* populations. In other words, allowing for different variances of the productivity distributions does not lead to more permissive testable implications. That is, in our setting, the variability hypothesis has no added explanatory power!

The redundant generality in our model can be rephrased in an additional way. Observe that, given any pair of productivity distributions $(H_1, H_2) \in \mathcal{H}_{=}$, every pair of mean-preserving contractions, $F_1 \succcurlyeq_2 H_1$ and $F_2 \succcurlyeq_2 H_2$ also belong to the set $\mathcal{H}_{=}$, that is $(F_1, F_2) \in \mathcal{H}_{=}$. This is simply because F_1 and F_2 have the same means and $\mathcal{H}_{=}$ contains every pair of distributions whose means are equal. Consequently,

$$\mathcal{H}_{=} = \{(F_1, F_2) \mid F_1 \succcurlyeq_2 H_1,\ F_2 \succcurlyeq_2 H_2 \text{ and } (H_1, H_2) \in \mathcal{H}_{=}\}.$$

This equality of sets has an important economic implication. It says that we cannot distinguish a model where employers have the prior beliefs (F_1, F_2) and *perfectly observe* the productivity of workers from a model where employers have the prior beliefs (H_1, H_2) and form the posterior beliefs (F_1, F_2) via informative signals.

We say that two wage distributions G_1 and G_2 are *rationalizable without signal discrimination* (given $\widehat{\mathcal{H}}$, $\widehat{\mathcal{W}}$) if there exist (i) productivity distributions $(H_1, H_2) \in \widehat{\mathcal{H}}$, (ii) perfectly informative signals ($F_1 = H_1$, $F_2 = H_2$) and (iii) a wage function $W \in \widehat{\mathcal{W}}$, such that these jointly induce the observed wage distributions.[12] The above argument implies the following result.

THEOREM 1 (CONTINUED). *The following statements are equivalent.*

(i) Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_=$, $\mathcal{W}$).

(ii) Neither G_1 nor G_2 strictly first-order stochastically dominates the other.

(iii) Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_\equiv$, $\mathcal{W}$).

(iv) Wage distributions G_1 and G_2 are rationalizable without signal discrimination (given $\mathcal{H}_=$, $\mathcal{W}$).

Statement (iv) says that, when condition (ii) holds, we *cannot conclude that discrimination in either signal informativeness or wage payments is present*. In other words, when wage distributions are not ordered by strict first-order stochastic dominance, it is possible that the wage differences arise simply from heterogeneous populations (with identical mean productivities) whose wages are determined in a group-neutral way by employers who perfectly observe productivities. Thus Theorem 1 implies that, when condition (ii) is rejected in the data, we can not only conclude that discrimination is present, we can also conclude that discrimination cannot be statistical alone! It is important to note that this claim only applies to the wages. We naturally cannot exclude the possibility that the heterogenous populations are themselves the product of past statistical or taste-based discrimination.

A version of this insight also holds in settings where, unlike wages, outcomes are binary. An example is the setting of correspondence studies in which researchers send fictitious CVs to employers and record the rates at which employers do or do not call back for interviews (the binary outcome). Differential interview callback rates are typically interpreted as evidence of discrimination (an additional statistical versus taste-based conclusion is usually not made). But consider the following situation. Suppose that the productivity of a worker is firm-specific and each firm can correctly assess the productivity of a worker with a given CV and group identity. Identical CVs from both groups may, however, correspond to different

[12]Note that it would be equally natural to assume that employers do not perfectly observe productivities, but learn about them via a single group-independent signal. The statement of our result applies verbatim with this definition of "rationalizability without signal discrimination."

firm-specific productivities. One possible reason is that universities have differential admission policies across groups and so the same educational qualification might nonetheless imply different skills. Aggregating across firms, the productivity distribution of one group might differ from the other group, even conditional on a given CV. If firms only call back for interviews those applicants whose productivities are above a threshold, the callback rates for both groups might differ even if the aggregate productivity distributions have the same mean. In other words, differential callback rates might occur even if both groups have the same average productivity and the employers can perfectly observe productivities.

Formally, it is easy to show that, given any two callback rates that lie strictly between 0 and 1, it is possible to construct two productivity distributions that have the same mean and a single group independent cutoff such that the mass of workers from each group above the cutoff are exactly the callback rates.[13] Thus, it is theoretically possible that the differential callback rates found in many correspondence studies evidence no discrimination (in our sense) whatsoever! One way to view this insight is that it formalizes critiques made by Heckman (1998) and Neumark (2012). An alternate implication is that, for audit or correspondence studies with non-binary outcomes (like wages), testing for strict first-order stochastic dominance allows the researcher to conclude not just that discrimination is present but also that discrimination is not statistical alone.

3.2. ORDERED MEAN PRODUCTIVITIES

While assuming equal mean productivities is a natural starting point for a theoretical study on discrimination, it is a restrictive assumption if one wants to take the theory to the data, as we do. In this subsection, we weaken this assumption and show that Theorem 1 extends immediately to an environment where we assume mean productivities to be ordered. We first present the result formally and then discuss why this generalization is helpful for empirical applications.

We define

$$\mathcal{H}_{\geq} := \{(H_1, H_2) \,|\, \mathbb{E}_{H_1}[\theta_1] \geq \mathbb{E}_{H_2}[\theta_2]\}$$

to be the set of productivity distribution pairs in which the mean productivity of group 1 is weakly greater than that of group 2. Since $\mathcal{H}_{\geq} \supset \mathcal{H}_{=}$, this is a weaker assumption than the one imposed in Theorem 1. The following result follows immediately.

THEOREM 2. *The following statements are equivalent.*

(i) Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_{\geq}$, $\mathcal{W}$).

(ii) G_2 does not strictly first-order stochastically dominate G_1.

[13] A formal statement is in an older version (Deb and Renou, 2022) of this paper.

(iii) Wage distributions G_1 and G_2 are rationalizable without signal discrimination (given $\mathcal{H}_{\geq}$, $\mathcal{W}$).

The discussion in the next paragraph serves as a proof of the equivalence of statements (i) and (ii). The equivalence of statement (iii) follows from the identical argument that showed the equivalence of the analogous statement (iv) of Theorem 1. Naturally, since productivity distributions can have unequal means, there is no natural analogue of statement (iii) of Theorem 1.

Suppose $G_2 \succ_1 G_1$. If the distributions are rationalizable, then the argument preceding (the first statement of) Theorem 1 implies that, for any $W \in \mathcal{W}$, the resulting distributions $F_1(\theta_1) = G_1(W(\theta_1))$, $F_2(\theta_2) = G_2(W(\theta_2))$ of posterior estimates satisfy $\mathbb{E}_{F_1}[\theta_1] < \mathbb{E}_{F_2}[\theta_2]$. This is a contradiction since every pair (H_1, H_2) that satisfies $F_i \succcurlyeq_2 H_i$ has mean $\mathbb{E}_{F_i}[\theta_i] = \mathbb{E}_{H_i}[\theta_i]$ for $i \in \{1, 2\}$, and so $(H_1, H_2) \notin \mathcal{H}_{\geq}$. If G_1 and G_2 are not ordered by strict first-order stochastic dominance, then Theorem 1 shows that these wage distributions are rationalizable (given $\mathcal{H}_{=}$, $\mathcal{W}$) and so also rationalizable (given $\mathcal{H}_{\geq}$, $\mathcal{W}$). Lastly, if $G_1 \succ_1 G_2$, then we can rationalize the data with the (perfectly competitive) wage function $W(\theta) = \theta$ and $(F_1, F_2) = (H_1, H_2) = (G_1, G_2)$.

In the next section, we derive and implement an empirical test for discrimination based on Theorem 2. For this test to be informative, we need to ensure that the ordered mean productivity assumption $\mathcal{H}_{\geq}$ is satisfied in our empirical application. For this reason, Theorem 2 is a better candidate than Theorem 1 to take to the data because it is hard to ensure that mean productivities are exactly equal, even with fine controls. Conversely, as mentioned earlier, there are reasons to believe that, due to discrimination (for instance in acquiring education), the disadvantaged group might be positively selected for a given vector of covariates. It is also possible to stack the deck in favor of the disadvantaged group (to ensure they have higher mean productivity) by, for instance, comparing more-educated workers from the disadvantaged group with less-educated workers from the advantage group (while, of course, controlling for everything else). The ordered means assumption would only be invalid under the unlikely situation that the advantaged group has higher average productivity despite having lower education.

As mentioned earlier, the null hypothesis of our empirical application is that wage distributions are rationalizable (given $\mathcal{H}_{\geq}$, $\mathcal{W}$) and we interpret rejections of the null as evidence of *taste-based discrimination*. Naturally, if wage distributions are not rationalizable, it means that the combination of assumptions of the model are rejected by the data. One assumption of the model is that the wage function is group independent. If wages are group dependent—that is, wages for group 1, 2 are determined by potentially different functions W_1, W_2—then *every* pair of wage distributions are rationalizable with group-dependent wage functions (given either $\mathcal{H}_{\geq}$, $\mathcal{W}$ or $\mathcal{H}_{\geq}$, $\mathcal{W}$). Group dependent wages can be thought of as taste-based discrimination (following Becker, 1957) because two workers with identical expected productivities in the eyes of the employer are paid differently. Interpreting $G_2 \succ_1 G_1$ as taste-based discrimination in this sense would not always be correct if there existed pairs of wage distributions that were not rationalizable by group-dependent wage functions.

Another assumption of the model is that employers are Bayesian with accurate beliefs. Bayesian updating is not needed for our results. Any updating rule, which satisfies the property that the expectation of posterior beliefs is the prior belief, induces the same testable implications. In addition, Theorem 2 allows for some belief inaccuracy. Indeed, rationalizing the observed wages does not involve finding the *true* beliefs of employers, it merely requires that there is *one* instance of the model that is consistent with the observed wages. Thus, a rejection of condition (ii) of Theorem 2 implies that we are rejecting the possibility of *any* prior belief (and signal) that assigns higher mean productivity to the disadvantaged group. Thus, employers can have inaccurate beliefs as long as the mean productivities remain correctly ordered. We are thus also labeling, as taste-based discrimination, beliefs that are so inaccurate that they assign lower mean productivity to the disadvantaged group. Ultimately, as in common in the literature (see Bohren, Haggag, Imas, and Pope, 2023), we cannot distinguish using basic wage data whether wage differences are the result of group-dependent wage functions or very inaccurate beliefs.

It is also possible that employers deliberately use different methods to screen candidates of both groups. For instance, they may interview candidates of one group in person and the other online leading to differentially informative signals. Since we do not model the source of the signals, we will label such deliberate differential treatment as statistical discrimination. In other words, even if the wage distributions are rationalizable, it is nonetheless possible that candidates from both groups are treated differently by employers. When wage distributions are not rationalizable, we are detecting taste-based discrimination (in the above sense), over and above potential discrimination in the signal choice.

Lastly, it is possible that wage distributions are not rationalizable because employers determine wages not just on the basis of posterior means but also on the basis of higher moments of the posterior distribution. As we discuss in our concluding remarks, this is a possibility that can never be ruled out in a model this general. One could impose enough structure on priors, signals and the relationship between wages and higher moments of the posterior distribution to ensure that the model has nontrivial testable implications. But, this would not insure the analysis from the criticism that non-rationalizability is interpreted incorrectly because the true model of the world is more general than that of the econometrician.

4. EMPIRICAL APPLICATION

In this section, we describe how Theorem 2 yields a simple empirical test to uncover taste-based discrimination on commonly available cross-sectional data. We first describe the methodology and then apply our test to Census and NLSY-79 data.

It is worth presenting a high level motivation for our approach before we provide specific details. There is a tradition in labor economics of using the Kitagawa-Oaxaca-Blinder decomposition to determine both the presence of discrimination and to measure its magnitude. As Guryan and Charles (2013) explain, this method "separates differences in average wages, for example, into the part that is explained by differences in characteristics (e.g., education), the part that is explained by differences in returns to those characteristics (e.g., returns to education) and unexplained differences. Many in this literature have called the unexplained differences, or both the unexplained and the differences in returns, the result of discrimination."

We follow this tradition with one notable difference: we decompose the differences in wages at *all quantiles* of the distributions.[14] Indeed, a key insight of our theory is that we need to compare the entire wage distributions, and not only the average wages, if we want to test whether statistical discrimination alone can rationalize the data. Ideally, we would like to decompose the wage distributions into the part that is explained by differences in (prior) mean productivities and the part that is not, so that we can apply our strict first-order stochastic dominance test to the latter part. Unfortunately, we do not observe individual productivities. We instead observe individual characteristics which are commonly used as proxies for productivities. We decompose the wage distributions with respect to these characteristics.

In standard decompositions, researchers have to be careful to make sure that they do not control for too little (omitted variables) or too much. As Guryan and Charles (2013) explain: "the variables the researcher controls for might themselves be affected by discrimination. Controlling for such variables can cause the unexplained differences to understate the role that discrimination in general plays in determining wage gaps." By contrast, controlling for too much is less of an issue for our test. To start with, we are only interested in detecting taste-based discrimination, and not in measuring its magnitude. In addition, even if the variables we control for have been affected by discrimination, we can use this to our advantage. For instance, if we think there is discrimination in education, this implies that, on average, Black workers (having to overcome more barriers) will be of higher ability compared to similarly educated White workers. Any such positive selection makes the assumption of Theorem 2 *more* likely to be true. In fact, we go one step further. We compare Black workers with more schooling to lesser educated White workers. If the wages of White workers still strictly first-order stochastically dominate those of Black workers, then a conclusion that taste-based discrimination is present is invalid only if we believe that White workers with less schooling are more productive on average.

4.1. THE METHODOLOGY

We compare the wage distributions for Black (group 1) and White (group 2) workers within a given occupation. The assumption here is that wages in a given occupation are governed by a single wage function

[14] Recall that the distribution G_2 first-order stochastically dominates the distribution G_1 if, and only if, all quantiles of G_2 are higher than the respective quantiles of G_1.

and the different characteristics of workers affect their expected productivity (and hence their wage). Our goal is to test Theorem 2: that is, under the assumption that Black workers are more productive on average, we want to test whether the wage distribution G_2 of White workers strictly first-order stochastically dominates the wage distribution G_1 of Black workers. Recall that, if $G_2 \succ_1 G_1$, then the wage distributions cannot be the result of statistical discrimination alone. Thus, our null hypothesis is that statistical discrimination alone rationalizes the data, and we are interested in rejecting the null. We interpret the rejection of the null as evidence of taste-based discrimination.

There are two difficulties to overcome. The first is that we do not observe the true wage distributions G_1 and G_2; we observe empirical samples, instead. The second is to ensure we can apply Theorem 2, which requires Black workers to be more productive on average. Indeed, since we do not observe the true productivity distributions of the workers, we cannot directly control for that assumption and we, instead, need to find an indirect way.

The first difficulty is easily dealt with. We can test whether $G_2 \succ_1 G_1$ by Kolmogorov-Smirnov-type tests. This is equivalent to testing whether G_2 first-order stochastically dominates G_1 but not vice versa. Examples of such tests can be found in McFadden (1989) and Barrett and Donald (2003).

The second difficulty is more challenging. To overcome this, we assume that *unobserved* individual productivities are correlated with the individual characteristics we observe. Now, since the distributions of individual characteristics in any occupation vary between Black and White workers, we employ a version of the KOB decomposition to adjust for the differences in characteristics and then test for first-order stochastic dominance. More precisely, we first compute the (counterfactual) distribution $\widehat{G}_1$ of wages for Black workers that we would have observed had they faced the wage setting process of White workers that is, had employers perceived them to be White. This counterfactual distribution captures both potential statistical discrimination (via the different signals for each group) and taste-based discrimination. We then test the null hypothesis of $\widehat{G}_1 \not\succ_1 G_1$: if the null is rejected, statistical discrimination alone cannot explain the difference between the observed wage distribution G_1 and the counterfactual distribution $\widehat{G}_1$.

Formally, let X be a vector of observable individual characteristics such as years of schooling, age and state of residence. Let $G_i(\cdot|X)$ be the observed distribution of wages conditional on characteristics X and F_i° the joint probability over expected productivity (in the eyes of the employer) and individual characteristics. The notation overloads are deliberately made to obviate the introduction of new notation. We perform the following decomposition:

$$
\begin{aligned}
G_1(w) - G_2(w) &= \int G_1(w|X)dF_1^\circ(X) - \int G_2(w|X)dF_2^\circ(X), \\
&= \left[\int G_1(w|X)dF_1^\circ(X) - \int G_2(w|X)dF_1^\circ(X)\right] \\
&\quad + \left[\int G_2(w|X)dF_1^\circ(X) - \int G_2(w|X)dF_2^\circ(X)\right], \\
&= \left[G_1(w) - \widehat{G}_1(w)\right] + \left[\widehat{G}_1(w) - G_2(w)\right],
\end{aligned} \tag{1}
$$

where $\widehat{G}_1(w) := \int G_2(w|X)dF_1^\circ(X)$ and the integrals are taken with respect to the marginal distributions over characteristics. A recent interpretation of this decomposition is provided by Bohren, Hull, and Imas (2022). They interpret the term on the left $G_1(w) - G_2(w)$ as *total* discrimination, the first term on the right $G_1(w) - \widehat{G}_1(w)$ as *direct* discrimination and the final term $\widehat{G}_1(w) - G_2(w)$ as *systemic* discrimination. Total discrimination compounds the differential treatment of Black and White workers before entering the job market (for instance, due to barriers in educational attainment) with the differential treatment after entering the job market. We are interested in the direct discrimination term, that is, the comparison of G_1 and $\widehat{G}_1$. The comparison is analogous to a correspondence study since we compare the wages of Black workers with the wages they would have obtained had they been treated as White workers (but retaining their individual characteristics). Conceptually, it is similar to exposing employers to resumes which differ only in names.

To compute the counterfactual distribution $\widehat{G}_1$, we follow the method of Chernozhukov, Fernández-Val, and Melly (2013). In a nutshell, this consists of estimating the marginal $F_i^\circ(X)$ over individual characteristics from the empirical distribution and the conditional $G_i(\cdot|X)$ from either distribution or quantile regressions. We then test whether $\widehat{G}_1 \succ_1 G_1$, which the approach of Chernozhukov, Fernández-Val, and Melly (2013) permits. Importantly, their method allows us to make statistical inference on the entire distributions, which is needed to test for first-order stochastic dominance.

Now, suppose we conclude that $\widehat{G}_1 \succ_1 G_1$. We can then infer that the wage distributions *cannot* be rationalized by statistical discrimination alone if

$$
\int \mathbb{E}_{F_1^\circ}[\theta_1|X]dF_1^\circ(X) \geq \int \mathbb{E}_{F_2^\circ}[\theta_2|X]dF_1^\circ(X).
$$

For instance, this hypothesis is satisfied whenever the productivity distribution θ_i is independent of race i, conditional on characteristics X. It is, however, weaker. The assumption allows for the expected productivity $\mathbb{E}_{F_1^\circ}[\theta_1|X] \geq \mathbb{E}_{F_2^\circ}[\theta_2|X]$ of Black workers to be higher for some individual characteristics X and lower for others. Economic theories support both scenarios. On the one hand, barriers to entry into the labor market and obtaining education suggest that Black workers are positively selected into the labor market (which is our belief). While we are not aware of empirical evidence supporting positive selection of Black workers, Ashraf, Bandiera, Minni, and Quintas-Martınez (2022) document that women are positively selected into the labor market (and that women are more comparatively productive than men in countries where female labor force participation is lower). On the other hand, classic theories of statistical discrimination (such as Coate and Loury, 1993) argue that Black workers may under-invest in productive skills and consequently make stereotypes self-fulfilling. To deal with this latter hypothesis, we also compare Black workers with more schooling to White workers with less; here, it is significantly harder (and, in our view, implausible) to argue that more educated Black workers have lower mean productivity.

We end this subsection with a brief discussion on the choice of the covariates X that should be included in the computation of the counterfactual distribution $\widehat{G}_1$. We suggest including as many covariates as necessary, subject to data and computational limitations, to make the ordered means assumption as plausible as possible. It is not essential that the employers observe all the variables in X when setting wages. Employers may not observe certain covariates, but may observe other informative signals (unobserved by the analyst) that are correlated with the unobserved variables in X. In other words, since the signals employers observe may be arbitrarily correlated with the the individual characteristics the analyst observes, it makes no differences to the empirical analysis whether we assume that employers observe the individual covariates X we use or not. A similar observation appears in Altonji and Pierret (2001) who employ NLSY-79 data and use Armed Forces Qualification Test (AFQT) scores, among other controls, to proxy the information employers obtain over time. We explicitly make this remark because one of our robustness checks uses NLSY-79 data and we include AFQT scores and past wages—variables that may not be observed by employers—as controls.

4.2. EVIDENCE OF TASTE-BASED DISCRIMINATION

We apply our test on two datasets. We conduct our main analysis with Census data. While these data have limitations, this serves as a proof of concept of how our test can be applied on commonly used and publicly available large datasets. We conduct an additional robustness check using NLSY-79 data and exploit the fact that it is a panel.

We begin by describing the sample construction. In each dataset, we focus on the wages of prime-aged men to avoid the typical selection issues associated with female labor force participation.[15]

[15] Examining wage distributions by race is also subject to selection as labor force participation is substantially lower for Black compared to White male workers. Since our test requires mean productivities to be ordered, the interpretation of our test results would only be invalid if selection is such that the more educated Black workers who participate in the labor force are, on average, less productive than their less educated White counterparts. This is contrary to the more reasonable assumption that the least productive Black men stay out of the labor force which is selection that works in favor of our ordered means requirement.

4.2.1. The data

Census: We use the American Community Survey (ACS) for the years 2007-2021 that contains all households and persons from the 1% sample. We restrict the data to men aged 30-55 years (to capture their prime labor market years) and to those who are employed, working full time (50-52 weeks), and working for a wage. The main income measure we use is the INCWAGE variable capturing "Wage and Salary Income."

Because the wage distribution decomposition method of Chernozhukov, Fernández-Val, and Melly (2013) that we employ is not completely non-parametric, we conduct our analysis at the occupation level. This allows for more flexibility of the wage functions across occupations (compared to using the full sample with occupation controls). We create eleven general categories of occupation status from the four digit codes (2010 basis) available in the Census.[16] The 11 groups are: (1) Management, Business, and Financial, (2) Professional, (3) Service, (4) Sales and related, (5) Office and administrative support, (6) Farming, fishing, and forestry, (7) Construction and extraction, (8) Installation, maintenance, and repair, (9) Production, (10) Transportation, and material, and (11) Armed forces occupations.

Race is measured from the self reported race category in the census. We also use data on age, census year, census region and highest grade attained; the latter is used to create samples with ordered means. Specifically, we consider the sample of White workers with a high school degree and Black workers who have attended some college or possess an associate's degree. The more educated Black workers are assumed to have higher mean productivity. It is, of course, possible to consider other types of samples with ordered means such as White and Black workers without and with high school diplomas or possessing bachelor's and master's degrees respectively. Our choice was governed by at least two reasons. First, compared to the two options in the previous sentence, our samples constitute a larger proportion of the data. Second, Census data does not have information on the university attended. Consequently, it would be hard to control for the very heterogeneous quality of college education were we to compare White and Black workers with bachelor's and master's degrees respectively.

NLSY-79: From the NLSY-79 we restrict our sample to men who worked full time (52 weeks) in 1998 and 2000 and who reported positive wages in those years. We selected those years as individuals in the NLSY-79 were between the ages of 14-22 at the time of their first interview in 1979. Hence, in 2000 these individuals were in the age range of 35-43, which were their prime labor market years. The main wage variable we use from the NLSY-79 is about the respondent's "amount of wages, salary, and tips" in the past year (so survey year 2000 relates to wages in 1999, etc.).

In addition to variables capturing occupation in 2000 and 1998 (3 digit CPS 1980 codes), we also have the highest grade completed (as of calendar year 2000), and the AFQT percentile score (all respondents

[16] This is based on the Census occupation classification available here.

took the AFQT in 1981 and we use the 2006 revised percentiles of this measure from the data).

4.2.2. Results

The one sentence summary of our results is that, across several broad occupation categories, the wage distribution of Black workers is strictly first-order stochastically dominated by the counterfactual distribution of wages they would receive were they to be treated as White *despite Black workers having higher educational attainment by construction*.[17] We interpret this as suggestive evidence of taste-based discrimination via the lens of Theorem 2.

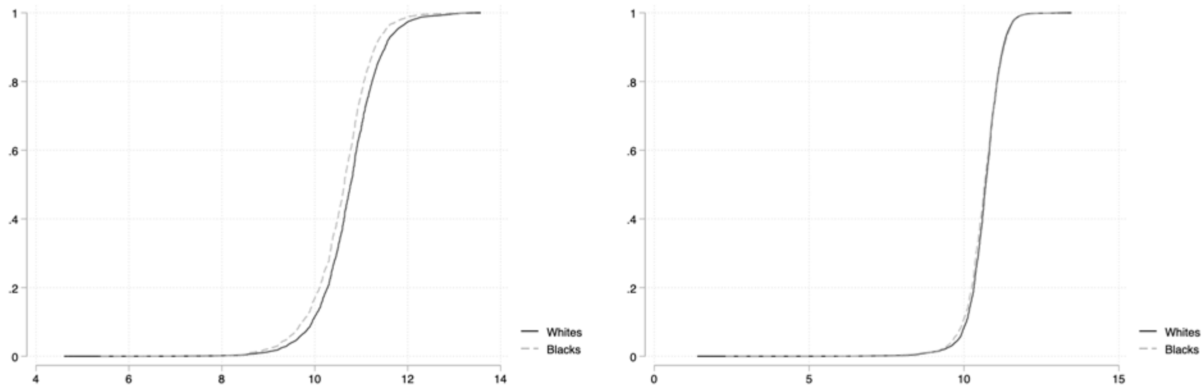
Before turning to the results from the quantile decomposition, we first present some descriptive and easy to visualize evidence. A strength of our test is that such visual evidence is straightforward for researchers and policy makers to generate. They can simply plot the wage distributions within well defined cells. As long as we are convinced that the assumption of equal or ordered means (depending on whether we are testing Theorem 1 or 2) is satisfied, a strict first-order stochastic dominance ordering can provide suggestive evidence of discrimination in general and taste-based discrimination in particular.

In Figure 1, we plot the distribution of wages of Black workers with some college or an associates' degree and White workers who have just a high school degree in two out of the eleven broad occupation groups mentioned above. Panel 1a shows the distribution of wages in jobs related to sales (such as retail salespersons, insurance sales agents, etc.) and Panel 1b shows the distribution of wages in jobs related to production (such as metal workers, painters, woodworkers, etc.). Panel 1a is visually stark: despite having lower years of schooling, White workers' wages clearly appear to strictly first-order stochastically dominate the wages of Black workers. The evidence in Panel 1b is less clear. We use a Kolmogorov-Smirnov (KS) test statistic to confirm these visual tests. For sales, the KS statistic for the null that White workers' wages first-order stochastically dominate those of Black workers is 0 (p-value of 1), while the KS statistic for the null that Black workers' wages first-order stochastically dominate those of White workers is -0.10 (p-value 0). Hence, we *cannot* reject the former but *can* reject the latter. This demonstrates *strict* first-order stochastic dominance. For production, the KS statistic for the null that White workers' wages first-order stochastically dominate those of Black workers is 0.01 (p-value of 0.14), while the KS statistic for the null that Black workers' wages first-order stochastically dominate those of White workers is -0.04 (p-value 0). Thus, the evidence is similar, but as the figure suggests, it is less stark.

While these descriptive graphs are informative, it is possible that these wage distributions are the result of different distributions of (non-educational) characteristics amongst Black and White workers. To hold these other characteristics constant, we turn to our main specification: the quantile decomposition dis-

[17] It is not the case that some college or associate's degrees are an overall negative signal for employers. In the Appendix, we show that wages of Black workers with some college or associate's degrees first order stochastically dominate the wages of Black workers with just a high school degree.

Figure 1: Descriptive Evidence with Ordered Means



(a) Sales (b) Production

Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks). Black workers with some college or associate's degrees and White workers with high school degrees.

cussed earlier using the methodology of Chernozhukov, Fernández-Val, and Melly (2013). We conduct the decomposition for the same two broad occupation categories above—"Sales" and "Production"—separately (and recall that in each case Black workers have some college or an associate's degree and White workers have a high-school diploma) with the following rich set of controls: census region of residence (9 regions), age, age squared, dummies for finer occupation controls (these are the most detailed occupation codes available in the ACS), and Census year.[18]

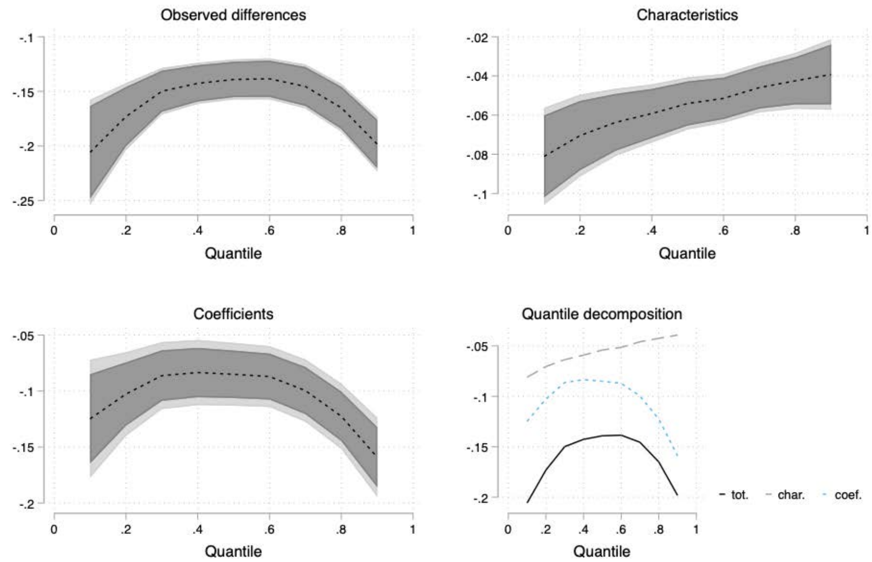
Figures 2 and 3 display the results for both occupation categories. Each figure contains four panels: in each panel, quantiles are on the x-axis and differences between the log annual wages of Black workers and White workers are on the y-axis. Note that an equivalent way of stating that the wage distribution for a group strictly first-order stochastically dominates the other is that the wages for the first group are weakly higher than the second at *all* quantiles and strictly higher at *some*. Formally, $G_i \succ_1 G_j$ is equivalent to $G_i^{-1}(q) \geq G_j^{-1}(q)$ for all $q \in [0,1]$ with the inequality strict for at least one q.[19] For the decompositions, we use the default settings of the CDECO command in Stata: estimation is based on linear quantile regressions based on Koenker and Bassett (1978), hundred bootstrap replications are performed for inference, and both pointwise and uniform confidence intervals at 95% are constructed.

The top left panel plots the difference in wages at each quantile of the wage distribution; this is the total discrimination we described above (that is, $G_1^{-1} - G_2^{-1}$). The top right panel is the share of that differ-

[18]For a sense of these finer occupation categories, the following are the top three finer occupations within each broader occupation category. Within "Sales": 1) First line supervisors of sales agents, 2) Sales representatives, 3) Retail salespersons. Within "Production": 1) First line supervisors of production, 2) Other production workers, 3) Welders.

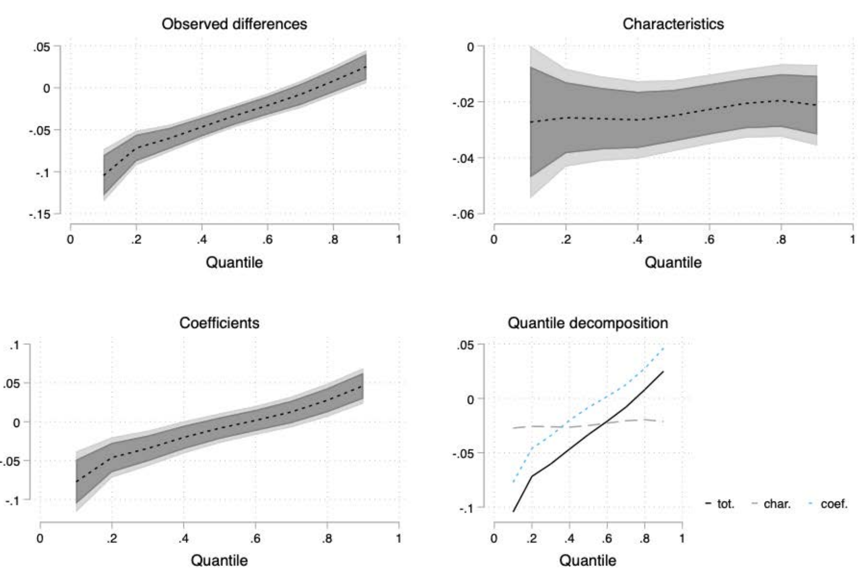
[19]We define $G_i^{-1}(q) = \inf\{w \in [0,\overline{w}] \mid G_i(w) \geq q\}$ since the wage distributions need not be strictly increasing.

Figure 2: Quantile Decompositions: Sales



Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

Figure 3: Quantile Decompositions: Production



Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

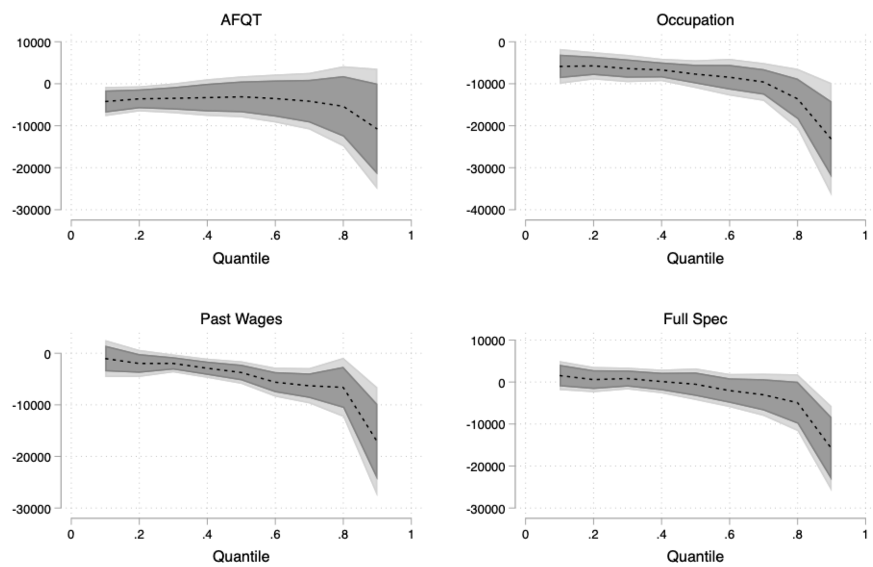
ence attributed to the different characteristics of Black and White workers, the aforementioned systemic discrimination (that is, $\widehat{G}_1^{-1} - G_2^{-1}$). The main panel of interest for us is the bottom left panel which captures the direct discrimination (that is, $G_1^{-1} - \widehat{G}_1^{-1}$). The bottom right panel in each figure summarizes these different pieces of the decomposition; to summarize, this is the equivalent quantile decomposition analogue of distribution decomposition (1).

In Figure 2, observe that the wage structure effect (bottom left panel labeled "Coefficients") is significantly less than 0 at all quantiles of the wage distribution. That is, Black workers are paid less at all quantiles of the wage distribution relative to the counterfactual wages they would receive were they to be treated as White. In other words, if we were to believe that conditioning on the above mentioned observables, while fixing the education of Black workers to be higher than that of White workers, the average productivity of Black workers is at least as high as White workers, then this is evidence of taste-based discrimination in sales based occupations. The same is not true in Figure 3. Once we factor in workers' characteristics in production occupations, the wage structure effect (the bottom left panel) is not below zero at all quantiles. Hence, we cannot conclude that there is strict first-order stochastic dominance. Intuitively these results resonate. In occupations such as sales, there are likely customer interactions and other aspects that might be subject to one's tastes and prejudices. In production occupations like welding and woodworking, perhaps the scope for taste-based discrimination is less (although, we stress again that not finding strict first-order stochastic dominance does not imply the absence of discrimination).

In the Appendix, we provide the decompositions for all the remaining broad occupation categories and the results are intuitive. We find evidence of strict first-order stochastic dominance in occupations where the scope for subjective evaluations might be higher (like professional occupations or management related jobs). In occupations like farming or the service sector where compensation structures might be less discretionary, it makes sense that taste-based discrimination plays a less significant role.

We end this section with a description of our results from the NLSY-79. This data set is not contemporary, is smaller than the Census but has the advantage of being a panel. This allows us to control on past wages and occupations. These controls allow us to compare Black and White workers who worked the *same* job in the previous period at the *same* wage. If there was no discrimination the previous period wages, two employees of each group who are paid the same must have the same expected productivity (since we assume that wages only depend on expected productivities). Conversely, if employers had engaged in discrimination against Black workers in the previous period, we expect Black workers to be *more* productive than White workers for them to have been paid the *same* wage. The remaining possibility is that Black workers have a lower expected productivity, but got paid the same as White workers (with higher mean productivity) in the previous period. We might expect this if we think there is strong bias in the opposite direction or forces like affirmative action at play. However, if this were the case, we would not expect a *reversal* of this pattern within one period.

Figure 4: Quantile Decompositions: Wage Structure Effects from the NLSY



Data is from the NLSY-79. Wages recorded in calendar year 2000 for men working full time (52 weeks). CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Only wage structure effect (coefficient plot from CDECO) is displayed. All estimates control for highest grade completed, age, and age-squared. In addition, the top left panel controls for AFQT percentile scores, top right panel controls for dummies for occupation in 2000, bottom left controls for wages in 1998, and the bottom right controls for all of these as well as dummies for occupation in 1998. The first three panels clearly show that White workers' wages strictly first-order stochastically dominate those of Black workers. For the full specification estimate, the KS statistic p-values for the null that White workers' wages FOSD Black workers' wages is 0.66 and the p-values for Black workers' wages FOSD White workers' wages is 0.03.

Figure 4 only displays the wage structure effect (which is required to conclude $\widehat{G}_1 \succ_1 G_1$) under increasingly stringent controls; all figures control for highest grade completed, age, and age-squared. The bottom right figure displays the full specification that controls not only for past wages and occupations but also AFQT scores. Strict first-order stochastic dominance is clearly visible in all plots; in the bottom right plot $\widehat{G}_1^{-1}$ and G_1^{-1} are not significantly different at low quantiles but at high quantiles $\widehat{G}_1^{-1}$ is significantly higher. The KS statistic p-values confirm the visual evidence.

5. EXTENDING THE THEORY

In this section, we generalize the results in Section 3 along several directions. We first impose natural shape restrictions on the set of permissible wage functions and derive stronger testable implications on the set of wage distribution pairs. We then characterize the set of rationalizable wage distribution pairs when we do not assume mean productivities are ordered. We show how these results can be inverted to derive bounds on the productivity differences required to rationalize the wage distributions. In all characterizations of rationalizable wage functions in this section, analogues of statement (iv) and (iii) of Theorems 1 and 2 respectively, apply. We do not explicitly state these for brevity.

5.1. CONCAVE AND CONVEX WAGES

The results from Section 3 allow for arbitrary wage functions. There might be settings where there is natural structure on the wage functions. Imposing such structure has the advantage of strengthening the testable implications of the model.

Motivated by increasing inequalities in income, one natural assumption to impose is that wages are convex in the expected productivity of the worker. We denote

$$\mathcal{W}_{conv} := \{W \in \mathcal{W} \,|\, W \text{ is strictly convex}\}$$

to be the set of strictly increasing and strictly convex (and therefore continuous) functions.[20] Likewise, we use $\mathcal{W}_{conc}$ to denote the set of strictly increasing and strictly concave functions.

We say that wage distribution G_i dominates distribution G_j in the *strict concave order*, if

$$\int_0^{\overline{w}} M(w)dG_i(w) > \int_0^{\overline{w}} M(w)dG_j(w)$$

for every strictly increasing, strictly concave function $M : \mathbb{R}_+ \to \mathbb{R}_+$. This order is closely related to the mean-preserving spread order $\succcurlyeq_2$. The latter implies second-order stochastic dominance but not vice versa since second-order stochastic dominance does not require equal means. The strict concave order implies strict second-order stochastic dominance, but does not require both distributions to have equal means. The *strict convex order* can be analogously defined when the above inequality holds for all strictly increasing, strictly convex function $M : \mathbb{R}_+ \to \mathbb{R}_+$.

We derive a version of Theorem 1 assuming wages are convex.

THEOREM 3. *Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_=$, $\mathcal{W}_{conv}$) if, and only if, neither G_1 nor G_2 dominates the other in the strict concave order.*

This result can also be used to derive the analogous version of Theorem 2 for the case of ordered means. If we were to instead assume wage functions were concave, the statement would be verbatim with $\mathcal{W}_{conc}$ replacing $\mathcal{W}_{conv}$ and the strict convex order replacing the strict concave order.

If $G_i \succ_1 G_j$, then G_i also dominates distribution G_j in both the strict concave and convex orders. Thus,

[20] The only possible point of discontinuity is at 0, but strict convexity and strict monotonicity guarantees that there is no jump at 0.

Theorem 3 shows that the testable implications are stronger under the assumption of convex wages but, once again, they take a form that is easy to take to the data. There are well-known tests for higher orders of stochastic dominance developed in the econometrics literature (once again, see Barrett and Donald, 2003).

5.2. NON-ORDERED MEANS

All of the above results required mean productivities to be either equal or ordered. Instead, suppose we assume that

$$\mathcal{H}_{|1-2|\leq d} := \{(H_1, H_2) \mid |\mathbb{E}_{H_1}[\theta_1] - \mathbb{E}_{H_2}[\theta_2]| \leq d\},$$

that is, both groups have productivity distributions whose mean differs by at most $d \in \mathbb{R}_+$. Additionally, we assume

$$\mathcal{W}_{L1} := \{W \in \mathcal{W} \mid |W(\theta) - W(\theta')| \leq |\theta - \theta'| \text{ for all } \theta, \theta' \in \mathbb{R}_+\},$$

that is, wage functions are 1-Lipschitz. This technical restriction imposes some discipline on the wage function, that is, it ties changes in wages with changes in productivities. It is weaker than assuming that wage functions are differentiable with slope less than 1.

We now characterize wage distributions that are rationalizable under these assumptions.

THEOREM 4. *Wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_{|1-2|\leq d}$, $\mathcal{W}_{L1}$) if, and only if, either*

(i) the wage gap is less than d, that is, $|\mathbb{E}_{G_1}[w] - \mathbb{E}_{G_2}[w]| \leq d$, or

(ii) neither G_1 nor G_2 strictly first-order stochastically dominates the other.

As Theorem 1 shows, if wages are not ordered by first-order stochastic dominance, they can be rationalized with equal mean productivities. Consequently, they will be rationalized (given $\mathcal{H}_{|1-2|\leq d}$, $\mathcal{W}_{L1}$) for any $d \geq 0$; the additional restriction on the wage function $\mathcal{W}_{L1} \subset \mathcal{W}$ does not affect Theorem 1 (the older version Deb and Renou, 2022 has a formal statement of this). So the real contribution of Theorem 4 is the first condition (i). This shows that the wage gap can be a useful statistic to test for the presence of taste-based discrimination, but *only* when the wage distributions are ordered by strict first-order stochastic dominance.

An issue with applying the result in Theorem 4 to data is that the researcher has to choose the appropriate mean productivity difference d to test for. However, the statement of this theorem can be inverted to show that the wage gap is a *tight* lower bound for mean productivity differences required to rationalize the wage distributions. In other words, if we want statistical discrimination alone to rationalize the data, the wage gap is the smallest difference in productivity means required, whenever one wage distribution first-order stochastically dominates the other.

THEOREM 4 (CONTINUED). *Suppose $G_2 \succ_1 G_1$. Then, there exist (H_1, F_1, H_2, F_2, W) that jointly induce G_1, G_2 such that $W \in \mathcal{W}_{L1}$ and $\mathbb{E}_{H_2}[\theta] - \mathbb{E}_{H_1}[\theta] = \mathbb{E}_{G_2}[w] - \mathbb{E}_{G_1}[w]$.*

Moreover, every (H_1, F_1, H_2, F_2, W) with $W \in \mathcal{W}_{L1}$, that jointly induce G_1, G_2 satisfy $\mathbb{E}_{H_2}[\theta] - \mathbb{E}_{H_1}[\theta] \geq \mathbb{E}_{G_2}[w] - \mathbb{E}_{G_1}[w]$.

In words, this result states that every (H_1, F_1, H_2, F_2, W) with $W \in \mathcal{W}_{L1}$ that induce wage distributions $G_2 \succ_1 G_1$ have the feature that the differences in mean productivities is at least the wage gap, and that this bound is tight. Thus, the wage gap is a useful measure of the minimum productivity difference required for statistical discrimination alone to rationalize the wage distributions. The data can be used to contextualize this bound and, in a sense, we already conduct such an exercise in Section 4.2.

To see this, let us revisit our empirical application. Compare prime aged (30-55) Black and White full time working males in sales with high school degrees. Controlling for their characteristics,[21] the mean gap in wages is $-11,807$ USD (Black workers earn less). So, if we want statistical discrimination alone to explain the data, we need the mean productivities to differ by at least this wage gap. To put this number in perspective, let us compare Black workers with high-school degrees to Black workers with associate's degrees in sales jobs. The difference in mean wages between these two groups is -4329 USD (workers with associate's degrees make more). This says that, absent taste-based discrimination, the difference in mean productivity between White and Black high-school educated workers corresponds to more than two additional years of schooling for Black workers.

Instead of bounding absolute productivity differences, one could also try to bound percentage productivity differences. Formally, for $\alpha \in (0, 1)$, let

$$\mathcal{H}_{1/2 \geq \alpha} := \{(H_1, H_2) \,|\, \mathbb{E}_{H_1}[\theta_1] \geq \alpha \mathbb{E}_{H_2}[\theta_2]\},$$

denote the set of productivity distribution pairs such that group 1's mean productivity is at least a fraction α of that of group 2. As with Theorem 4, one could first characterize the set of rationalizable wage distributions and then invert result to derive the bound. Unfortunately, there is too little structure in the model to bound percentage differences.

THEOREM 5. *Suppose $G_1(0) = G_2(0) = 0$. Then, for every $\alpha \in (0, 1)$, every pair of wage distributions is rationalizable (given $\mathcal{H}_{1/2 \geq \alpha}$, $\mathcal{W}_{L1}$).*

[21] All wage gaps in this paragraph control for age, age-squared, Census region, 2010 occupation codes, and Census year.

This result states that, for any fraction α, every pair of wage distributions are rationalizable with productivity distributions whose means are within a fraction α of each other. Loosely speaking, this is because we can consider wage functions that are very flat for low wages but have slope one for higher wages. With such a wage function, absolute differences in higher wages correspond to absolute differences in productivities but because the values of the latter are high, the percentage difference between mean productivities becomes small. The additional assumption $G_1(0) = G_2(0) = 0$ is trivially satisfied in any application since we restrict attention to working adults and no one works for zero wages.

6. CONCLUDING REMARKS

In this paper, we developed a simple but general framework that lends itself to testing for taste-based discrimination in widely available cross-sectional wage data. We view our contributions to be several. First, unlike a bulk of the literature, our modeling choices allow for unrestricted signals and can flexibly capture imperfectly competitive labor markets. Despite this generality, the testable implications of the model are easy to describe and test. We demonstrate how an ordered means assumption can be validated on either cross sectional or panel data by comparing less educated White workers to more educated Black workers or by conditioning on past wages and occupations respectively. Our theoretical results provide a lens through which the decompositions of wage distributions can be used to uncover evidence of discrimination in general and taste-based discrimination in particular. Our empirical results document stark patterns in US Census data: controlling for observables, the wages of Black workers are strictly first-order stochastically dominated by those of White workers across occupations. This provides suggestive evidence of taste-based discrimination against Black workers in US labor markets. Finally, we demonstrate the flexibility of our framework by deriving the testable implications of the model under difference assumptions on the set of permissible productivity distribution pairs and wage functions.

We end the paper by discussing a few remaining assumptions of the model and suggest some directions for future research. Throughout, we assumed that there were two groups. One could, in principle, test whether the wage distributions for three or more groups—each with their own productivity distribution and signal but with there being a common wage function—are rationalizable. Considering more than two groups leads to stronger testable implications. Indeed, in the Online Appendix, we generalize Theorem 1 and show that wage distributions from $n \geq 2$ groups are rationalizable (given $\mathcal{H}_=$ defined for n groups and $\mathcal{W}$) if, and only if, we cannot find two subsets of the groups and convex combinations of their wage distributions that are ordered by strict first-order stochastic dominance. It is similarly possible to generalize Theorem 2 by, for instance, comparing a disadvantaged group with higher mean productivity against multiple (comparatively) advantaged groups.

A critical assumption in our model is that wages only depend on the posterior mean which is, of course, an assumption that is commonly made when assuming perfectly competitive labor markets.[22] With that said, one could allow wages to depend on higher moments of the posterior distribution inferred by employers upon observing a signal realization. Indeed, the seminal work of Aigner and Cain (1977) allows wages to depend on both the mean and the variance of the posterior. We show, in the Online Appendix, that such additionally generality renders vacuous the testable implications of our model. Specifically, even if wages are linear in the mean and variance of the posterior, every pair of wage distributions can be rationalized under the assumption of equal mean productivities.

While we considered many combinations of assumptions on the sets of productivity distributions and wage functions, there are of course other restrictions that one could impose based on the context. For instance, information or estimates about labor market competition may allow the researcher to bound slopes of the wage function. This in turn, could allow the researcher to infer bounds on mean productivity differences (absolute or percentage) required to rationalize the wage distributions absent taste-based discrimination. We demonstrate how to do so in the Online Appendix.

Lastly, our empirical application demonstrated (by comparing less educated White workers to more educated Black workers) how the data can validate assumptions on the set $\widehat{\mathcal{H}}$ of productivity distributions (in this case, $\widehat{\mathcal{H}} = \mathcal{H}_{\geq}$). Our approach can be adapted to richer data sets where there may be additional information about worker productivity. In a sense, the robustness check we conduct using NLSY-79 data does precisely this by exploiting the panel data structure to control for past wages and occupations. But this of course, does not employ the full richness of the panel as we only use information in one previous period. To do so, we would need to introduce dynamics into our framework which is something we hope to do in future research.

[22] This assumption also requires that productivities are one dimensional. This too can be relaxed although once again, the testable restrictions of the model become vacuous without additional structure.

A. PROOF OF THEOREM 1

In the text, we already proved the equivalence of statements (i), (iii) and (iv). We also argued that if the wage distributions G_1 and G_2 are rationalizable with $(\mathcal{H}_=, \mathcal{W})$, then neither G_1 nor G_2 strictly first-order stochastically dominates the other. That is, we showed that statement (i) implies statement (ii). We now prove the converse. Throughout, when we use the indices i and j, we assume that $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$.

We start with a preliminary observation about strict first-order dominance. It is well-known that G_i first-order stochastically dominates G_j if, and only if,

$$\int_0^{\overline{w}} M(w) dG_i(w) \geq \int_0^{\overline{w}} M(w) dG_j(w),$$

for all (not necessarily strictly) increasing (and not necessarily continuous) functions $M : \mathbb{R}_+ \to \mathbb{R}_+$ (see Chapter 1 of Shaked and Shanthikumar, 2007).[23] In addition, it is easy to see that G_i *strictly* first-order stochastically dominates G_j if, and only if, G_i first-order stochastically dominates G_j and is not first-order stochastically dominated by G_j. It follows immediately that G_i *strictly* first-order stochastically dominates G_j if, and only if,

$$\int_0^{\overline{w}} M(w) dG_i(w) \geq \int_0^{\overline{w}} M(w) dG_j(w),$$

for all increasing functions $M : \mathbb{R}_+ \to \mathbb{R}_+$ with a strict inequality for some M. The following lemma is an immediate consequence.

LEMMA 1. *Suppose $G_i \neq G_j$ and $G_i \not\succ_1 G_j$. Then there exists a strictly increasing and continuous function M such that*

$$\int_0^{\overline{w}} M(w) dG_i(w) < \int_0^{\overline{w}} M(w) dG_j(w).$$

PROOF OF LEMMA 1. Since $G_i \not\succ_1 G_j$, either there exists an increasing function $\widehat{M} : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\int_0^{\overline{w}} \widehat{M}(w) dG_i(w) < \int_0^{\overline{w}} \widehat{M}(w) dG_j(w),$$

[23] Since the distributions G_1 and G_2 are supported on a subset of $[0, \overline{w}]$, the expectation of these functions always exist. Further, there is no loss in taking the domain and range of M to be $\mathbb{R}_+$ as opposed to $\mathbb{R}$.

or for all increasing functions $\widehat{M} : \mathbb{R}_+ \to \mathbb{R}_+$,

$$\int_0^{\overline{w}} \widehat{M}(w) dG_i(w) = \int_0^{\overline{w}} \widehat{M}(w) dG_j(w).$$

The latter case is equivalent to $G_i = G_j$, since we can choose $\widehat{M}$ to be the indicator of the intervals $(w, \overline{w}]$ for all $w \in [0, \overline{w})$ (recall that $G_i(\overline{w}) = G_j(\overline{w}) = 1$ and $G_i(0) = G_j(0)$). Since $G_i \neq G_j$, the former must therefore hold. Since the inequality is strict, we can approximate the increasing function $\widehat{M}$ by a strictly increasing and continuous function $M : \mathbb{R}_+ \to \mathbb{R}_+$ such that the inequality remains strict (this can be done constructively). This completes the proof. ■

We now complete the proof of Theorem 1 by showing statement (ii) implies statement (i). If $G_1 = G_2$, we can just take $W(\theta) = \theta$ along with $F_i = H_i$. So henceforth, we assume $G_1 \neq G_2$.

Since $G_1 \not\succ_1 G_2$ and $G_2 \not\succ_1 G_1$, Lemma 1 implies there exist strictly increasing and continuous functions $M' : \mathbb{R}_+ \to \mathbb{R}_+$ and $M'' : \mathbb{R}_+ \to \mathbb{R}_+$ such that

$$\int_0^{\overline{w}} M'(w) dG_1(w) > \int_0^{\overline{w}} M'(w) dG_2(w) \quad \text{and}$$
$$\int_0^{\overline{w}} M''(w) dG_1(w) < \int_0^{\overline{w}} M''(w) dG_2(w).$$

Therefore, by the intermediate value theorem, there exists an $\alpha \in (0, 1)$ such that

$$\int_0^{\overline{w}} [\alpha M'(w) + (1-\alpha) M''(w)] dG_1(w) = \int_0^{\overline{w}} [\alpha M'(w) + (1-\alpha) M''(w)] dG_2(w).$$

We denote $M := \alpha M' + (1-\alpha) M''$. Since, affine transformations of M will not affect the above equation, we can, without loss of generality, assume that $M(0) = 0$

Define $\overline{\theta} := M^{-1}(\overline{w})$. Take a continuous, strictly increasing function $W : \mathbb{R}_+ \to \mathbb{R}_+$ that satisfies $W(\theta) := M^{-1}(\theta)$ for $\theta \in [0, \overline{\theta}]$ and define $F_i(\theta) := G_i(W(\theta))$. Clearly, F_i is a well-defined distribution.

Finally, observe that

$$\int_0^{\overline{\theta}} \theta dF_1(\theta) = \int_0^{\overline{w}} M(w) dG_1(w) = \int_0^{\overline{w}} M(w) dG_2(w) = \int_0^{\overline{\theta}} \theta dF_2(\theta)$$

where the first and last equalities follow from the change of variable $\theta = M(w)$ and the fact that $F_i(M(w)) = G_i(w)$ by construction. Taking $H_i = F_i$ completes the proof.

A remark is in order. While this proof is non-constructive, it is possible to show (as we do in the older version Deb and Renou, 2022) that there always exists a piece-wise linear 1-Lipschitz wage function, which rationalizes the wage distributions.

B. PROOF OF THEOREM 2

This result was proved in the body of the paper.

C. PROOF OF THEOREM 3

(Only if.) The proof is by contradiction. Suppose that the wage distributions G_1 and G_2 are rationalizable (given $\mathcal{H}_=$, $\mathcal{W}_{conv}$) and yet G_i dominates G_j in the strict concave order ($i \neq j$). Since the wage distributions are rationalizable, there exists a strictly increasing, strictly convex wage function W, distributions (F_1, F_2) such that $G_k(w) = F_k(W^{-1}(w))$ for all $w \in [0, \overline{w}]$ and priors (H_1, H_2) such that $\mathbb{E}_{H_1}[\theta] = \mathbb{E}_{H_2}[\theta]$ and $F_k \succcurlyeq_2 H_k$ for $k \in \{1, 2\}$.

Since W is strictly increasing and strictly convex, W^{-1} is strictly increasing and strictly concave. To see the latter, choose $(w, w') \in [0, \overline{w}]^2$ with $w \neq w'$. From the strict convexity of W, we have that

$$\alpha W(W^{-1}(w)) + (1-\alpha)W(W^{-1}(w')) > W(\alpha W^{-1}(w) + (1-\alpha)W^{-1}(w')).$$

for all $\alpha \in (0, 1)$. Since W^{-1} is strictly increasing, it follows that

$$W^{-1}\left(\alpha W(W^{-1}(w)) + (1-\alpha)W(W^{-1}(w'))\right) > W^{-1}\left(W(\alpha W^{-1}(w) + (1-\alpha)W^{-1}(w'))\right),$$

which is equivalent to

$$W^{-1}(\alpha w + (1-\alpha)w') > \alpha W^{-1}(w) + (1-\alpha)W^{-1}(w'),$$

as required.

Finally, since we have assumed G_1 and G_2 are rationalizable, we have that

$$\int_0^{\overline{w}} W^{-1}(w)dG_i(w) = \int_0^{\overline{\theta}} \theta dF_i(\theta) = \int_0^{\overline{\theta}} \theta dH_i(\theta) = \int_0^{\overline{\theta}} \theta dH_j(\theta) = \int_0^{\overline{\theta}} \theta dF_j(\theta) = \int_0^{\overline{w}} W^{-1}(w)dG_j(w),$$

which is the required contradiction since G_i dominates G_j in the strict concave order.

(If.) The argument is almost identical to the part of the proof of Theorem 1 showing statement (ii) implies statement (i). Since neither G_1 nor G_2 dominates the other in the strict concave order, there exists a strictly increasing, strictly concave function M such that

$$\int_0^{\overline{w}} M(w)dG_1(w) = \int_0^{\overline{w}} M(w)dG_2(w).$$

Since, affine transformations of M will not affect the above equation, we can, without loss of generality, assume that $M(0) = 0$ and $M(\overline{w}) = \overline{\theta}$.

Define $\overline{\theta} := M^{-1}(\overline{w})$. Take a strictly increasing and strictly convex function $W : \mathbb{R}_+ \to \mathbb{R}_+$ that satisfies $W(\theta) := M^{-1}(\theta)$ for $\theta \in [0, \overline{\theta}]$ and define $F_i(\theta) := G_i(W(\theta))$. Clearly, F_i is a well-defined distribution.

Finally, observe that

$$\int_0^{\overline{\theta}} \theta dF_1(\theta) = \int_0^{\overline{w}} M(w)dG_1(w) = \int_0^{\overline{w}} M(w)dG_2(w) = \int_0^{\overline{\theta}} \theta dF_2(\theta)$$

where the first and last equalities follow from the change of variable $\theta = M(w)$ and the fact that $F_i(M(w)) = G_i(w)$ by construction. Taking $H_i = F_i$ completes the proof.

D. PROOF OF THEOREM 4

We prove the first statement of Theorem 4. The restated version of the theorem follows immediately from the argument below.

(Only if.) Suppose the given wage distributions G_1 and G_2 are induced by model primitives H_1, H_2, F_1, F_2 and W where $|\mathbb{E}_{H_i}[\theta] - \mathbb{E}_{H_j}[\theta]| \leq d$.

Since the wage function is Lipschitz continuous, it is differentiable almost everywhere. Therefore, for each group $i \in \{1, 2\}$, we can write

$$\begin{aligned}\mathbb{E}_{G_i}[w] &= [wG_i(w)]_0^{\overline{w}} - \int_0^{\overline{w}} G_i(w)dw \\ &= \overline{w} - \int_0^{\overline{\theta}} W'(\theta)F_i(\theta)d\theta,\end{aligned}$$

where the second equality follows by a change of variable from w to θ.

In light of Theorem 1, we only need to consider the case where $G_i \succ_1 G_j$ for some $i \in \{1, 2\}$ and

$j \in \{1, 2\} \setminus \{i\}$. In this case, the above equation implies that

$$\begin{aligned}
\mathbb{E}_{G_i}[w] - \mathbb{E}_{G_j}[w] &= \int_0^{\overline{\theta}} W'(\theta)[F_j(\theta) - F_i(\theta)]d\theta \\
&\leq \int_0^{\overline{\theta}} [F_j(\theta) - F_i(\theta)]d\theta \\
&= \mathbb{E}_{F_i}[\theta] - \mathbb{E}_{F_j}[\theta] \\
&= \mathbb{E}_{H_i}[\theta] - \mathbb{E}_{H_j}[\theta] \\
&\leq d,
\end{aligned}$$

where the first inequality follows from the fact that W is 1-Lipschitz and $F_j(\theta) \geq F_i(\theta)$ (since $G_j(\theta) \geq G_i(\theta)$). As required, this shows that two wage distributions (ordered by strict first-order stochastic dominance) are rationalizable (given $\mathcal{H}_{|1-2|\leq d}$, $\mathcal{W}_{L1}$) only if the wage gap is less than d.

(If.) Once again, in light of Theorem 1, we only need to consider the case where $G_i \succ_1 G_j$ for $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$ and the wage gap satisfies $|\mathbb{E}_{G_i}[w] - \mathbb{E}_{G_j}[w]| \leq d$.

These wage distributions are induced by $H_i = F_i = G_i$ along with $W(\theta) = \theta$ and these chosen productivity distributions satisfy $|\mathbb{E}_{H_i}[\theta] - \mathbb{E}_{H_j}[\theta]| = |\mathbb{E}_{G_i}[w] - \mathbb{E}_{G_j}[w]| \leq d$ as required.

E. PROOF OF THEOREM 5

We will construct a wage function such that distribution of posterior estimates obtained satisfies $\mathbb{E}_{F_1}[\theta] > \alpha\mathbb{E}_{F_2}[\theta]$. We then take $H_i = F_i$.

Since $G_1(0) = G_2(0) = 0$ (and cumulative distributions are right continuous), there exists a small enough $\tilde{w} > 0$ such that

$$\frac{1 - G_1(w)}{1 - G_2(w)} > \alpha \qquad \text{for all} \quad w \leq \tilde{w}.$$

This in turn implies that

$$\frac{\int_0^{\tilde{w}}[1 - G_1(w)]dw}{\int_0^{\tilde{w}}[1 - G_2(w)]dw} > \alpha.$$

Now take a $\varepsilon \in (0, 1)$ and consider the following wage function

$$W_\varepsilon(\theta) = \begin{cases} \varepsilon\theta & \text{when } \theta \in [0, \tilde{w}/\varepsilon], \\ \tilde{w} + \left(\theta - \frac{\tilde{w}}{\varepsilon}\right) & \text{when } \theta > \tilde{w}/\varepsilon. \end{cases}$$

Note that, by construction, it is 1-Lipschitz.

With this wage function, the ratios of the mean productivities becomes

$$\begin{aligned}\frac{\mathbb{E}_{F_1}[\theta]}{\mathbb{E}_{F_2}[\theta]} &= \frac{\int_0^{\overline{\theta}} \theta dF_1(\theta)}{\int_0^{\overline{\theta}} \theta dF_2(\theta)} \\ &= \frac{\int_0^{\overline{\theta}} [1 - F_1(\theta)] d\theta}{\int_0^{\overline{\theta}} [1 - F_2(\theta)] d\theta} \\ &= \frac{\frac{1}{\varepsilon} \int_0^{\tilde{w}} [1 - G_1(w)] dw + \int_{\tilde{w}}^{\overline{w}} [1 - G_1(w)] dw}{\frac{1}{\varepsilon} \int_0^{\tilde{w}} [1 - G_2(w)] dw + \int_{\tilde{w}}^{\overline{w}} [1 - G_2(w)] dw} \\ &= \frac{\int_0^{\tilde{w}} [1 - G_1(w)] dw + \varepsilon \int_{\tilde{w}}^{\overline{w}} [1 - G_1(w)] dw}{\int_0^{\tilde{w}} [1 - G_2(w)] dw + \varepsilon \int_{\tilde{w}}^{\overline{w}} [1 - G_2(w)] dw}.\end{aligned}$$

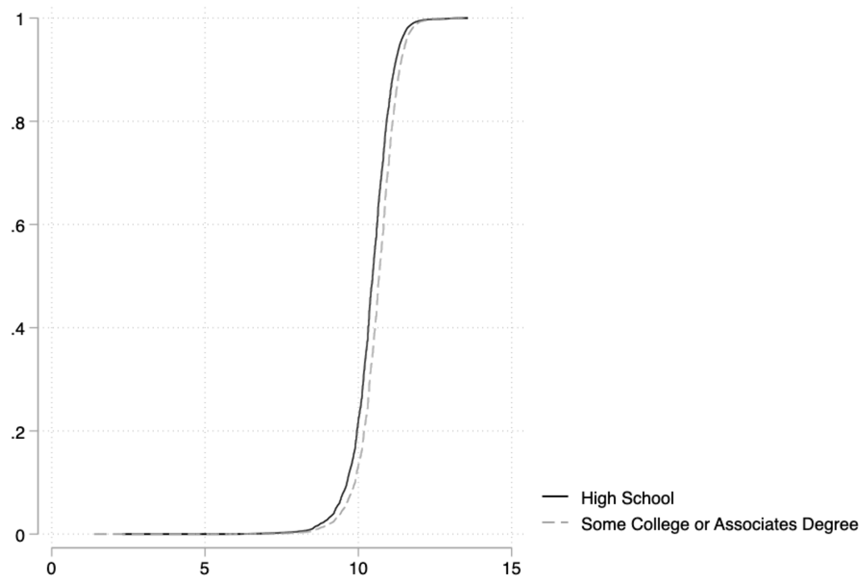
Since $\frac{\int_0^{\tilde{w}} [1 - G_1(w)] dw}{\int_0^{\tilde{w}} [1 - G_2(w)] dw} > \alpha$, we can find small enough ε such that

$$\frac{\int_0^{\tilde{w}} [1 - G_1(w)] dw + \varepsilon \int_{\tilde{w}}^{\overline{w}} [1 - G_1(w)] dw}{\int_0^{\tilde{w}} [1 - G_2(w)] dw + \varepsilon \int_{\tilde{w}}^{\overline{w}} [1 - G_2(w)] dw} = \frac{\mathbb{E}_{F_1}[\theta]}{\mathbb{E}_{F_2}[\theta]} > \alpha$$

which completes the proof because W_ε is the requisite wage function.

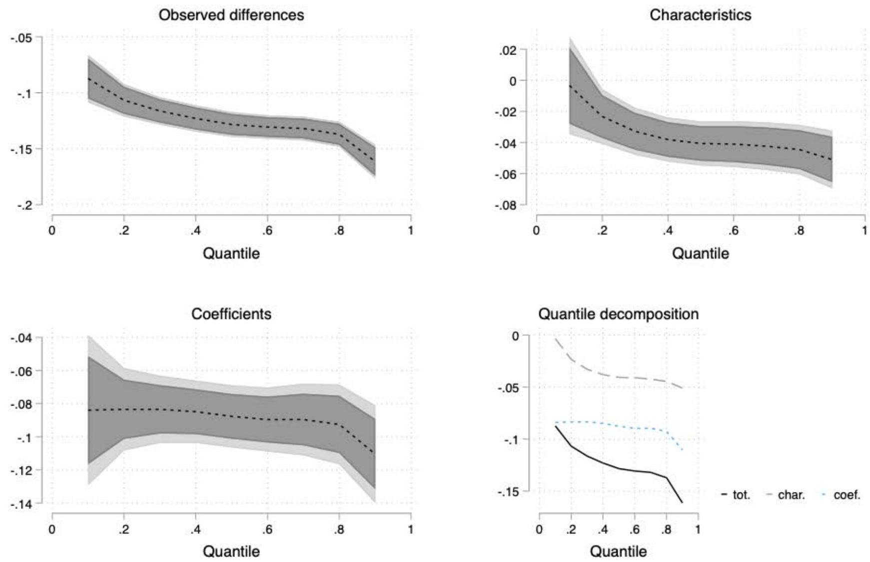
F. WAGE DISTRIBUTION DECOMPOSITIONS OF OTHER OCCUPATION CATEGORIES

Figure 5: Black workers' wages



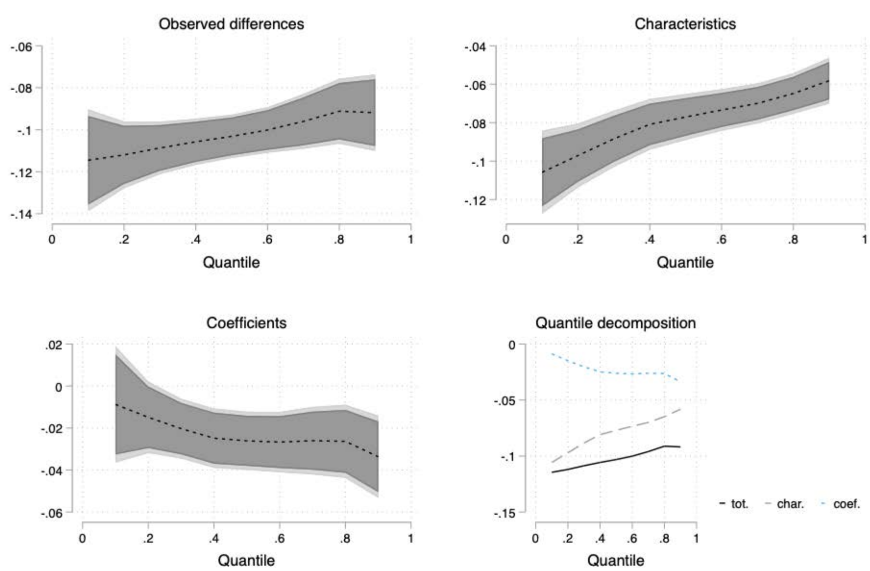
Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages.

Figure 6: Quantile Decompositions: Management, Business, and Financial Occupations



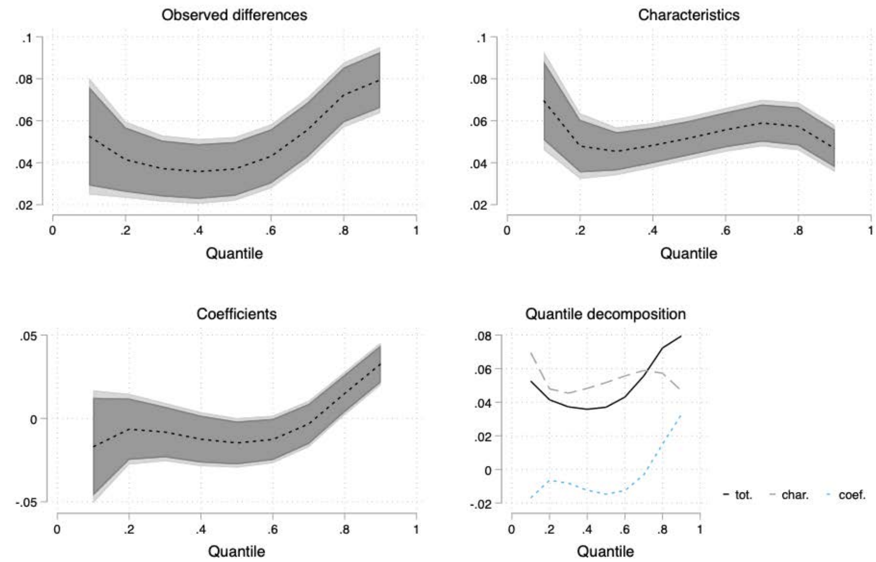
Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

Figure 7: Quantile Decompositions: Professional Occupations



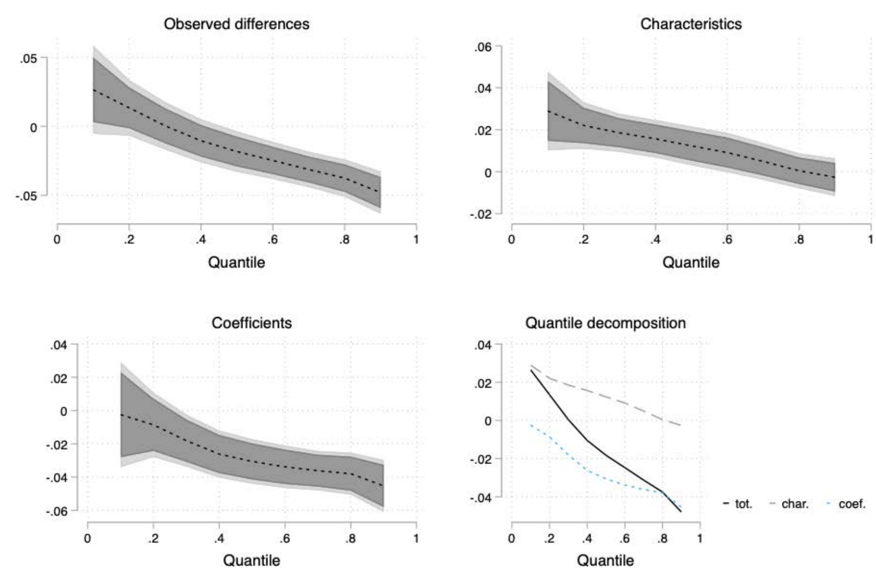
Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

Figure 8: Quantile Decompositions: Service



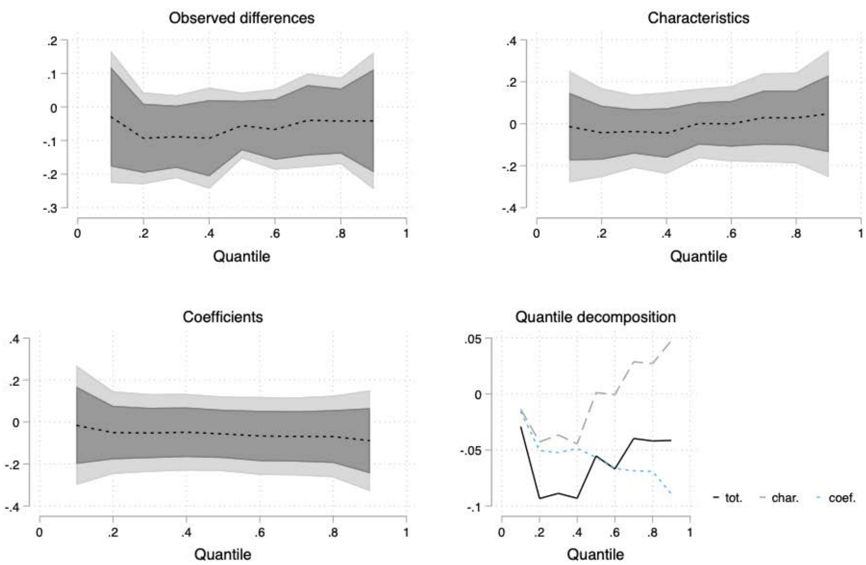
Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

Figure 9: Quantile Decompositions: Office and Admin Support



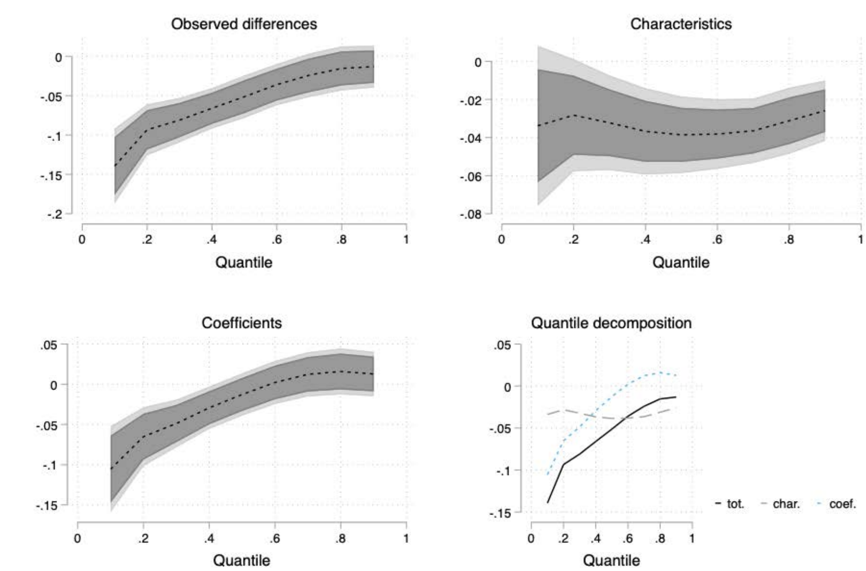
Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

Figure 10: Quantile Decompositions: Farming, Fishing, and Forestry



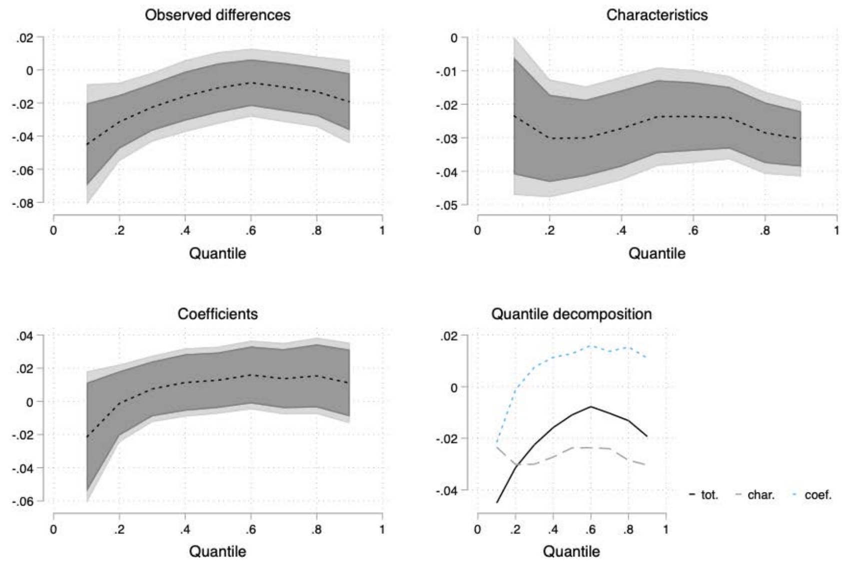
Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

Figure 11: Quantile Decompositions: Construction and Extraction Occupations



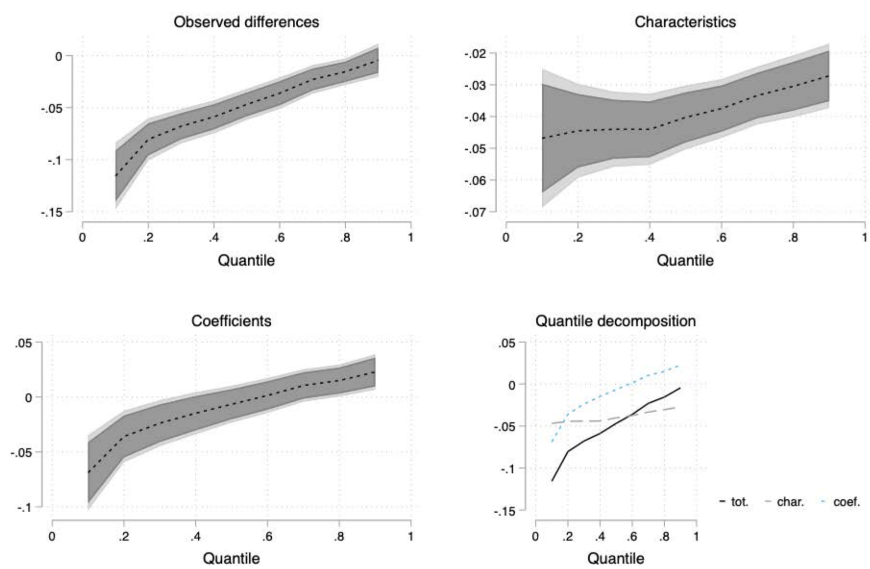
Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, and Census year. 95% confidence bands.

Figure 12: Quantile Decompositions: Installation, Maintenance, and Repair



Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

Figure 13: Quantile Decompositions: Transportation and Material Moving



Data is from the ACS 2007-2021. Men aged 30-55, working full time (52 weeks), and working for wages. White workers have high school degree, Black workers have some college or associate's degrees. CDECO State command based on Chernozhukov, Fernández-Val, and Melly (2013). Controls include age, age-squared, Census region, 2010 occupation codes, and Census year. 95% confidence bands.

G. ONLINE APPENDIX

G.1. HIGHER MOMENTS

In the main text, we assumed that the wage is a strictly increasing function of the posterior mean productivity alone. We now show that if wages were to depend on higher moments of the posterior productivity distribution, then all wage distributions are rationalizable.

Once again, we overload notation but the meaning should be clear. Let $W_\alpha(\mu, \sigma^2) = \mu - \alpha\sigma^2$ with $\alpha \geq 0$ be a wage function that assigns a wage based on both the posterior mean productivity μ and the posterior variance σ^2. Note that W_α is a linear function of the posterior mean and variance, is increasing in the former and decreasing in the latter. We define

$$\mathcal{W}_{\mu,\sigma^2} := \{W_\alpha \mid \alpha \geq 0\}$$

to be the set of all such wage functions. Rationalizability (given $\mathcal{H}_=, \mathcal{W}_{\mu,\sigma^2}$) is defined as before, except it is no longer sufficient to simply consider the distribution of posterior mean productivities F_i (since the variance matters too).

We now show that meaningful testable implications of the model disappear even with this limited class of wage functions that are an affine function of only the posterior mean and variance. This shows that a general (non-Gaussian) version of the model of Aigner and Cain (1977) is not refutable. In the following result, a *nontrivial* wage distribution is one that does not have mass 1 on the highest wage $w = \overline{w}$.

THEOREM 6. *Every pair of nontrivial wage distributions* (G_1, G_2) *is rationalizable (given* $\mathcal{H}_=$, $\mathcal{W}_{\mu,\sigma^2}$*).*

PROOF. If neither G_1 nor G_2 strictly first-order stochastically dominates the other, then the result follows from Theorem 1. So, assume that $G_i \succ_1 G_j$ for $i \neq j$ and hence $\mathbb{E}_{G_i}[w] > \mathbb{E}_{G_j}[w]$.

To ease notation, we normalize the wages to be distributed in $[0, 1]$. (This is without loss of generality since we can consider the re-scaled values $\widehat{w} := w/\overline{w}$ and the re-scaled distribution $\widehat{G}_k$, given by $\widehat{G}_k(\widehat{w}) = G_k(\overline{w} \times \widehat{w})$ for $k \in \{1, 2\}$.)

In what follows, we construct a wage function $W_\alpha \in \mathcal{W}_{\mu,\sigma^2}$ that rationalizes the wage distributions.

Take an $\alpha \geq 0$. For all $w \in [0, 1]$, let $\theta^\alpha(w) \in [0, 1]$ be the greatest solution to

$$w = \theta - \alpha\theta(1 - \theta).$$

Note that the right-hand side of the above equation is continuous in θ, and takes values 0 and 1 at $\theta = 0$

and $\theta = 1$ respectively. Thus, a greatest solution $\theta^\alpha(w) \in [0,1]$ exists. Moreover, except when $w = 0$, the above equation has a unique solution since the right side is strictly increasing whenever its value is strictly greater than 0. Then, note that $\theta^\alpha(w)$ is strictly increasing in w. Finally, observe that $\theta^\alpha(w) \geq w$ and $\lim_{\alpha\to\infty} \theta^\alpha(w) = 1$.

Let $\Theta = [0,1]$. Define the prior distribution H_j^α of group j to be the Bernoulli distribution (so supported on $\{0,1\}$) such that the probability of $[\theta = 1]$ is $\int_0^1 \theta^\alpha(w)dG_j(w)$. Consider the signal (S_j, π_j) where $S_j = [0,1]$ and π_j is a joint distribution over $\Theta \times S_j$ such that

(i) the marginal distribution of π_j over Θ is the prior H_j^α,

(ii) the marginal distribution over S_j is G_j,

(iii) the posterior distribution upon observing any $s_j \in [0,1]$ is a Bernoulli distribution with the probability of $[\theta = 1]$ being $\theta^\alpha(s_j)$.

Observe that this is a valid signal because the average of the posterior distributions $\int_0^1 \theta^\alpha(s_j)dG_j(s_j)$ assigns the same probability to $[\theta = 1]$ as the prior. Observe also that, by construction, the prior distribution H_j^α, the signal (S_j, π_j) and the wage function W_α induce the wage distribution G_j.

For group i, we assume that the signal is perfectly informative and that the prior distribution H_i is G_i. Note that since the experiment is perfectly informative, the posterior variance is always zero so the wage at any signal realization is simply the posterior estimate. Clearly, the prior H_i, a perfectly informative signal (S_j, π_j) and the wage function W_α induce the wage distribution G_i.

It remains to argue that we can choose α such that

$$\mathbb{E}_{H_j^\alpha}[\theta] = \mathbb{E}_{H_i}[\theta].$$

By construction, we have:

$$\begin{aligned}
\mathbb{E}_{H_j^\alpha}[\theta] &= \int \theta^\alpha(w)dG_j(w) \geq \mathbb{E}_{G_j}[w], \\
\lim_{\alpha\to+\infty} \mathbb{E}_{H_j^\alpha}[\theta] &= 1, \\
\mathbb{E}_{H_i}[\theta] &= \mathbb{E}_{G_i}[w] > \mathbb{E}_{G_j}[w].
\end{aligned}$$

Since $\mathbb{E}_{G_i}[w] < 1$ (as distribution G_i is nontrivial), there exists a requisite $\alpha \geq 0$. This completes the proof. ■

We end this section by noting that it is possible to show an analogous result if we instead assume $\alpha \leq 0$.

Our choice of $\alpha \geq 0$ is motivated by Aigner and Cain (1977) and is natural if employers are risk averse. If we instead allow $\alpha \in \mathbb{R}$ to be either positive or negative, then all distributions (including those that are not nontrivial) can be rationalized with equal mean productivities.

G.2. MULTIPLE GROUPS

In the main text, we have restricted our attention to two groups. The purpose of this section is to extend Theorem 1 to n groups. To ease the arguments in this and the next section, we assume that all wage distributions are finitely supported. Without loss of generality, we assume that the support is included in $\mathbf{W} = \{0, 1, 2, \ldots, L\}$. (In empirical applications, wages are supported on finitely many rationals and we can always redefine them to be supported on finitely many integers.)

Let $I = \{1, \ldots, n\}$ be the set of groups. We use $g_{i,\ell}$ to denote the probability that group i receives wage ℓ and $G_{i,\ell}$ to denote the cumulative probability, that is, $G_{i,\ell}$ is the probability that wages are less than or equal to ℓ.

We once again assume that all groups have the same mean productivity and we denote this using the same notation

$$\mathcal{H}_= := \{(H_1, H_2, \ldots, H_n) \mid \mathbb{E}_{H_1}[\theta_1] = \mathbb{E}_{H_2}[\theta_2] = \cdots = \mathbb{E}_{H_2}[\theta_n]\}.$$

As with Theorem 1, we consider unrestricted wage functions in the set $\mathcal{W}$.

Theorem 1 generalizes to $n \geq 2$ groups as follows.

THEOREM 7. *The wage distributions $(G_1, \ldots, G_n)$ supported on a subset of $\mathbf{W}$ are rationalizable (given $\mathcal{H}_=, \mathcal{W}$) if, and only if, there do not exist nonempty subsets $I \cap I' = \emptyset$, $I \cup I' = I$ and weights $\beta \in \Delta^{|I|-1}$, $\beta' \in \Delta^{|I'|-1}$ such that*

$$\sum_{i \in I} \beta_i G_i \succ_1 \sum_{i' \in I'} \beta_{i'} G_{i'}.$$

In words, this condition states that we cannot find two subsets of the groups and convex combinations of their wage distributions that are ordered by strict first-order stochastic dominance.

PROOF. Since we have assumed wage distributions are discrete, we can rewrite the definition of rationalizability in terms of a system of equalities and inequalities. We then use a version of the Farkas lemma to take the alternative and the latter yields the requisite characterization.

Given any strictly increasing wage function $W \in \mathcal{W}$, there exists a unique θ_ℓ such that $W(\theta_\ell) = \ell$. Thus, to rationalize the wage distributions $(G_i)_{i \in I}$, we need to find a vector $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_\ell, \ldots, \theta_L)$

with $0 \leq \theta_\ell < \theta_{\ell+1}$ such that the mean productivities

$$\sum_{\ell=0}^{L} \theta_\ell g_{i,\ell} = \sum_{\ell=0}^{L} \theta_\ell g_{j,\ell}, \tag{2}$$

of all groups $(i, j) \in I \times I$ are equal. In other words, the wage distributions are rationalizable if, and only if, there is a solution to the above equations as then we can choose $W(\theta_\ell) = \ell$, $F_i(\theta_\ell) = G_{i,\ell}$ and $H_i = F_i$ for all $i \in I$ and $0 \leq \ell \leq L$ to rationalize the wage distributions.

We now rewrite rationalizability conditions as a linear system of equalities and inequalities. To do so, we define two matrices. The matrix $\mathbf{A}$ is

$$\mathbf{A} = \begin{bmatrix} g_{1,0} - g_{2,0} & \cdots & g_{1,\ell} - g_{2,\ell} & \cdots & g_{1,L} - g_{2,L} \\ g_{2,0} - g_{3,0} & \cdots & g_{2,\ell} - g_{3,\ell} & \cdots & g_{2,L} - g_{3,L} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ g_{n-1,0} - g_{n,0} & \cdots & g_{n-1,\ell} - g_{n,\ell} & \cdots & g_{n-1,L} - g_{n,L} \end{bmatrix}$$

and the matrix $\mathbf{C}$ is

$$\mathbf{C} = \begin{bmatrix} -1 & 1 & 0 & \dots & \dots & 0 \\ 0 & -1 & 1 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \dots & -1 & 1 \end{bmatrix}.$$

The matrix $\mathbf{A}$ has $n - 1$ rows and $L + 1$ columns. The matrix $\mathbf{C}$ has L rows and $L + 1$ columns. Wage distributions $(G_i)_{i \in I}$ are rationalizable (given $\mathcal{H}_=, \mathcal{W}$) iff we can find a solution to the system

$$\mathbf{A} \cdot \boldsymbol{\theta} = 0 \text{ and } \mathbf{C} \cdot \boldsymbol{\theta} \gg 0.$$

The first constraint captures the equal mean productivity requirement (2) and the second constraint captures the requirement that $\theta_\ell < \theta_{\ell+1}$ for $\ell \in \{0, \dots, L - 1\}$. Note that we do not need to include a constraint that ensures $\boldsymbol{\theta} \geq 0$ since if we find a solution $\boldsymbol{\theta}$, we can also find a positive solution by adding a large enough positive constant to $\boldsymbol{\theta}$.

A general version of the Farkas lemma (see Theorem 1.6.1 in Stoer and Witzgall, 1970) states that either there exists $\boldsymbol{\theta}$ that satisfies $\mathbf{A} \cdot \boldsymbol{\theta} = 0$ and $\mathbf{C} \cdot \boldsymbol{\theta} \gg 0$ or there exists a solution $(\boldsymbol{\alpha}, \boldsymbol{\gamma})$ to

$$\boldsymbol{\alpha} \cdot \mathbf{A} + \boldsymbol{\gamma} \cdot \mathbf{C} = 0,$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_{n-1})$, $\boldsymbol{\gamma} = (\gamma_0, \ldots, \gamma_{L-1})$ are row vectors of dimension $n-1$, L respectively and $\boldsymbol{\gamma} > 0$ (that is, all coordinates are nonnegative with at least one positive).

The vector equation of the alternative $\boldsymbol{\alpha} \cdot \mathbf{A} + \boldsymbol{\gamma} \cdot \mathbf{C} = 0$ can be written in the scalar form

$$\sum_{i=1}^{n-1} \alpha_i (g_{i,\ell} - g_{i+1,\ell}) + (\gamma_{\ell-1} - \gamma_\ell) = 0,$$

for all $0 \leq \ell \leq L-1$. (With the normalization, $\gamma_{-1} = 0$.)

If we sum up these equations from 0 to $0 \leq \bar{\ell} \leq L-1$, we get

$$\sum_{i=1}^{n-1} \alpha_i (G_{i,\bar{\ell}} - G_{i+1,\bar{\ell}}) + \sum_{\ell=0}^{\bar{\ell}} (\gamma_{\ell-1} - \gamma_\ell) = \sum_{i=1}^{n-1} \alpha_i (G_{i,\bar{\ell}} - G_{i+1,\bar{\ell}}) - \gamma_{\bar{\ell}} = 0. \tag{3}$$

Thus, the equation $\boldsymbol{\alpha} \cdot \mathbf{A} + \boldsymbol{\gamma} \cdot \mathbf{C} = 0$ of the alternative is equivalent to (3) holding for all $0 \leq \bar{\ell} \leq L-1$.

Further, since $\boldsymbol{\gamma} > 0$, a solution to the alternative exists iff there exists $\boldsymbol{\alpha}$ such that

$$\sum_{i=1}^{n-1} \alpha_i (G_{i,\ell} - G_{i+1,\ell}) \geq 0, \tag{4}$$

for all $0 \leq \ell \leq L-1$ with at least one strict inequality. To see this, observe that if the there was an $\boldsymbol{\alpha}$ satisfying the above system of inequalities, we could simply set $\gamma_\ell = \sum_{i=1}^{n-1} \alpha_i (G_{i,\ell} - G_{i+1,\ell})$ and this would satisfy $\boldsymbol{\gamma} > 0$.

For a given ℓ, condition (4) can be rewritten as

$$\sum_{i=1}^{n} (\alpha_i - \alpha_{i-1}) G_{i,\ell} \geq 0, \tag{5}$$

with the convention that $\alpha_0 = \alpha_n = 0$. Thus, a necessary and sufficient condition for wage distributions $(G_1, \ldots, G_n)$ to be rationalizable (given $\mathcal{H}_=, \mathcal{W}$) is that there does not exist $\boldsymbol{\alpha}$ that satisfies (5) for all $0 \leq \ell \leq L-1$ with at least one strict inequality. We now show that this condition is equivalent to the condition in the statement of the theorem.

First, suppose the wage distributions $(G_1, \ldots, G_n)$ cannot be rationalizable (given $\mathcal{H}_=, \mathcal{W}$). This implies that there exists $\boldsymbol{\alpha}$ that satisfies (5) for all $0 \leq \ell \leq L-1$ with at least one strict inequality. Now since $\sum_{i=1}^{n} (\alpha_i - \alpha_{i-1}) = \alpha_n - \alpha_0 = 0$, the $\alpha_i - \alpha_{i-1}$ terms cannot all be negative or positive and moreover, they cannot all be zero since condition (4) has to hold strictly for at least one ℓ.

Without loss, let the groups be labeled such that there is a $1 < \bar{n} < n$ for which $\alpha_i - \alpha_{i-1} \geq 0$ for $i \leq \bar{n}$ and $\alpha_i - \alpha_{i-1} < 0$ for $i > \bar{n}$. We can rewrite inequality (5) as

$$\sum_{i=1}^{\bar{n}} (\alpha_i - \alpha_{i-1}) G_{i,\ell} \geq \sum_{j=\bar{n}+1}^{n} (\alpha_{j-1} - \alpha_j) G_{j,\ell}. \tag{6}$$

Observe that $0 < \sum_{i=1}^{\bar{n}} (\alpha_i - \alpha_{i-1}) = \alpha_{\bar{n}} = \sum_{j=\bar{n}+1}^{n} (\alpha_{j-1} - \alpha_j)$. Define

$$\beta_i = \begin{cases} (\alpha_i - \alpha_{i-1})/\alpha_{\bar{n}} & \text{if } i \leq \bar{n}, \\ (\alpha_{i-1} - \alpha_i)/\alpha_{\bar{n}} & \text{if } i > \bar{n} \end{cases}$$

and observe that $(\beta_1, \dots, \beta_{\bar{n}})$ and $(\beta_{\bar{n}+1}, \dots, \beta_n)$ are $\bar{n}$ and $n - \bar{n} + 1$ dimensional probability vectors. Thus, because (6) holds for all $0 \leq \ell \leq L - 1$ with at least one strict inequality, we get that

$$\sum_{i=1}^{\bar{n}} \beta_i G_i \succ_1 \sum_{j=\bar{n}+1}^{n} \beta_j G_j.$$

Conversely, suppose, without loss, that for some $1 < \bar{n} < n$, we have

$$\sum_{i=1}^{\bar{n}} \beta_i G_i \succ_1 \sum_{j=\bar{n}+1}^{n} \beta_j G_j.$$

for some $(\beta_1, \dots, \beta_{\bar{n}}) \in \Delta^{\bar{n}-1}$, $(\beta_{\bar{n}}, \dots, \beta_n) \in \Delta^{n-\bar{n}-1}$. Define

$$\begin{aligned} &\alpha_1 = \beta_1,\ \alpha_2 = \beta_1 + \beta_2,\ \dots,\ \alpha_{\bar{n}} = \beta_1 + \cdots + \beta_{\bar{n}} = 1 \quad \text{and} \\ &\alpha_{\bar{n}+1} = 1 - \beta_{\bar{n}+1},\ \alpha_{\bar{n}+2} = 1 - (\beta_{\bar{n}+1} + \beta_{\bar{n}+2}),\ \dots,\ \alpha_{n-1} = 1 - (\beta_{\bar{n}+1} + \cdots + \beta_{n-1}) = \beta_n \end{aligned}$$

and observe that this implies that (6) holds for all $0 \leq \ell \leq L - 1$ with at least one strict inequality. Thus, these wage distributions cannot be rationalized (given $\mathcal{H}_=, \mathcal{W}$). ■

It is also possible to derive an analogue of Theorem 2. Here, we assume that there are non-empty disjoint subsets $(I_1, \dots, I_m)$ of I with $\cup_{j=1}^{m} I_j = I$ and that the set of productivity distribution is given by

$$\{(H_1, H_2, \dots, H_n) \,|\, \mathbb{E}_{H_j}[\theta_j] \geq \mathbb{E}_{H_{j'}}[\theta_{j'}] \text{ when } j \in I_k,\ j' \in I_{k'},\ k < k' \text{ and } \mathbb{E}_{H_j}[\theta_j] = \mathbb{E}_{H_{j'}}[\theta_{j'}] \text{ when } k = k'\}.$$

We omit the statement for brevity.

G.3. RATIONALIZATION WHEN THE SLOPE OF THE WAGE FUNCTION IS BOUNDED

In the setting of Theorem 5, we restricted attention to 1-Lipschitz wage functions, which imposes an upper bound on the slope of the wage functions. However, no lower bound was imposed. This made it possible for arbitrarily large changes in expected productivities to induce arbitrarily small changes in wages paid – a fact we exploit in the proof. The purpose of this section is to extend the analysis to wage functions, whose slopes take values in an interval $[a,b] \subset \mathbb{R}_{++}$. The lower and upper bounds may be taken from the literature for instance, from Mincerian wage equation estimates. As in the previous section, we assume (to simplify arguments) that all wage distributions are finitely supported on a subset of $\mathbf{W}$.

Specifically, we assume that the wage functions are strictly increasing and Lipschitz continuous with a modulus of continuity in $[a,b]$. We denote the set of such wage functions by $\mathcal{W}_{Lab} \subset \mathcal{W}$. Therefore, wherever differentiable, $W'(\theta) \in [a,b]$ for all $\theta \in \mathbb{R}_+$, $W \in \mathcal{W}_{Lab}$. Throughout this section, we repeatedly use the following two observations.

Observation 1: When $x, x', y, y', \alpha \geq 0$, the fraction

$$\frac{\alpha x + x'}{\alpha y + y'} = \frac{\frac{x}{y}(\alpha y + y') + y'\left(\frac{x'}{y'} - \frac{x}{y}\right)}{\alpha y + y'} = \frac{x}{y} + \frac{y'\left(\frac{x'}{y'} - \frac{x}{y}\right)}{\alpha y + y'}$$

is increasing in α if $\frac{x}{y} > \frac{x'}{y'}$ and decreasing in α if $\frac{x}{y} < \frac{x'}{y'}$.

Observation 2: When $x, x', y, y' \geq 0$, then

$$\frac{x}{y} \geq (\text{resp.}, \leq)\frac{x+x'}{y+y'} \iff \frac{x}{y} \geq (resp., \leq)\frac{x}{y} + \frac{y'\left(\frac{x'}{y'} - \frac{x}{y}\right)}{y+y'} \iff \frac{x}{y} \geq (resp., \leq)\frac{x'}{y'}.$$

We first argue that, because wages are discrete, it is without loss to restrict attention to piecewise linear wage functions.

LEMMA 2. *Consider two wage distributions G_1 and G_2 supported on a subset of $\mathbf{W}$ and a wage function $W \in \mathcal{W}_{Lab}$. Let F_1 and F_2 be the distributions over posterior estimates induced by this wage function, that is, $F_i(\theta) = G_i(W(\theta))$.*

Then, there exists a piecewise-linear wage function $\widehat{W} \in \mathcal{W}_{Lab}$ that has kink points in $\mathbf{W}$ such that this wage function also induces distributions over posterior estimates F_1 and F_2, that is, $F_i(\theta) = G_i(\widehat{W}(\theta))$.

PROOF. Let $\{\theta_0, \ldots, \theta_L\} = \{W^{-1}(0), \ldots, W^{-1}(L)\}$ be the pre-image of the wage function W.

Consider the function

$$\widehat{W}(\theta) = \begin{cases} \frac{\theta-\theta_0}{\theta_1-\theta_0} & \text{if } \theta \in [\theta_0, \theta_1), \\ 1 + \frac{\theta-\theta_1}{\theta_2-\theta_1} & \text{if } \theta \in [\theta_1, \theta_2), \\ \vdots & \vdots \\ L-1+\frac{\theta-\theta_{L-1}}{\theta_L-\theta_{L-1}} & \text{if } \theta \in [\theta_{L-1}, \infty). \end{cases}$$

Note that this function has the feature that $\widehat{W}(\theta_\ell) = \ell$ for all $0 \leq \ell \leq L$.

We now argue that that $\widehat{W} \in \mathcal{W}_{Lab}$. It is clearly continuous by construction. Now, observe that for any $0 \leq \ell \leq L-1$ and $(\theta', \theta) \in [\theta_\ell, \theta_{\ell+1}]^2$ with $\theta' > \theta$, we have

$$\frac{\widehat{W}(\theta') - \widehat{W}(\theta)}{\theta' - \theta} = \frac{\widehat{W}(\theta_{\ell+1}) - \widehat{W}(\theta_\ell)}{\theta_{\ell+1} - \theta_\ell} = \frac{W(\theta_{\ell+1}) - W(\theta_\ell)}{\theta_{\ell+1} - \theta_\ell} \in [a, b]$$

where the first equality follows from the piecewise linearity of $\widehat{W}$ and the value of rightmost expression is in $[a, b]$ because $W \in \mathcal{W}_{Lab}$.

Now take any $\theta < \theta'$ such that $\theta \in [\theta_\ell, \theta_{\ell+1}]$ and $\theta' \in [\theta_{\ell'}, \theta_{\ell'+1}]$ where $0 \leq \ell < \ell' \leq L-1$. We then have

$$\begin{aligned} \frac{\widehat{W}(\theta') - \widehat{W}(\theta)}{\theta' - \theta} &= \frac{\widehat{W}(\theta') - \widehat{W}(\theta_{\ell'}) + \widehat{W}(\theta_{\ell'}) - \widehat{W}(\theta_{\ell'-1}) + \cdots + \widehat{W}(\theta_{\ell+1}) - \widehat{W}(\theta)}{\theta' - \theta_{\ell'} + \theta_{\ell'} - \theta_{\ell'-1} + \cdots + \theta_{\ell+1} - \theta} \\ &\in \left[\min_{\ell \leq k \leq \ell'} \left\{ \frac{\widehat{W}(\theta_{k+1}) - \widehat{W}(\theta_k)}{\theta_{k+1} - \theta_k} \right\}, \max_{\ell \leq k \leq \ell'} \left\{ \frac{\widehat{W}(\theta_{k+1}) - \widehat{W}(\theta_k)}{\theta_{k+1} - \theta_k} \right\} \right] \\ &\subseteq [a, b]. \end{aligned}$$

The second line follows from two facts. The first is that $\frac{\widehat{W}(\theta')-\widehat{W}(\theta_{\ell'})}{\theta'-\theta_{\ell'}} = \frac{\widehat{W}(\theta_{\ell'+1})-\widehat{W}(\theta_{\ell'})}{\theta_{\ell'+1}-\theta_{\ell}}$ and $\frac{\widehat{W}(\theta_{\ell+1})-\widehat{W}(\theta)}{\theta_{\ell+1}-\theta} = \frac{\widehat{W}(\theta_{\ell+1})-\widehat{W}(\theta_\ell)}{\theta_{\ell+1}-\theta_\ell}$ because of the piecewise-linearity of $\widehat{W}$. The second is that the value of the fraction $\frac{x+x'}{y+y'}$ with $x, x', y, y' \geq 0$ lies between min $\left\{\frac{x}{y}, \frac{x'}{y'}\right\}$ and max $\left\{\frac{x}{y}, \frac{x'}{y'}\right\}$ (because of Observation 2).

The proof of the lemma is completed by the observation that the distributions of posterior estimates F_1 and F_2 obtained by inverting G_1 and G_2 are the same under either W or $\widehat{W}$ since, by construction $W(\theta_\ell) = \widehat{W}(\theta_\ell) = \ell$ for all $0 \leq \ell \leq L$. ■

A direct implication of the lemma is that it is without loss of generality to restrict attention to piecewise linear wage functions. Note that, if W is piecewise linear, so it W^{-1}. Moreover, the set of kink points of W^{-1} is a subset of $\mathbf{W}$ and the slope of W^{-1} between any $\ell, \ell+1 \in \mathbf{W}$ (which is simply $W^{-1}(\ell+1) - W^{-1}(\ell)$) lies in the interval $[1/b, 1/a]$.

We use $\mathcal{M}$ to denote the set of strictly increasing, piecewise-linear continuous functions (the notation captures that M is W upside down) whose finite kink points are a subset of **W** and whose slope at every differentiable point lies between $[1/b, 1/a]$. Each $M \in \mathcal{M}$ takes the form

$$M(w) = \left\{ \begin{array}{cc} \alpha_0 w + \kappa & \text{if } 0 \leq w \leq 1, \\ \alpha_1(w-1) + M(1) & \text{if } 1 < w \leq 2, \\ \vdots & \vdots \\ \alpha_{L-1}(w-(L-1)) + M(L-1) & \text{if } w > L-1 \end{array} \right\}$$

where all $\alpha_i \in [1/b, 1/a]$ and $\kappa \geq 0$.

Our aim is to bound the ratio

$$\frac{\mathbb{E}_{H_1}[\theta]}{\mathbb{E}_{H_2}[\theta]} = \frac{\mathbb{E}_{F_1}[\theta]}{\mathbb{E}_{F_2}[\theta]}.$$

For any $M \in \mathcal{M}$, this ratio equals

$$\frac{\mathbb{E}_{F_1}[\theta]}{\mathbb{E}_{F_2}[\theta]} = \frac{\int_0^{M(L)}[1-F_1(\theta)]d\theta}{\int_0^{M(L)}[1-F_2(\theta)]d\theta} = \frac{\int_0^{L}[1-G_1(w)]M'(w)dw}{\int_0^{L}[1-G_2(w)]M'(w)dw} = \frac{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_1(\ell)]}{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_2(\ell)]}.$$

Thus, the bounds for $\frac{\mathbb{E}_{H_1}[\theta]}{\mathbb{E}_{H_2}[\theta]}$ are

$$\left[\min_{\alpha_0,\ldots,\alpha_{L-1}\in\left[\frac{1}{b},\frac{1}{a}\right]} \frac{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_1(\ell)]}{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_2(\ell)]}, \max_{\alpha_0,\ldots,\alpha_{L-1}\in\left[\frac{1}{b},\frac{1}{a}\right]} \frac{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_1(\ell)]}{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_2(\ell)]}\right].$$

In what follows, we describe how to derive the values for these bounds. Specifically, we solve

$$\max_{\alpha_0,\ldots,\alpha_{L-1}\in\left[\frac{1}{b},\frac{1}{a}\right]} \frac{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_1(\ell)]}{\sum_{\ell=0}^{L-1}\alpha_\ell[1-G_2(\ell)]}. \tag{7}$$

The value of the lower bound can be derived analogously. The maximization problem (1) is essentially a knapsack problem as the following two lemmata show. We use the convention that $\frac{1-\widehat{G}_1(\ell)}{1-\widehat{G}_2(\ell)} = 0$ whenever both the numerator and denominator are 0.

LEMMA 3. *There is a solution* $(\alpha_0^*, \ldots, \alpha_{L-1}^*)$ *to* (7) *where every* $\alpha_\ell^* \in \{\frac{1}{b}, \frac{1}{a}\}$. *In words, the slopes of each linear piece takes either the maximum or minimum possible value.*

PROOF. Suppose the solution has some $\alpha_\ell^* \in \left(\frac{1}{b}, \frac{1}{a}\right)$. Then, from Observation 1, if

$$\frac{1 - G_1(\ell)}{1 - G_2(\ell)} \geq \frac{\sum_{\ell' \neq \ell} \alpha_{\ell'}^* [1 - G_1(\ell')]}{\sum_{\ell' \neq \ell} \alpha_{\ell'}^* [1 - G_2(\ell')]}$$

the value of the objective function in (7) will weakly increase if we set $\alpha_\ell^* = \frac{1}{a}$. Conversely if the above inequality is reversed, the value of the objective function in (7) will weakly increase if we set $\alpha_\ell^* = \frac{1}{b}$. ■

Now, sort the series

$$\left\{ \frac{1 - G_1(0)}{1 - G_2(0)}, \frac{1 - G_1(1)}{1 - G_2(1)}, \ldots, \frac{1 - G_1(L-1)}{1 - G_2(L-1)} \right\}$$

in increasing order and denote this sorted series by

$$\left\{ \frac{1 - \widehat{G}_1(0)}{1 - \widehat{G}_2(0)}, \frac{1 - \widehat{G}_1(1)}{1 - \widehat{G}_2(1)}, \ldots, \frac{1 - \widehat{G}_1(L-1)}{1 - \widehat{G}_2(L-1)} \right\}.$$

In other words, $\frac{1-\widehat{G}_1(\ell)}{1-\widehat{G}_2(\ell)}$ is the $(\ell+1)$th highest value in the series $\left\{ \frac{1-G_1(0)}{1-G_2(0)}, \frac{1-G_1(1)}{1-G_2(1)}, \ldots, \frac{1-G_1(L-1)}{1-G_2(L-1)} \right\}$. Note that the value of the objective in the solution to the problem

$$\max_{\alpha_0, \ldots, \alpha_{L-1} \in \left[\frac{1}{b}, \frac{1}{a}\right]} \frac{\sum_{\ell=0}^{L-1} \alpha_\ell [1 - \widehat{G}_1(\ell)]}{\sum_{\ell=0}^{L-1} \alpha_\ell [1 - \widehat{G}_2(\ell)]} \tag{8}$$

is identical to (7).

LEMMA 4. *There is a solution to* (8) *where there is a cutoff index* $\hat{\ell}$ *such that* $\alpha_\ell^* = \frac{1}{b}$ *for* $\ell \leq \hat{\ell}$ *and* $\alpha_\ell^* = \frac{1}{a}$ *for* $\ell > \hat{\ell}$.

PROOF. From Lemma 3, there is a solution in which every $\alpha_k^* \in \{\frac{1}{b}, \frac{1}{a}\}$ for $0 \leq k \leq L-1$. So suppose, for contradiction, that there exist $\ell < \ell'$ such that $\alpha_\ell^* = \frac{1}{a}$ and $\alpha_{\ell'}^* = \frac{1}{b}$. Clearly, it must be the case that

$$\frac{1 - \widehat{G}_1(\ell)}{1 - \widehat{G}_2(\ell)} \geq \frac{\sum_{k \neq \ell} \alpha_k^* [1 - \widehat{G}_1(k)]}{\sum_{k \neq \ell} \alpha_k^* [1 - \widehat{G}_2(k)]}$$

as, otherwise, Observation 1 implies that $\alpha_\ell^* = \frac{1}{a}$ would not be a solution to (8). It then follows from Observation 2 that:

$$\frac{1 - \widehat{G}_1(\ell)}{1 - \widehat{G}_2(\ell)} \geq \frac{\sum_{k=0}^{L-1} \alpha_k^* [1 - \widehat{G}_1(k)]}{\sum_{k=0}^{L-1} \alpha_k^* [1 - \widehat{G}_2(k)]}.$$

In turn, this implies that

$$\frac{1-\widehat{G}_1(\ell')}{1-\widehat{G}_2(\ell')} \geq \frac{1-\widehat{G}_1(\ell)}{1-\widehat{G}_2(\ell)} \geq \frac{\sum_{k=0}^{L-1}\alpha_k^*[1-\widehat{G}_1(k)]}{\sum_{k=0}^{L-1}\alpha_k^*[1-\widehat{G}_2(k)]} \geq \frac{\sum_{k\neq\ell'}^{L-1}\alpha_k^*[1-\widehat{G}_1(k)]}{\sum_{k\neq\ell'}^{L-1}\alpha_k^*[1-\widehat{G}_2(k)]}$$

where the first inequality is a consequence of the fact that $\frac{1-\widehat{G}_1(\cdot)}{1-\widehat{G}_2(\cdot)}$ is sorted in increasing order and, the second and third inequalities are a consequence of Observation 2.

Finally, the above inequality implies that the value of the objective function in (8) is weakly increasing in $\alpha_{\ell'}$ (fixing all α_k^* with $k \neq \ell'$) so $\alpha_{\ell'}^* = \frac{1}{a}$ must also be a solution to (8).

This in turn implies that there is a solution to (8) with the requisite cutoff $\hat{\ell}$ in the statement of the lemma. ■

The previous two lemmas yield a simple algorithm to determine the solution to (7).

Algorithm:

1. Sort the series $\left\{\frac{1-G_1(0)}{1-G_2(0)}, \frac{1-G_1(1)}{1-G_2(1)}, \ldots, \frac{1-G_1(L-1)}{1-G_2(L-1)}\right\}$ in increasing order and set up optimization problem (8).
2. Evaluate the objective in (8) for all cutoffs $\hat{\ell}$ described in Lemma 3. The cutoff that maximizes the value of the objective yields solution to (8).

This algorithm is computationally inexpensive: the time complexities of the first and second steps are $O(L\log(L))$ and $O(L)$ respectively.

REFERENCES

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

Altonji, J. G. and R. M. Blank (1999): "Race and gender in the labor market," *Handbook of labor economics*, 3, 3143–3259.

Altonji, J. G. and C. R. Pierret (2001): "Employer learning and statistical discrimination," *Quarterly Journal of Economics*, 116, 313–350.

Anwar, S. and H. Fang (2006): "An alternative test of racial prejudice in motor vehicle searches: Theory and evidence," *American Economic Review*, 96, 127–151.

Arnold, D., W. Dobbie, and C. S. Yang (2018): "Racial bias in bail decisions," *Quarterly Journal of Economics*, 133, 1885–1932.

Ashraf, N., O. Bandiera, V. Minni, and V. Quintas-Martinez (2022): "Gender roles and the misallocation of labour across countries," *Unpublished Manuscript*, 2.

Ayalew, S., S. Manian, and K. Sheth (2023): "Discrimination and Access to Capital: Experimental Evidence from Ethiopia," .

Barrett, G. F. and S. G. Donald (2003): "Consistent tests for stochastic dominance," *Econometrica*, 71, 71–104.

Becker, G. S. (1957): *The economics of discrimination*, University of Chicago press.

——— (1993): "Nobel lecture: The economic way of looking at behavior," *Journal of Political Economy*, 101, 385–409.

Bertrand, M. and E. Duflo (2017): "Field experiments on discrimination," *Handbook of Economic Field Experiments*, 1, 309–393.

Blinder, A. S. (1973): "Wage discrimination: reduced form and structural estimates," *Journal of Human resources*, 436–455.

Bohren, J. A., K. Haggag, A. Imas, and D. G. Pope (2023): "Inaccurate statistical discrimination: An identification problem," *Review of Economics and Statistics*, 1–45.

Bohren, J. A., P. Hull, and A. Imas (2022): "Systemic discrimination: Theory and measurement," Tech. rep., National Bureau of Economic Research.

Bohren, J. A., A. Imas, and M. Rosenberg (2019): "The dynamics of discrimination: Theory and evidence," *American economic review*, 109, 3395–3436.

Canay, I. A., M. Mogstad, and J. Mountjoy (2023): "On the use of outcome tests for detecting bias in decision making," *Review of Economic Studies*, forthcoming.

Chambers, C. P. and F. Echenique (2021): "A characterisation of 'Phelpsian' statistical discrimination," *Economic Journal*, 131, 2018–2032.

Chernozhukov, V., I. Fernández-Val, and B. Melly (2013): "Inference on counterfactual distributions," *Econometrica*, 81, 2205–2268.

Coate, S. and G. C. Loury (1993): "Will affirmative-action policies eliminate negative stereotypes?" *American Economic Review*, 1220–1240.

Deb, R. and L. Renou (2022): "Which wage distributions are consistent with statistical discrimination?" *CEPR Discussion Paper No. 17676*.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach," *Econometrica*, 64, 1001–1044.

Guryan, J. and K. K. Charles (2013): "Taste-based or statistical discrimination: The economics of discrimination returns to its roots," *Economic Journal*, 123, F417–F432.

Heckman, J. J. (1998): "Detecting discrimination," *Journal of Economic Perspectives*, 12, 101–116.

Juhn, C., K. M. Murphy, and B. Pierce (1993): "Wage inequality and the rise in returns to skill," *Journal of Political Economy*, 101, 410–442.

Kitagawa, E. M. (1955): "Components of a difference between two rates," *Journal of the American Statistical Association*, 50, 1168–1194.

Knowles, J., N. Persico, and P. Todd (2001): "Racial bias in motor vehicle searches: Theory and evidence," *Journal of Political Economy*, 109, 203–229.

Koenker, R. and G. Bassett (1978): "Regression quantiles," *Econometrica*, 33–50.

McFadden, D. (1989): "Testing for stochastic dominance," in *Studies in the economics of uncertainty: In honor of Josef Hadar*, Springer, 113–134.

Neumark, D. (2012): "Detecting discrimination in audit and correspondence studies," *Journal of Human Resources*, 47, 1128–1157.

Oaxaca, R. (1973): "Male-female wage differentials in urban labor markets," *International Economic Review*, 693–709.

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

Shaked, M. and J. G. Shanthikumar (2007): *Stochastic orders*, Springer.

Stoer, J. and C. Witzgall (1970): "Convex Polyhedra," *Convexity and Optimization in Finite Dimensions I*, 31–81.