



**Barcelona School of Economics**

**Màster Universitari en Economia i Finances**

**“A Short Literature Review on Workplace  
Discrimination”**

Authors:

Maddalena Grignani, Wei-Liang Hsu,  
Jiaxun Liu, Georgii Zhrebilov

Director:

Kadir Ozen

*Date*

July 22, 2022

## **ABSTRACT IN ENGLISH:**

In this literature review, we focus on workplace discrimination with particular attention paid to the hiring process. We classified the hiring process into two stages: CV-screening and interview. Following the theoretical discussion of stereotypes supporting the existence of discrimination, two empirical pieces of research are introduced: the first paper shows that race affects applicants' callback rate during the screening process, while the second study shows that the probability of a female musician being hired increases with the introduction of blind auditions. Lastly, from a sociological aspect, we claim that cultural similarities can affect employers' hiring decisions, from which workplace discrimination might also appear.

## **ABSTRACT IN CATALAN/ SPANISH**

En esta revisió de la literatura, nos enfocamos en la discriminaci3n en el lugar de trabajo con especial atenci3n al proceso de contrataci3n. Clasificamos el proceso de contrataci3n en dos etapas: selecci3n de CV y entrevista. Siguiendo la discusi3n te3rica de los estereotipos que respaldan la existencia de discriminaci3n, se presentan dos investigaciones empíricas: el primer artícolo muestra que la raza influncia la probabilidad de conseguir entrevistas de trabajo, mientras que el segundo estudio muestra que la probabilidad de que una música sea contratada aumenta con la introducci3n de audiciones a ciegas. Por último, desde un aspecto sociol3gico, afirmamos que las similitudes culturales pueden afectar las decisiones de contrataci3n de los empleadores, aumentando la discriminaci3n laboral.

## **KEYWORDS IN ENGLISH:**

stereotypes, workplace discrimination, hiring process

## **KEYWORDS IN CATALAN/ SPANISH:**

estereotipos, discriminaci3n laboral, proceso de contrataci3n



## **MASTER PROJECT**

### **A Short Literature Review on Workplace Discrimination**

Authors: Maddalena Grignani, Wei-Liang Hsu, Jiaxun Liu, Georgii Zherebilov

Program: PhD Track Program

Advisor: Kadir Ozen

Academic Year: 2021-2022

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theoretical Background</b>	<b>3</b>
2.1	How Stereotypes Appear . . . . .	3
2.2	The Model . . . . .	5
<b>3</b>	<b>Empirical Evidence</b>	<b>9</b>
3.1	Proving the Existence and the Extent of Racial Discrimination in Hiring Processes through a Correspondence Study. . . . .	9
3.2	Showing Job-finding Improvement for Females Brought by Fostering Impartiality during Interview. . . . .	13
<b>4</b>	<b>Sociological Perspective: Culture-based Discrimination</b>	<b>17</b>
<b>5</b>	<b>Conclusion</b>	<b>20</b>
	<b>References</b>	<b>22</b>

# Abstract

In this literature review, we focus on workplace discrimination, with special attention paid to the hiring process. Discrimination in hiring practices has been treated extensively from the theoretical and empirical literature (e.g. Neumark, 2018; Bertrand and Duflo, 2017), and we further expanded them with some of the most relevant articles. We classified the hiring process into two stages: CV-screening and interview. Following the theoretical discussion of stereotypes supporting the existence of discrimination (Bordalo et al., 2016), two empirical researches are introduced to present empirical evidence that discrimination exists in the two stages of the hiring process. The first paper shows that race, proxied by ethics-based names, affects applicants' callback rate during the screening process (Bertrand and Mullainathan, 2004), while the second study shows that the probability of a female musician being advanced and finally hired increases with the introduction of a blind audition (Goldin and Rouse, 2000). Lastly, from a sociological aspect we claim that cultural similarities can affect employers' hiring decisions, from which workplace discrimination might also come into existence (Rivera, 2012).

## 1 Introduction

Discrimination in hiring practices has been treated extensively from the literature, both under a theoretical and an empirical point of view.

From a theoretical point of view, current economic theories concerning discrimination can be classified into two main sets: taste-based and statistical discrimination models (Bertrand and Duflo, 2017). Statistical discrimination can be attributed to a prediction problem. Namely, an employer, not being able to perfectly observe a candidate's productivity, relies on the average productivity of their group (which can be determined by their race, sex, age and so on) to make inferences on them. Bordalo et al. (2016) propose a model in which discrimination is a consequence representative heuristics. In conditions of imperfect information, people are evaluated according to their representative types - which are stereotypical and rooted in the dif-

ferences across groups - and such information is over-weighted, causing judgements to be distorted. On the contrary, taste-based discrimination relies on mere preferences - regardless of candidates' objective productivity. In the psychology literature, it has been described as an evolutionary phenomenon or as an unconscious automatic negative response to the exposition of someone coming from a different group.

Statistical discrimination can be inferred if, once having provided the relevant information to assess productivity, the likelihood of a job offer or a call-back, conditionally on a set of such individual relevant characteristics, is not statistically different across groups. If that is the case, discrimination can thus be framed as a problem of asymmetric information and signalling. On the contrary, taste-based discrimination persists even after such relevant information is provided.

From an empirical perspective, Neumark (2018) and Bertrand and Duflo (2017) wrote two recent literature reviews on existing studies over field experiments on discrimination, while Walton et al. (2015) focused on discrimination in organizational settings, providing extensive sources of information. FitzGerald et al. (2019), instead, review several articles with interventions designed to reduce implicit prejudices and stereotypes mainly measured through an Implicit Association Test (IAT).

In this paper, we cover some of the most seminal articles in the literature concerning both aspects. We focus on four key papers: Bordalo et al. (2016), Bertrand and Mullainathan (2004), Goldin and Rouse (2000) and Rivera (2012).

We start by analyzing Bordalo et al. (2016), which provides a theoretical framework of stereotypes to discuss discrimination. This paper builds on the concept of representativeness heuristic, introduced in Tversky and Kahneman (1983). Following this approach, one can see discrimination as a result of stereotypical group assessment, rooted in overweighting the most representative (i.e. the most different) types of a given group. Secondly, we present Bertrand's and Mullainathan's paper (2004), as it proved for the first time the extent of racial discrimination in the labor market

through a randomized experiment, introducing a novel (and currently broadly used) methodology: correspondence. Thirdly, we consider Goldin and Rouse (2000). On one side, this paper uncovers the persistence of gender discrimination even in a field in which employers are extremely qualified and sophisticated such as the one of music auditions. On the other side, authors also implicitly suggest a way to overcome such discrimination: blind auditions. Finally, we adopt a more sociological perspective and discuss Rivera’s work 2012, which further shows how cultural similarities might lead to taste-based discrimination.

## 2 Theoretical Background

### 2.1 How Stereotypes Appear

We start with the paper of Bordalo et al. (2016) which introduces a theoretical framework of stereotypes. They borrow the definition of a stereotype from the Oxford English Dictionary, which defines it as a “widely held but fixed and oversimplified image or idea of a particular type of person or thing”. Stereotypes usually cover large groups such as race (“Asians are good at math”), gender (“Women are emotional”), and so on. Therefore, by oversimplifying and generalizing particular group traits stereotypes may create basis for persistent discrimination.

The authors use the so called “social cognition approach” rooting from social psychology (Schneider, 2005). This approach presents stereotypes as a special case of cognitive schemes which build on intuitive generalizations routinely used in everyday life, saving cognitive resources when making decisions. The main property of such stereotypical schemes comes from Hilton and Von Hippel (1996). They argue that “Stereotypes [...] are localized around group features that are the most distinctive, that provide the greatest differentiation between groups, and that show the least within-group variation.” This definition closely follows the concept of *representativeness* heuristic introduced by Tversky and Kahneman (1983). They say that “an attribute is representative of a class if it is very diagnostic; that is, the relative frequency of this attribute is much higher in that class than in the relevant reference

class.” Gennaioli and Shleifer (2010) formalize this idea by assuming that a type  $t$  is representative for group  $G$  relative to a comparison group  $-G$  if it maximizes the likelihood ratio

$$\frac{P(t|G)}{P(t|-G)}$$

The authors argue that the most likely types come to mind first and are overweighted in judgements. For example, consider the stereotype “Florida residents are elderly”. Figure 1 presents the age distributions for Florida and US. One can see that the distributions are quite similar, yet the elderly group “65+” stands out in terms of the highest likelihood ratio  $\frac{P(t|Florida)}{P(t|US)}$ . Therefore, the 65+ type comes first to mind when thinking about Florida, creating the stereotype that the Floridians are mostly old.

Figure 1: Age Distributions for Florida and US

<i>age</i>	0 – 19	20 – 44	45 – 64	65+
<b>Florida</b>	24.0%	31.7%	27.0%	17.4%
<b>US</b>	26.9%	33.6%	26.4%	13.1%

Source: Bordalo et al. (2016)

What is important, if this hypothesis of overweighting representative types is correct, then subjects do not only mistakenly inflate the probability of a random Floridian being old, but also they overestimate the average age for the whole population of Florida. In this regard, discrimination in the workplace may come from the fact that decision makers inadequately assess not only the probability of facing a certain type in a given group (e.g. what is the probability a candidate is good in quantitative skills given it’s a woman), but also the average type of a given group (e.g. how good are women in math on average). Given this, discrimination may occur both at the screening and at the interviewing stages of hiring process.



## 2.2 The Model

The authors provide the following model of stereotypes. A decision maker (DM) faces a prediction problem, that is to assess a characteristic (type) of some group. Consider a set of types  $T$  and a group  $G$  belonging to an overall population  $\Omega$ . The set of types can be unordered (e.g. occupations) or ordered (e.g. earning levels). There is a frequency distribution  $\pi \in \Delta(T \times \Theta)$  inducing conditional distributions  $P(T = t|G) := \pi_{t,G}$ . The DM's goal is to assess  $\pi_G := (\pi_{t,G})_{t \in T}$ . The main assumption is that the DM retrieves a distorted version of  $\pi_G$  that overweights the most representative types of  $G$  with respect to a comparison group  $-G \subseteq \Theta/G$ . The representatives of type  $t$  for group  $G$  given comparison group  $-G$  is defined as:

$$R(t, G, -G) := \frac{\pi_{t,G}}{\pi_{t,-G}}$$

Denote  $\mathbf{R}(t, G, -G) := (\frac{\pi_{t,G}}{\pi_{t,-G}})_{t \in T}$  and define a distorted probability attached to each  $t \in T$  in  $G$  as:

$$\pi_{t,G}^{st} = \pi_{t,G} \frac{h_t(\mathbf{R}(t, G, -G))}{\sum_{s \in T} \pi_{s,G} h_s(\mathbf{R}(t, G, -G))}$$

Here  $h_t : \mathbb{R}_+^T \rightarrow \mathbb{R}_+$  defines a weighting function such that:

1. it's symmetric with respect to  $s \neq t$
2. it increases in own representativeness and decreases in the representativeness of other types

Then the distribution  $(\pi_{t,G}^{st})_{t \in T}$  is the stereotype for group  $G$ . From this framework then some properties of stereotypes follow. Specifically, the authors analyze when the most representative types are a good fit for the group, that is, when they are modal. Also they assess the overall accuracy of the stereotypical distribution by comparing the stereotypical mean and variance to the true ones.

To answer the first question, define an exemplar for  $G$  given  $-G$  as

$$t^* \in \arg \max_t \frac{\pi_{t,G}}{\pi_{t,-G}}$$

Note that under any specification of  $h_t$  satisfying the given definition, overweighting weakly increases when moving towards the exemplar. For the given definition of  $h_t$  the following proposition arises.

PROPOSITION 1. Suppose the conditional distributions  $\pi_{t,G}$  and  $\pi_{t,-G}$  are not identical. Consider two extreme cases:

1. If  $\forall t, t' \in T \pi_{t,G} > \pi_{t',G} \iff \pi_{t,-G} > \pi_{t',-G}$ , then the modal type is not an exemplar for at least one group.
2. If  $\forall t, t' \in T \pi_{t,G} > \pi_{t',G} \iff \pi_{t,-G} < \pi_{t',-G}$  then for each group the modal type is the exemplar.

That is, when groups have the same likelihood ranking, the most representative type is unlikely (not modal) for at least one group. On the other hand, when the likelihood ranking differs significantly in two groups, then the most representative type should be likely (modal). Therefore, stereotypes are more misleading when the distributions of two comparison groups are similar (e.g. Florida example). Instead, when distributions are very different, stereotypes may be accurate (e.g. stereotype “Swedes are blond” is accurate as it’s a majority trait in comparison to the rest of Europe).

Now, to assess the first two moments of the real and stereotypical distributions, consider two cases. First, assume  $R(t, G, -G)$  is monotone, that is the monotone likelihood ratio property (MLRP) holds for the comparison groups. Then the following proposition is true.

PROPOSITION 2. If MLRP holds (assume w.l.o.g. it is strictly increasing in  $t$ ) then for any  $h_t$  that is not constant in the relevant range:

$$\mathbb{E}^{st}(t|G) > \mathbb{E}(t|G) > \mathbb{E}(t|-G) > \mathbb{E}^{st}(t|-G)$$

That is, under MLRP, when the ratio is increasing (decreasing) the right (left)

tail of the distribution for  $G$  is the most representative and therefore the most overweighted. This creates the corresponding bias of the stereotypical mean in the direction of the exemplar. Now assume that the comparison distributions share the same support and are symmetric around  $\mathbb{E}(t|G) = \mathbb{E}(t|-G) = \mathbb{E}(t)$ . Then the following proposition is true.

PROPOSITION 3. Suppose  $R(t, G, -G)$  is U-shaped in  $t$  around  $\mathbb{E}(t)$ . Then, for any  $h_t$  that is not constant in the relevant range:

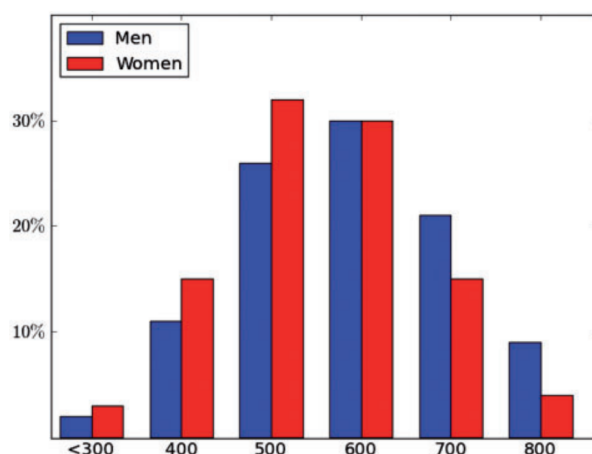
$$\text{Var}^{st}(t|G) > \text{Var}(t|G) > \text{Var}(t|-G) > \text{Var}^{st}(t|-G)$$

$$\mathbb{E}^{st}(t|G) = \mathbb{E}^{st}(t|-G) = \mathbb{E}(t)$$

That is, when the exemplars are located at both tails of the true distribution, by overweighting them, DM inflates the estimated variance, while correctly assessing the mean. For example, if the inflow of immigrants in some country features either relatively unskilled or relatively very skilled workers, one can overexaggerate the variance of immigrants' skills, while having a correct perception about the average level of skill.

To illustrate these properties the authors provide the following example. Many papers provide evidence of a widespread belief that women are worse than men in mathematics (e.g. Guiso et al., 2008). Figure 2 shows the score distributions from the math section of 2013's Scholastic Aptitude Test (SAT) for men and women.

Figure 2: SAT Math Scores for Men and Women, 2013



Source: Bordalo et al. (2016)

As one can see, the distributions are extremely similar (531 points on average for men and 499 for women). Where the distributions differ most is the tails: men are twice as likely to have a perfect score than women and women are almost twice likely to score very low. Hence, in the framework of the presented model the most representative type for men is the one performing very well in the test, while for women the most representative type is the one performing poorly. Therefore, the model predicts exaggerated stereotypical difference between math performance of men and women. Adding to the existing literature on distorted beliefs about quantitative performance of men and women (e.g. Reuben et al., 2014), the model suggests that this pattern comes mainly from the tails of the distribution. Therefore, this example, framed with the presented model, highlights irrational nature of stereotypes: while for most decisions outliers (that is, distribution tails) are irrelevant, it can significantly support stereotypical thinking, “whether through self-stereotyping [...] or through discrimination.”

Concluding, the paper by Bordalo et al. (2016) provides an intuitive theoretical framework for stereotypical thinking, consisting in overweighting the types that differ most relative to a comparison group. This heuristic assessment in many cases gives ground for discrimination in many areas, including workplace discrimination.

## 3 Empirical Evidence

### 3.1 Proving the Existence and the Extent of Racial Discrimination in Hiring Processes through a Correspondence Study.

Bertrand and Mullainathan (2004) aim at assessing whether discrimination in hiring practices exists and, if so, whether there is heterogeneity in terms of gender, industry, quality of the resume, geographic location, and size of the firm, in order to provide some insights on which discrimination model is best supported by evidence.

We believe this paper to be particularly seminal in the discrimination literature for two main reasons: the novel methodology proposed in the study and the consequent evidence-based findings. For what concerns the former, this is the first large-scale correspondence study, which allowed to compare candidates that differed only in their race, keeping “all other things constant”, and in a relatively low-cost and replicable way. Such a novel method allowed the authors to provide scientific evidence that African Americans are less likely to be called back by employers, despite presenting the same quality of resumes as White job seekers. Furthermore, the heterogeneity in resume quality and in the firms’ characteristics allowed to further inquire into the possible causes and types of discrimination. In the following paragraphs we explain more in details the points above.

Isolating the race component in a candidate is challenging, since often many other aspects of a job seeker might affect hiring decisions. Hence, the authors decide to focus on call-back rates from firms who had posted a job vacancy on local news papers in Boston and Chicago requiring only a curriculum vitae, removing all the confounding factors that derive from in-person interactions. A pool of resumes is then created by exploiting real resumes that were sent by job seekers in response to some fictitious job postings published by the authors in local newspapers. The original curriculum vitae (CVs) were slightly modified in order to protect the ap-

plicants' privacy and to remove the components that might suggest the applicants' race, such as the address or the names of the schools attended. Resumes are also divided in two groups: a low quality and a high quality one, with the latter presenting less gaps in employment history, some honors received in school, volunteering or military experience or an email address.

A list of African American-sounding and White-sounding names is created starting from census data and further improved through surveys. Then CVs are then randomly assigned to a name, ensuring comparability across races. Finally, authors send four resumes – a high and a low-quality resume associated with a White sounding and an African American sounding name – to over 1300 job postings found in local newspapers. Since the number of job applications exceeds the number of CVs, the same resume will sometimes be assigned to an African American name and sometimes to a White name, ensuring an even stronger comparability.

Callback rates significantly vary with race: on average, applicants with a White sounding name receive a request for an interview after having sent 10 CVs, while there is a 50% increase in the figure for individuals with an African American sounding name, who have to send about 15 CVs. As shown in table 3, the gap in callback rates is about 3,35 percentage points, or around 50%, and it varies by city, with Boston presenting a larger one.

Table 3. Mean Call-Back Rates By Racial Soundingness of Names

	<i>Call-Back Rate for White Names</i>	<i>Call-Back Rate for African American Names</i>	<i>Ratio</i>	<i>Difference (p-value)</i>
Sample:				
All sent resumes	<b>10.06%</b> [2445]	<b>6.70%</b> [2445]	<b>1.50</b>	<b>3.35%</b> (.0000)
Chicago	<b>8.61%</b> [1359]	<b>5.81%</b> [1359]	<b>1.48</b>	<b>2.80%</b> (.0024)
Boston	<b>11.88%</b> [1086]	<b>7.83%</b> [1086]	<b>1.52</b>	<b>4.05%</b> (.0008)
Females	<b>10.33%</b> [1868]	<b>6.87%</b> [1893]	<b>1.50</b>	<b>3.46%</b> (.0001)
Females in administrative jobs	<b>10.93%</b> [1363]	<b>6.81%</b> [1364]	<b>1.60</b>	<b>4.12%</b> (.0001)
Females in sales jobs	<b>8.71%</b> [505]	<b>6.99%</b> [529]	<b>1.25</b>	<b>1.72%</b> (.1520)
Males	<b>9.19%</b> [577]	<b>6.16%</b> [552]	<b>1.49</b>	<b>3.03%</b> (.0283)

The table reports, for the entire sample and different sub-samples of sent resumes, the call-back rates for applicants with a White sounding name (column 1) and an African American sounding name (column 2), as well as the ratio (column 3) and difference (column 4) of these call-back rates. In brackets in each cell is the number of resumes sent in that cell. Column 4 also reports the p-value for a test of proportion testing the null hypothesis that the call-back rates are equal across racial groups.

Moreover, sending a high-quality resume significantly increases the chances of a White sounding candidate to be called back by 30%, but this is not true for the African American counterpart, suggesting that this racially motivated gap *widens* with the resume quality, making it difficult for the discriminated part to improve their employability. We report the callback rates by racial soundingness and resume quality in table 4.

Table 4. Average Call-Back Rates By Racial Soundingness of Names and Resume Quality

	Low	High	Ratio	Difference (p-value)
White Names	<b>8.80%</b> [1216]	<b>11.31%</b> [1229]	<b>1.29</b>	<b>2.51%</b> (.0391)
African American Names	<b>6.41%</b> [1216]	<b>6.99%</b> [1229]	<b>1.09</b>	<b>0.58%</b> (.5644)

The table reports the mean call-back rates for applicants with a White sounding name (row 1) and African American sounding name (row 2) depending on whether the resume was subjectively qualified as a lower quality (column 1) or higher quality (column 2). In brackets is the number of resumes sent for each race/quality group. Column 4 reports the p-value of a test of proportion testing the null hypothesis that the call-back rates are equal across quality groups within each racial group.

Discrimination occurs in all sectors, regardless of the level of interaction with the customers required by the job, suggesting that discrimination is not motivated by presumed customers' preferences. Moreover, whether employers are federally controlled entities or whether they declare to be Equal Opportunity Employer does not affect the level of discrimination. Finally, the callback gap seems to be lower when firms are located in a neighborhood with a higher percentage of African Americans, which seems to indicate that discrimination can be mitigated by social interactions. This effect is mainly driven by the sample in Chicago, a city that is highly segregated and in which some neighborhoods present a percentage of African American sufficiently relevant to make the detection of some effects possible.

The variety of the above-presented results allow the authors to comment on the relationship between the observed discrimination and the existing theories. Firstly, the callback gap seems unlikely to be motivated by statistical discrimination: indeed, better CVs do not contribute to increase the callback rate, suggesting that providing more information about the job seekers does not correct employers' prior, as we would expect if discrimination was due to a lack of information and lower



expectations in terms of productivity for African American people. However, at the same time, also taste-based models do not apply, as discrimination does not vary across sectors, nor types of employers – such as Equal Opportunity Employers. Authors advance a behavioral theory: employers pay less attention to resumes presenting an African American sounding name, which would explain why only CVs with a White sounding names are ”rewarded” for a higher quality.

It is important to notice that this study also presents some relevant limitations: in the first place, it does not measure the impact of race directly, but of names that suggest belonging to an ethnicity. Authors check whether names that suggest belonging to a low socioeconomic status affect the result, and they do not. Secondly, the outcome is in term of callback rates and not actual hires, so the measure is somehow crude. Lastly, newspapers advertisements represent only one possible channel to find a job, and social networks, which are crucial, are ignored here. Nevertheless, this study provides strong evidence that racial discrimination in hiring practices is indeed happening, regardless of the industry, and that the quality of the resume does not reduce the callback gap.

### **3.2 Showing Job-finding Improvement for Females Brought by Fostering Impartiality during Interview.**

Gender discrimination appears to be one of the most discussed topics within the contemporary context. Bordalo et al. (2019), by conducting laboratory experiments and forming a theoretical model nesting stereotypes and distorted confidence, provided a framework of thinking why women turn out to be less confident in the so-called “male-typed fields”. This inaccurate belief which resembles stereotypes to a large extent lies at the heart of discrimination practices (Bohren et al., 2019).

Empirically, many occupations have been claimed to carry the feature of sex-biased hiring process, even though it appears rather hard to prove its actual existence. Seminal articles dating back to the 1970s discussed the earning difference between

men and women (Becker, 2010; Arrow, 2015), leaving, however, the discussion of actual “bias during hiring” untouched. Although a few audit researches directly showed race-bias (Kenney and Wissoker, 1994) and sex-bias (Neumark et al., 1996) during hiring process, rare has it been for us to see researches addressing bias during hiring with real-life data. At the same time, the hot debate over the cause of occupational segregation by sex went on - “hiring discrimination” and “self-selection” both have their supporting empirical evidence.

Goldin and Rouse (2000) turned our attention to a specific industry: symphony orchestras. The turnover rate of orchestra personnel has been extremely low, but female composition increased from 12% in the 1970s to 35% after the 1980s. Combined with a continuing downward trend in new hires within this industry (regardless of sex) since the 1960s, it implies that the proportion of female new hires must have been extremely high. Within this context, the change in recruiting procedure among major US orchestras throughout the second half of the 20th century then provided an interesting piece of evidence to test for the existence of sex-biased hiring.

The essence of the research relies on its unique data set - exclusive real audition records of eight major US orchestras forming a complete pool of candidates during the research window period, from the late 1950s to 1995. The variation across time, audition rounds and orchestras of the audition data has been exploited to identify the effect of introducing “an extra screen” between audition committees and candidates on the likelihood of a female musician being advanced or finally hired. Meanwhile, there is also the “roster data” containing publicly known year-wise personnel information of eleven major US orchestras, through which the effect of a blind screen on female composition among new-hires has been evaluated.

Table 5. Linear Probability Estimates of the Likelihood of Being Advanced: with Individual Fixed Effects

	Preliminaries							
	Without semifinals		With semifinals		Semifinals		Finals	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Blind	-0.017 (0.039)	0.003 (0.046)	0.109 (0.172)	0.224 (0.242)	0.026 (0.089)	0.102 (0.096)	-0.154 (0.150)	-0.060 (0.149)
Female $\times$ Blind	0.125 (0.068)	0.111 (0.067)	0.013 (0.215)	-0.025 (0.251)	-0.179 (0.126)	-0.235 (0.133)	0.308 (0.196)	0.331 (0.181)
Number of auditions attended		-0.020 (0.014)		0.010 (0.010)		0.015 (0.030)		0.126 (0.028)
Years since last audition		-0.005 (0.007)		-0.006 (0.005)		-0.006 (0.013)		0.016 (0.015)
Automatic placement						-0.096 (0.064)		-0.069 (0.073)
“Big Five” orchestra		-0.154 (0.035)		-0.059 (0.024)		0.006 (0.081)		-0.059 (0.084)
Total number of auditioners in round ( $\div 100$ )		-0.003 (0.081)		0.014 (0.031)		-0.371 (0.521)		-0.262 (0.756)
Proportion female at the audition round		0.118 (0.139)		0.312 (0.134)		0.104 (0.218)		0.067 (0.159)
Principal		-0.079 (0.037)		-0.078 (0.019)		-0.082 (0.066)		-0.185 (0.076)
Substitute		0.165 (0.081)		0.123 (0.093)		0.167 (0.183)		0.079 (0.217)
$p$ -value of $H_0$ : Blind + (Female $\times$ Blind) = 0	0.053	0.063	0.342	0.285	0.089	0.170	0.222	0.042
Year fixed effects?	No	Yes	No	Yes	No	Yes	No	Yes
$R^2$	0.748	0.775	0.687	0.697	0.774	0.794	0.811	0.878
Number of observations	5,395	5,395	6,239	6,239	1,360	1,360	1,127	1,127

The unit of observation is a person-round. The dependent variable is 1 if the individual is advanced to the next round and 0 if not. Standard errors are in parentheses. All specifications include individual fixed effects, an interaction for the sex being missing and a blind audition round, a dummy indicating if years since last audition is missing, and in columns (3)–(8) whether an automatic placement is missing.

On the one hand, they showed, using a probit model with fixed effect framework and the actual audition data focusing on the difference-in-difference estimator (female  $\times$  blind audition), that the introduction of a blind screen increases the probability of a women being advanced to the next round of the audition by 50 percent, and it increases the probability of a women being finally hired by several-fold (shown in table 5). On the other hand, by exploiting year-wise orchestra personnel data with similar probit setting, they also showed that the switch to blind auditions could explain, after adjustment with respect to the current-year female audition participation rate, about 30 percent of the increase in “the proportion of female among new

hires”, and about 25 percent of the overall increase in female composition within an orchestra from 1970 to 1996 (shown in table 6).

Table 6. Probit Estimates of the Effect of Blind Auditions on the Sex of New Members: 1970 to 1996

	Any blind auditions	Only blind preliminaries and/or semifinals vs. completely blind auditions
	(1)	(2)
Any blind auditions	0.238 (0.183) [0.075]	
Only blind preliminaries and/or semifinals		0.232 (0.184) [0.074]
Completely blind auditions		0.361 (0.438) [0.127]
Section:		
Woodwinds	-0.187 (0.114) [-0.058]	-0.188 (0.114) [-0.058]
Brass	-1.239 (0.157) [-0.284]	-1.237 (0.157) [-0.284]
Percussion	-1.162 (0.305) [-0.235]	-1.164 (0.305) [-0.235]
<i>p</i> -value of test: only blind preliminaries and/or semifinals = completely blind		0.756
pseudo $R^2$	0.106	0.106
Number of observations	1,128	1,128

The dependent variable is 1 if the individual is female and 0 if male. Standard errors are in parentheses. All specifications include orchestra fixed effects and orchestra-specific time trends. Changes in probabilities are in brackets. New members are those who enter the orchestra for the first time.

Their findings survived several robustness tests ranging from individual time effect and sample bias to endogeneity (self-selection bias) and sex misclassification bias. Even though some of the estimators have large standard errors, the resulted statistical insignificance cannot obscure the weight of the evidence that the switch to blind auditions did foster impartiality within hiring process, which in turn not only increased the probability of a female being advanced and eventually hired in

an audition but also could help explain the raised female composition in symphony orchestras throughout the second half of the 20th century.

As gender discrimination in general has been paid attention, some corresponding interventions have emerged. Since implications from experimental analyses suggest that there are psychological factors that shape 1. self-confidence within a gender group and 2. discrimination across gender groups, at the same time, mechanisms that aim to reduce self-confidence gap (towards a specific task) between gender groups might work in the direction of alleviating discrimination (Bordalo et al., 2019). A more practical proposal comes from Porter and Serra (2020): role models and manifestations of “similar performance between men and women” in more advanced tasks can have big impacts on the alleviation of the confidence-gap between gender groups, thus reducing the entailed gender discrimination. Nonetheless, apart from suggestions on reducing gender discrimination in general, interventions targeting gender discrimination *in workplace*, or hiring process in particular, still welcome more attention and empirical validation.

## 4 Sociological Perspective: Culture-based Discrimination

Aside from gender or race discrimination, research in the past decades have also shown that culture plays a significant role on labor market stratification. Rivera (2012) first empirically investigated the widely-taken but unexamined hypothesis by past scholars that cultural similarities, including shared tastes, experiences, leisure pursuits, and self-presentation styles (Bourdieu, 1984), between employers and job candidates affect employers’ hiring decisions. Even though scholars often portray the hiring process as the matching between firms’ characteristics, job demands, and candidates’ skills (Tilly and Tilly, 1998), the drivers of employers’ hiring decision have remained mysterious. Thus, the paper investigates the potential factors that determine this decision-making process through conducting qualitative field work in elite professional service firms, focusing the interviewing stage of the hiring process.

This paper critically examines the shortfalls of past research on hiring issues. First, studies on culture and stratification used to have disproportionate focus on examining cultural similarities under education settings (Stevens et al., 2008), which should not be assumed to have analogous impact on workplaces and hiring scenarios. Second, possibly due to data limitation, the focus of hiring studies has been exclusive on racial and sexual similarities, as data for individuals' tastes and experiences can be difficult to collect (Mitchell L. Stevens and Arum, 2008). Furthermore, past studies taking demographic similarities as proxies for cultural similarities overlooked the variations of culture and experiences within the demographic groups (Lamont and Small, 2008), which, due to their relevance, should be taken into account (Turco, 2010; Wilson, 1997). Last but not least, the assumption in some hiring studies that sex and race similarities surmount all other similarities is empirically difficult to sustain since the former have inconsistent effects on hiring processes (Huffcut, 2011). As a result, this paper proposes that investigating shared culture can be influential in labor market sorting, since shared culture is considered as a critical source of perceived similarity (Lamont and Molnár, 2002), which has been thought to be more important than actual similarities in hiring scenarios (Laura M. Graves, 1995).

Elite professional service firms, including law firms, investment banks, and consulting firms, are deliberately chosen to address the following shortcomings in past hiring studies. First, past hiring research mainly focuses on low-wage or low-skilled workers, as hiring process tends to be labor-market specific (Bills, 2003), conducting research on these institutes allow us to make up the deficiency of the lack of understanding in privileged job tracks. Second, most job applicants for these elite firms come from pre-screened pools, enabling us to analyze cultural similarities removing the noise due to candidates' human and social capital differences. Last, research also shows that cultural qualities tend to be more significant among elites (Lamont, 1992), and thus elite professional service firms provide a unique environment to investigate hiring processes from an insightful perspective.

Given the subjective nature of observing perceptions, the methodology implemented in this paper relies on interviews and participant observations. Interviews are conducted with the purpose of understanding which qualities of the interviewees the interviewers look for and which of the mock candidates would be selected by the evaluators who screen résumés. Participant observations are aimed to reveal patterns that the evaluators ignore and record subjective impressions that the evaluators have in interaction with the applicants. Data including the biographic and demographic information and the preferred candidates' qualities of the evaluators are analyzed to reveal its frequencies.

This paper presents evidence through three channels that cultural similarities are valued highly by employers during hiring process. First, the 'organizational processes' channel concerns the relevance of cultural fit, inviting people in charge of screening candidates to pay attention to applicants' performance outside of the workplace, with the goal of ensuring a more enjoyable working environment. The cultural fit highlighted in this channel can be classified into two categories, cultural similarities to firm and cultural similarities to "Self". The former aims at looking for candidates that fit the firms' image, while the latter at searching for those who share similarities with the evaluator representing the firm. The second channel, the 'cognitive processes', shows that experiential similarities are used as frames when evaluators measure the applicants. The main reasoning is that candidates sharing similar experiences with the evaluators are considered more likely to achieve a successful career, as the evaluators themselves. Lastly, the 'affective processes' channel states that cultural similarities affect hiring processes by yielding excitement to evaluators, an ability viewed as a proxy of client skills. The level of excitement perceived through affective processes can also influence evaluators' willingness to advocate for the applicants to receive a job offer among the candidate pool after the interview.

In conclusion, cultural similarities are considered to be with economic conversion value (Bourdieu, 1986), and this paper claims that cultural fit, experiential similarities, and emotional excitement perceived by the evaluators can affect the decision

whether to hire a specific job candidate. In hiring, evaluators not only tend to look for compatible workers but also favor having enjoyable playmates as their colleagues. However, selecting candidates based on shared culture in hiring processes acts as a double-edge sword when considering inequality issues: one side, it challenges sex and race inequalities, but on the other, it may raise new concerns about social mobility. A limitation of this paper is that due to the qualitative design of the research, there is no evidence regarding the extent of the impact of cultural similarities and the potentially following discrimination or inequality issues. Nevertheless, this paper shows the importance of considering shared culture in hiring processes by exploring that alongside the widely-studied effects of gender and racial discrimination on labor market sorting, the inequality caused by cultural similarities between the firms and the applicants should also be noted when studying this issue.

## 5 Conclusion

Discrimination in the workplace as a sub-field of general discrimination has been invoking increasing attention in our contemporary world. As mentioned earlier, workplace discrimination can happen in different stages, including but not limited to hiring process, salary negotiation and promotion throughout employees' careers. In our literature review, focus has been put on the hiring practice, which has been categorized again into two different stages: 1. application stage and 2. interview stage. In particular, the former stage often involves non face-to-face screening according to applicants' CV when the latter usually features in-person interactions.

The discussion on discrimination during hiring process has been carried out around four seminal papers. Starting from a theoretical framework that addresses the potential formation of stereotypes, which has been considered to be one of the major drive for discrimination, (Bordalo et al., 2016), we then move on to two seminal empirical works that address discrimination happening in, as we defined, the two stages of hiring practice.



Firstly, we discussed discrimination in the first stage of hiring process via the seminal work of Bertrand and Mullainathan (2004), which took racial discrimination as the focal aspect. By conducting large-scale correspondence study, they provided scientific evidence that racial discrimination in hiring practices, omnipresent in all sectors, is indeed happening. Moving on to the second stage of interview, Goldin and Rouse (2000) provided a gender-focused research on discrimination in hiring practices. As one of the few researches exploiting real-life data when studying discrimination in hiring process, their study showed that the introduction of blind auditions not only significantly increased the probability of a female musician being finally hired, but also could explain about one third of the increase in female compositions in orchestras during the second half of the 20th century.

Lastly, a sociological research of Rivera (2012) provided a cultural perspective when looking at workplace discrimination. By conducting a qualitative analysis relying on interview and participant observations targeting elite service firms, the study presented evidence that cultural similarities and the invoked inequality, can affect evaluators' final decisions.

Workplace discrimination indeed exists in various dimensions. While what we mainly discussed paid special attention to the hiring process, workplace discrimination in other stages along an employee's career-path is worth further investigation. Meanwhile, most literature has been focusing on providing theoretical framework or empirical evidence for the existence of workplace discrimination and few has proposed potential intervention, or evaluated policy impacts. This suggests that future studies could potentially pay more attention to 1. the linkage between theory and practice: how empirical evidence can help us understand the nature of discrimination, and 2. intervention / policy evaluation: in which way can we alleviate workplace discrimination and to what extent different policies exert their impact.

## References

- Arrow, K. J. (2015). The theory of discrimination. In *Discrimination in Labor Markets*, pp. 1–33. Princeton University Press.
- Becker, G. S. (2010). *The Economics of Discrimination*. University of Chicago press.
- Bertrand, M. and E. Dufló (2017). Field experiments on discrimination. In *Handbook of Economic Field Experiments*, Volume 1, pp. 309–393. Elsevier.
- Bertrand, M. and S. Mullainathan (2004). Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination. *American Economic Review* 94(4), 991–1013.
- Bills, D. B. (2003). Credentials, signals, and screens: Explaining the relationship between schooling and job assignment. *Review of Educational Research* 73, 441–469.
- Bohren, J. A., A. Imas, and M. Rosenberg (2019). The dynamics of discrimination: Theory and evidence. *American Economic Review* 109(10), 3395–3436.
- Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). Stereotypes. *The Quarterly Journal of Economics* 131(4), 1753–1794.
- Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2019). Beliefs about gender. *American Economic Review* 109(3), 739–73.
- Bourdieu, P. (1984). Distinction a social critique of the judgement of taste. In *Inequality Classic Readings in Race, Class, and Gender*, pp. 287–318. Routledge.
- Bourdieu, P. (1986). The forms of capital. In *Handbook of Theory and Research for the Sociology of Education*, pp. 241–258. Greenwood Press.
- FitzGerald, C., A. Martin, D. Berner, and S. Hurts (2019). Interventions designed to reduce implicit prejudices and implicit stereotypes in real world contexts: a systematic review. *BMC psychology* 7(1), 29.

- Gennaioli, N. and A. Shleifer (2010). What comes to mind. *The Quarterly Journal of Economics* 125(4), 1399–1433.
- Goldin, C. and C. Rouse (2000). Orchestrating impartiality: The impact of “blind” auditions on female musicians. *American Economic Review* 90(4), 715–741.
- Guiso, L., F. Monte, S. P., and Z. L (2008). Culture, gender, and math. *Science* 320, 1164–1165.
- Hilton, J. L. and W. Von Hippel (1996). Stereotypes. *Annual review of psychology* 47(1), 237–271.
- Huffcut, A. I. (2011). An empirical review of the employment interview construct literature. *International Journal of Selection and Assessment* 19(1), 62–81.
- Kenney, G. M. and D. A. Wissoker (1994). An analysis of the correlates of discrimination facing young hispanic job-seekers. *American Economic Review* 84(3), 674–683.
- Lamont, M. (1992). *Money, morals, and manners: The culture of the french and the american upper- middle class*. University of Chicago Press.
- Lamont, M. and V. Molnár (2002). The study of boundaries in the social sciences. *Annual review of sociology* 28, 167–195.
- Lamont, M. and M. Small (2008). How culture matters: Enriching our understanding of poverty. In *The Colors of Poverty: Why Racial and Ethnic Disparities Exist*, pp. 76–102. Russell Sage.
- Laura M. Graves, G. N. P. (1995). The effect of sex similarity on recruiters’ evaluations of actual applicants: A test of the similarity-attraction paradigm. *International Journal of Selection and Assessment* 48, 85–98.
- Mitchell L. Stevens, E. A. A. and R. Arum (2008). Sieve, incubator, temple, hub: Empirical and theoretical advances in the sociology of higher education. *Annual Review of Sociology* 34, 127–151.

- Neumark, D. (2018). Experimental research on labor market discrimination. *Journal of Economic Literature* 56(3), 799–866.
- Neumark, D., R. J. Bank, and K. D. Van Nort (1996). Sex discrimination in restaurant hiring: An audit study. *The Quarterly Journal of Economics* 111(3), 915–941.
- Porter, C. and D. Serra (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics* 12(3), 226–54.
- Reuben, E., P. Sapienza, and L. Zingales (2014). How stereotypes impair women’s careers in science. *Proceedings of the National Academy of Sciences* 111(12), 4403–4408.
- Rivera, L. A. (2012). Hiring as cultural matching: The case of elite professional service firms. *American Sociological Review* 77(6), 999–1022.
- Schneider, D. (2005). The psychology of stereotyping.
- Stevens, M. L., E. A. Armstrong, and R. Arum (2008). Sieve, incubator, temple, hub: Empirical and theoretical advances in the sociology of higher education. *Annual Review of Sociology* 34(1), 127–151.
- Tilly, C. and C. Tilly (1998). *Work under capitalism*. Routledge.
- Turco, C. J. (2010). Cultural foundations of tokenism: Evidence from the leveraged buyout industry. *American Sociological Review* 75(6), 894–913.
- Tversky, A. and D. Kahneman (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review* 90(4), 293.
- Walton, G. M., M. C. Murphy, and A. M. Ryan (2015). Stereotype threat in organizations: Implications for equity and performance. *Annu. Rev. Organ. Psychol. Organ. Behav.* 2(1), 523–550.
- Wilson, W. J. (1997). *When work disappears: The world of the new urban poor*. Vintage.