# I Z A Institute of Labor Economics

## DISCUSSION PAPER SERIES

# An Arab, an Asian, and a Black Guy Walk into a Job Interview: Ethnic Stigma in Hiring after Controlling for Social Class

Hannah Van Borm
Louis Lippens
Stijn Baert

# DISCUSSION PAPER SERIES

# An Arab, an Asian, and a Black Guy Walk into a Job Interview: Ethnic Stigma in Hiring after Controlling for Social Class

**Hannah Van Borm**
*Ghent University*

**Stijn Baert**
*Ghent University, University of Antwerp,
Université Catholique de Louvain and IZA*

**Louis Lippens**
*Ghent University*

# ABSTRACT

# An Arab, an Asian, and a Black Guy Walk into a Job Interview: Ethnic Stigma in Hiring after Controlling for Social Class*

Over the last decades, researchers have found compelling evidence of hiring discrimination toward ethnic minorities based on field experiments using fictitious job applications. Despite increasing efforts to discover why ethnic minorities experience hiring penalties, the academic world has not yet found a satisfying answer. With this study, we aim to close this gap in the literature by conducting a state-of-the-art scenario experiment with genuine American recruiters. In the experiment, we ask recruiters to assess fictitious job applicants of various race-ethnicities but consistent social class. The applicants are rated on 22 statements related to the dominant explanations for ethnic discrimination in hiring that the models of taste-based and statistical discrimination have offered. We find that different race-ethnicity groups are evaluated rather similarly, except for Asian Americans, who are perceived to have better intellectual abilities and organizational skills and to be more ambitious, motivated, efficient, and open. These results suggest that the hiring discrimination found in previous experimental research might be overestimated because part of the reported hiring penalty may be attributed to aspects other than race-ethnicity.

**Corresponding author:**
Stijn Baert
Ghent University
Sint-Pietersplein 6
9000 Ghent Belgium
E-mail: stijn.baert@ugent.be

# 1. Introduction

Over the last decades, researchers have found compelling evidence of hiring discrimination toward ethnic minorities.[1] A recent meta-analysis of 143 correspondence experiments on ethnic discrimination conducted between 2005 and 2020 showed that minority applicants receive, overall, 34% fewer positive reactions to their job applications than equally qualified majority candidates (Lippens, Vermeiren, & Baert, 2023).[2] This discrimination, however, varies remarkably by country (Bessudnov & Shcherbak, 2019; Lancee, 2019; Quillian et al., 2019; Zschirnt and Ruedin, 2016) and ethnic group (Booth, Leigh, & Varganova, 2012; Koopmans, Veit, & Yemane, 2019; Lancee, 2019; Thijssen, Coenders, & Lancee, 2020; Yemane, 2020). In the United States, for example, varying levels of hiring discrimination have been found toward African Americans, Hispanic Americans, Arab Americans, and Asian Americans (Boyd-Swan & Herbst, 2019; Darolia, Koedel, Martorell, Wilson, & Perez-Arce, 2016; Decker, Ortiz, Spohn, & Hedberg, 2015; Flake, 2019; Gorsuch & Rho, 2018; Lambert & Akinlade, 2019; Mobasseri, 2019; Yemane, 2020).

In recent years, ever more researchers have attempted to empirically identify different explanations for the varying levels of hiring discrimination toward different ethnic minority groups.[3] Two economic models underpin these attempts: the model of taste-based discrimination (Becker, 1957) and the model of statistical discrimination (Arrow, 1973; Phelps, 1972). Taste-based discrimination (Becker, 1957) implies that hiring discrimination toward ethnic minorities is driven by a distaste that employers, employees, or customers feel toward them. Due to these negative attitudes, employers, employees, and customers experience disutility when interacting with minority employees, which might reduce their hiring chances. Becker's model might also serve as an explanation for different degrees of discrimination that multiple ethnic minority groups experience as employers, employees, and

---

[1] For reviews, see Baert (2018a); Bartkoski, Lynch, Witt, and Rudolph (2018); Bertrand and Duflo (2017); Bonoli and Fossati (2020); Heath and Di Stasio (2019; Lippens et al. (2023); Neumark (2018); Quillian and Midtbøen (2021); Quillian, Pager, Hexel, and Midtbøen (2017); and Zschirnt and Ruedin (2016).

[2] A correspondence test is a type of field experiment that is often used to measure hiring discrimination, although recently it has also been used to measure the impact of characteristics on which unequal treatment in hiring is permitted such as educational outcomes and to measure discrimination in other markets such as the housing market and the dating market (Baert & Verhaest, 2019; Neyt, Baert, and Vynckier, in press). In these tests, researchers send fictitious résumés varying only in terms of a specific characteristic of interest to actual job openings. The callbacks involving these profiles are then examined (Neumark, 2018). Many view correspondence tests as the gold standard for measuring hiring discrimination (Baert, 2018a; Neumark, 2018).

[3] For an overview of the existing literature, see Lippens, Baert, Ghekiere, Verhaeghe, and Derous (2022).

customers might have varying attitudes toward different ethnic groups. Sociological and psychological research has argued that people tend to prefer others belonging to their own in-groups (i.e., the social identity model; Tajfel, 1974) and people similar to themselves (i.e., the model of homophily; McPherson, Smith-Lovin, & Cook, 2001, and the model of homosocial reproduction; Bird, 1996). As a result, individuals who belong to a minority group that is more similar to the majority group in terms of appearance, socioeconomic status, cultural customs or values and norms (i.e., sociocultural closeness; Bogardus, 1925) might be discriminated against less than members of minority groups who differ substantially from the majority group.

Second, the model of statistical discrimination (Arrow, 1973; Phelps, 1972) implies that negative perceptions and stereotypes about ethnic minorities drive their reduced hiring chances. When making hiring decisions, employers typically have limited information. As a result, they will use the available information on the résumé (e.g., race, age, gender, and education) as signals of unobserved factors (e.g., motivation and work ethics). Existing perceptions and stereotypes about ethnicity could signal lower productivity to employers—which might explain their unfair treatment of ethnic minorities. As Katz and Braly (1933) and Lee and Fiske (2006) proved, the content of stereotypes associated with ethnic minorities often differs between minority groups. African Americans, for example, are often perceived as incompetent, untrustworthy, undisciplined, and lacking motivation and work ethic (Ghavami & Peplau, 2012; Federal Glass Ceiling Commission, 1995; Kirschenman & Neckerman, 1991; Moss & Tilly, 2001; Pager & Karafin, 2009; Pedulla, 2018). Asian Americans, on the other hand, are often perceived as highly competent, disciplined, motivated, and hardworking (Ghavami & Peplau, 2012; Federal Glass Ceiling Commission, 1995; Lin, Kwan, Cheung, & Fiske, 2005). According to the model of statistical discrimination, it seems valid to believe that these differences in stereotypes might explain the observed differences in the degree of discrimination experienced by Asian American and African American job seekers, as well as other distinct ethnic groups (Zschirnt & Ruedin, 2016).

Throughout the years, researchers have attempted to gain insights into the empirical validity of these theoretical models as explanations for ethnic discrimination in hiring. More concretely, in addition to measuring hiring discrimination toward ethnic minorities, researchers conducting correspondence tests also tried to gain insights into the underlying mechanisms of this discrimination with secondary analyses to test certain predictions based on the abovementioned theoretical models, such as investigating higher levels of ethnic discrimination in jobs with significant customer contact as indications of taste-based discrimination).[4] Survey research has also been conducted to investigate the potential drivers of

---

[4] Examples of studies that use correspondence tests to investigate taste-based discrimination as an explanation for ethnic discrimination in hiring include Koopmans, Veit, and Yemane (2019); Lancee (2019); Veit and Thijsen (2021); and Zschirnt and

this discrimination (Lippens, Baert, Ghekiere, Verhaeghe, & Derous, 2022).[5]

However, these methods are ideal for investigating the underlying reasons for multiple ethnic minorities' lower—and different—hiring chances. Although correspondence tests are excellent for measuring hiring discrimination, they are less suitable for explaining that discrimination because they can only provide indirect evidence of the mechanisms underlying the phenomenon. Directly surveying the tested employers about their hiring motivations is impossible because, in these experiments, only an employer's binary decision to offer a job interview to an applicant or not can be measured (Van Belle et al., 2020; Van Borm, Burn, & Baert, 2021). Additionally, non-experimental self-reported survey data can never be fully confirmed as free of bias due to the presence of possible confounding variables or socially desirable answering. Therefore, the results found in these studies are only suggestive and should be interpreted carefully.

Despite increasing efforts to understand why different ethnic minority job applicants are (to a lesser or greater extent) penalized when looking for jobs, the academic world does not seem to have found a satisfying and unambiguous answer. However, knowing this information is crucial to developing suitable and effective policy measures to combat this severe and ongoing societal problem.

To gain new insights into the potential explanations for ethnic discrimination in hiring, in this study, we conducted a state-of-the-art scenario experiment based on Baert and De Pauw (2014) and Hosoda, Stone, and Stone-Romero (2003) with genuine recruiters from the United States.[6] In the experiment, we asked our participants to assess a set of fictive résumés that varied in the job applicants' gender, race-ethnicity, and some additional characteristics for a hypothetical vacancy. The applicants' gender and race-ethnicity were signaled to participants via the job candidates' first and last names. The participants evaluated the fictitious job candidates' hireability and 22 statements corresponding to

---

Ruedin (2016). Examples of studies that use correspondence tests to investigate statistical discrimination as an explanation for ethnic discrimination in hiring include Agerström, Björklund, Carlsson, and Rooth (2012); Baert and Vujić (2016); Bartoš, Bauer, Chytilová, and Matějka (2016); Edo, Jacquemet, and Yannelis (2019); Kaas and Manger (2012); Koopmans, Veit, and Yemane (2019); Thijssen, Coenders, and Lancee (2020); Thijssen, Coenders, and Lancee (2021); Thijssen, Lancee, Veit, and Yemane (2019); Veit, Arnu, Di Stasio, Yemane, and Coenders (2022); Yemane and Fernández-Reino (2019); and Zschirnt and Ruedin (2016).

[5] Examples of studies that use non-experimental survey data to investigate explanations for ethnic discrimination in hiring include Aeberhardt, Coudin, and Rathelot (2016); Åslund and Rooth (2005); Combes, Decreuse, Laouénan, and Trannoy (2016); Horvath and Huber (2018); Laouénan (2017); Longley (2003); and Ritter and Taylor (2011).

[6] To the best of our knowledge, Baert and De Pauw (2014) and Hosoda et al. (2003) are the sole experimental studies that use a vignette design to survey employers about the underlying reasons they would hire or not hire ethnic minority job applicants.

stereotypes (in line with the model of statistical discrimination) and attitudes (in line with the model of taste-based discrimination) toward the race-ethnicity minority groups included in our experiment (German Americans, Asian Americans, African Americans, Hispanic Americans, and Arab Americans) and identified in the academic literature.[7] By using candidate names from a name categorization experiment, we ensured that candidates' names signaled the same social class.

We contribute to the existing literature in several meaningful ways. First, we are the first to empirically study the underlying reasons for ethnic discrimination in hiring toward multiple ethnic minorities directly and unambiguously—that is, by using a scenario experiment rather than a correspondence test or survey data. In doing so, we fill a gap in the literature and add to the growing understanding of the reasons underlying ethnic discrimination in hiring.

Second, we add to the literature by including three less-studied race-ethnicity minority groups in the United States (German Americans, Asian Americans, and Arab Americans) in our study. Although German Americans are the largest European immigrant group in the United States and the number of immigrants from Asia and MENA countries (i.e., Middle Eastern and North African countries) has notably increased throughout the years in the United States, most of the studies on this topic that are conducted in the United States investigate hiring discrimination toward African American or Hispanic American job applicants only as these are the two largest race-ethnicity minority groups in the U.S. (Migration Policy Institute, 2021). No study has yet investigated potential hiring discrimination toward German Americans. Furthermore, only three studies have considered Asian Americans and two have considered Arab Americans as study objects (Gorsuch & Rho, 2018; Kang, DeCelles, Tilcsik, & Jun, 2016; Lambert & Akinlade, 2019; Widner & Chicoine, 2011; and Yemane, 2020). By focusing on these three less-studied but growing minority groups (next to African Americans and Hispanic Americans), we gain deeper insights into their hiring experiences in the U.S. labor market.

Third, although this is not the primary goal of our study, we supplement the limited literature on the effect of résumé whitening—i.e., downplaying racial cues in a résumé—on race-ethnicity minority job candidates' hiring chances (Fossati, Liechti, & Auer, 2020; Kang et al., 2016). More concretely, by varying the first-name types (i.e., ethnic- or Anglo-sounding) in the minority job candidates' résumés, we explore whether résumé whitening positively influences employers' perceptions and attitudes toward job applicants, as well as their hiring chances.

---

[7] We must note that, because we conducted a vignette experiment and not a field experiment, we did not measure employers' actual hiring behavior but rather employers' hiring intentions in a simplified simulation of the hiring context.

Fourth, we factored in Gaddis' (2017a, 2017b, 2019) criticism that the names used in previous experiments (i.e., both in the field and in the lab) can present noisy signals for race and may not properly signal the intended race-ethnicity alone, instead signaling differences in social class as well. By using names that are perceived as belonging to people of similar social class, we increase the internal validity of our experiment remarkably over that of previous experimental research (Bertrand and Mullainathan, 2004; Baert and De Pauw, 2014; Leasure and Andersen, 2020; and Hosoda et al., 2003).

Last, we contribute methodologically by dramatically improving Baert and De Pauw's (2014) and Hosoda et al.'s (2003) designs. While both Baert and De Pauw (2014) and Hosoda et al. (2003) only investigated one ethnic group, we use five distinct race-ethnicity minority groups that differ in their sociocultural closeness to the majority group (white Anglo-Saxon Americans) simultaneously. In doing so, we can explain the differences in the degree of discrimination these minority groups experience by looking at how the reasons for discrimination vary between them. Moreover, where Baert and De Pauw (2014) and Hosoda et al. (2003) only considered a limited number—and highly selective set—of possible explanations for ethnic discrimination, we investigate a more systematically acquired and extensive set of stereotypes and attitudes drawn from the relevant literature, considering both the taste-based and statistical discrimination models. Our design is also an improvement on Baert and De Pauw (2014) and Hosoda et al. (2003) in terms of external validity and scale. Both previous studies used a somewhat small sample of students (268 and 298 students, respectively) and focused on a limited number of jobs or internships (one in Baert and De Pauw (2014) and two in Hosoda et al. (2003)). By using a large sample of individuals with genuine hiring experience— 473 participants who evaluated five profiles each, resulting in no less than 2,365 different observations—and ten jobs from distinct economic sectors, we substantially improved the generalizability of our results.


## 2. The Experiment

As mentioned in the introduction, we conducted a state-of-the-art scenario experiment in the United States to explore the empirical value of the theoretical explanations for the varying levels of discrimination experienced by multiple ethnic groups that are outlined by the models of taste-based and statistical discrimination. By using names from a name categorization experiment, we ensured that candidates' names signaled the same social class.

Our experiment applies the factorial survey method developed by Rossi (1951) and often used to study human attitudes, judgments, and behavioral intentions (Auspurg & Hinz, 2014; Rossi & Nock, 1982). This kind of experiment has been used in diverse academic fields, including sociology (Auspurg, Hinz, &

Sauer, 2017; Jasso, 2006; Liebe, Moumouni, Bigler, Ingabire, & Bieri, 2020; McDonald, 2019; Rivera & Tilcsik, 2016; Wallander, 2009), psychology (Derous, Nguyen, & Ryan, 2009; Derous, Ryan, & Nguyen, 2012; Webster, O'Toole, O'Toole, & Lucal, 2005), and economics (Ambuehl & Ockenfels, 2017; Eriksson & Kristensen, 2014; Mathew, 2017), and is extensively conducted to study hiring discrimination and decisions in the labor market (Baert & De Pauw, 2014; Damelang, Abraham, Ebensperger, & Stumpf, 2019; Derous et al., 2009; Derous et al., 2012; Di Stasio, 2014; Hosoda et al., 2003; McDonald, 2019; Van Belle, Di Stasio, Caers, De Couck, & Baert, 2018; Van Belle, Caers, De Couck, Di Stasio, & Baert, 2019; Van Belle et al., 2020; Van Borm et al., 2019; Van Hoye & Lievens, 2003).

In these experiments, participants judge short descriptions of hypothetical situations, objects, or persons that are portrayed in vignettes, for which the attributes (vignette factors) vary systematically or randomly over a predefined number of categories (the vignette levels) (Sauer, Auspurg, Hinz, & Liebig, 2011). By implementing this experimental design within a survey—the main feature of a vignette experiment—the advantages of both experimental and survey research can be combined. That is, by experimentally manipulating the vignette levels, a causal interpretation of the effect of each vignette factor on participants' evaluations can be generated, thus guaranteeing internal validity and avoiding the common issue of confounding variables in survey research. Moreover, by implementing this design with an online survey, higher external validity can be obtained than in typical lab experiments. This is because online surveys can reach relatively large and heterogeneous populations (Auspurg & Hinz, 2014; Damelang & Abraham, 2016; Van Belle et al., 2018; Wallander, 2009).

## 2.1. Vignette Design

In this experiment, the vignettes used were profiles of fictitious job applicants. We asked participants to evaluate a set of five unique vignettes (a deck) comprised of small tables with summarized information about each of the five fictitious job candidates, with one table per vignette. [8,9] The vignettes

---

[8] We chose to present five unique vignettes to each participant based on Auspurg and Hinz's (2014) recommendation to use no more than ten vignettes per participant. Because the participants had to rate many statements for each vignette, we decided to present only five vignettes per participant to avoid excessive survey time, learning effects, and participant fatigue.

[9] As Auspurg and Hinz (2014) recommended, we worked with tabular vignettes instead of text vignettes because these are better suited to decision-making tasks involving lists of decision criteria, such as evaluating fictitious résumés. Tabular vignettes' more straightforward presentation of vignette factors helps participants form consistent judgments, which is especially useful in the context of an experiment where participants evaluate multiple vignettes, such as ours (Auspurg & Hinz, 2014). Furthermore, previous research has shown that tabular vignette designs produce similar evaluations to text vignettes (Auspurg & Hinz, 2014; Shamon, Dülmer, & Giza, 2022; Sauer, Auspurg, & Hinz, 2020).

differed in five applicant characteristics with two to 24 levels each. We present an overview of the different vignette factors and their corresponding levels in Table 1.

<Table 1 about here>

The factor of prime interest in our experiment was applicants' full names, through which we signaled their race-ethnicity and gender. As mentioned in Section 1, we worked with six distinct race-ethnicity groups: one majority group (white Anglo-Saxon Americans) and five minority groups (German Americans, Asian Americans, African Americans, Hispanic Americans, and Arab Americans). We decided to focus on Asian, African, Hispanic, and Arab Americans because these populations are the largest race-ethnicity minority groups in the United States (U.S. Census Bureau, 2021b). We included German American job applicants in our experiment as a kind of control group that was socioculturally close to the majority group. We chose to include German Americans because they are the most populous European immigrant group in the United States (Migration Policy Institute, 2021). Moreover, we believed their first and last names would be distinctive enough to distinguish them from those of white Anglo-Saxon job applicants. We also decided to include both male and female job applicants to increase the ecological validity of our experiment. [10]

As Table 1 shows, we used 24 different name levels, each corresponding to one of the twelve gender-race-ethnicity groups. Fourteen names (seven for each gender group) were associated with majority group candidates (about 60% of the sample) and ten (five for each gender group) with minority group candidates (about 40% of the sample). Half of the names corresponded to male job applicants and the other half to female job applicants. We decided to work with these 24 name levels because we wanted to obtain an applicant pool with a similar composition as the current American population to increase the ecological validity of our experiment. About 60% of people in the United States are 'only white,' and about 50% are female (U.S. Census Bureau, 2021b).

For each of the 24 name levels, we decided to work with more than one name to avoid name effects. Therefore, we used eight full names for the white Anglo-Saxon American male applicants and eight for the white Anglo-Saxon American female candidates. We also used four distinct combinations of first and last names for each of the ten gender-race-ethnicity groups corresponding to one of the five minority groups—for example, we used four names for African American men and four for African American

---

[10] Combining the six race-ethnicity groups with the two gender groups results in twelve distinct gender-race-ethnicity groups: (i) white Anglo-Saxon American men, (ii) white Anglo-Saxon American women, (iii) German American men, (iv) German American women, (v) Asian American men, (vi) Asian American women, (vii) African American men, (viii) African American women, (ix) Hispanic American men, (x) Hispanic American women, (xi) Arab American men, and (xii) Arab American women.

women. To see whether our results would change if we 'whitened' the minority job applicants' names, two of the minority applicants' first names were Anglo-sounding and two were ethnic-sounding. When presenting the vignettes to the participants, one first-and-last-name pair was randomly drawn from the pool of possible name combinations corresponding to the appropriate name level.[11]

A crucial point in our design is that all combinations of first and last names are based on Baert et al.'s (2022) name categorization experiment. They identified 56 combinations of first and last names with variations in perceived race-ethnicity and gender while signaling the same social class (working or middle class). These 56 combinations comprise eight female and eight male 'white'-sounding Anglo-Saxon American first and last names as well as four female and four male names for five ethnic minority groups: German Americans, Asian Americans, African Americans, Hispanic Americans, and Arab Americans. In each gender-race-ethnicity group, two of the name combinations have an Anglo-sounding first name (e.g., Susan Wong), while the other two name combinations have more ethnic-sounding first names (e.g., Mei Lin). We drew the 56 names used in our vignette experiment from Baert et al.'s (2022) Table 2.

As can be seen in Table 1, we also varied the applicants' details in four additional factors: their (i) commuting distance (i.e., the distance the applicants would have to commute between the job and their home; 0–5 miles, 5–10 miles, 10–50 miles, or More than 50 miles), (ii) experience in the occupation (None, About two years, About five years, or About ten years), (iii) recent period of unemployment (Yes or No), and (iv) extracurricular activities (None, Volunteering, Sports activities, or Cultural activities).[12,13] We selected these extra factors and corresponding levels based on (i) a screening of American résumés for elements typically included in these résumés and (ii) findings in the literature (Olian, Schwab, & Haberfeld, 1988; Lahey, 2008; Nuijten, Poell, & Alfes, 2017; Carlsson, Reshid, & Rooth, 2018; Van Belle et al., 2018; Van Belle et al., 2019; Van Borm et al., 2020). To check whether these extra factors were

---

[11] For example, if the job applicant was an African American man, one African American male name was randomly drawn out of the pool of four possible African American male names.

[12] Varying these additional factors has two advantages: (i) we increase the ecological validity of our experiment by mimicking real-life hiring decisions as closely as possible, and (ii) we avoid the influence of social desirability by concealing our primary goal.

[13] The choice to vary the candidate characteristics over five vignette factors was made based on Auspurg and Hinz's (2014) recommendation to work with vignettes of middling complexity, i.e., vignettes in which approximately seven (plus or minus two) vignette dimensions varied. This avoids overburdening participants with a too-complex vignette design and, at the same time, assures enough stimulation to prevent participants from dropping out because of boredom or fatigue, as might occur with an overly simple design (Auspurg & Hinz, 2014).

perceived to be (i) relevant, (ii) realistic, and (iii) informative for employers, we conducted a pilot test of our survey with 40 American Prolific users who were experienced in hiring.[14] We will discuss this pilot study further in Section 2.2. We selected the factors and their levels such that illogical or implausible cases could not occur (Auspurg & Hinz, 2014).

By combining all vignette levels for the five factors (24 × 4 × 4 × 2 × 4), we were able to construct 3,072 unique vignettes (the vignette universe). However, Auspurg and Hinz (2014) recommended that each vignette be evaluated at least five times. In the context of our experiment, this would require either an enormous sample of participants or each participant to assess an immense number of vignettes—which could cause fatigue among the participants and compromise the quality of our data (Auspurg & Hinz, 2014). Therefore, following Auspurg and Hinz (2014), we decided to sample vignettes from the vignette universe (deriving a vignette fraction) using a D-efficient resolution V design. A D-efficient resolution V design selects the vignettes with the most statistical power while ensuring the estimation of all relevant parameters (Auspurg & Hinz, 2014).[15,16,17] This resulted in a more efficient design that required fewer participants or vignette evaluations per participants than a less efficient design would (such as a design obtained through random sampling). In total, we sampled 480 vignettes from the vignette universe using the Kuhfeld (2010) computer algorithm as explained in Auspurg and Hinz (2014). [18] In doing so, we obtained a D-efficient resolution V design with a sufficiently high D-efficiency (92.13) that also allowed

---

[14] Prolific Academic (see Section 3.1. for a more elaborate discussion of Prolific), is one of the many online crowdsourcing platforms on which researchers can recruit participants to perform certain tasks in return for financial compensation (Palan & Schitter, 2018).

[15] Sampling a vignette fraction from the vignette universe typically leads to a loss of information and could cause partial confounding (i.e., the occurrence of moderate or strong correlations between different vignette dimensions; Atzmüller & Steiner, 2010; Auspurg & Hinz, 2014). However, using a D-efficient resolution V design, instead of random sampling, allows us to retain much of the available information and avoid partial confounding.

[16] A D-efficient design enhances statistical precision by maximizing both orthogonality and level balance (i.e., equal frequencies of all levels).

[17] The resolution of a design defines which effects, including possible interaction effects, are identifiable in a given vignette fraction (Dülmer, 2007; Kuhfeld, 1997; Auspurg & Hinz, 2014). Resolution V corresponds to a design where all main effects and two-way interactions are identifiable.

[18] To select the 480 vignettes, we used the freeware macro %Mktex developed by Kuhfeld (2010). The algorithm considers the desired number of vignettes for the experiment, the identified parameters, and the defined number of factors and associated levels to first build a set of potential designs and, subsequently, search for the design with the highest D-efficiency. D-efficiency is sufficiently high when it exceeds 0.90 (Auspurg &Hinz, 2014; Kuhfeld, 1997; Kuhfeld, Tobias, & Garratt, 1994). For more information, see Auspurg and Hinz (2014).

for the estimation of all main effects and two-way interactions. Subsequently, we blocked the 480 vignettes into 96 sets of 5 vignettes (or decks), again using Kuhfeld's (2010) algorithm. To avoid order effects, we randomized the order of the five vignettes within each deck. During the experiment, we randomly assigned one of the 96 decks to each of the participants.

## 2.2. Online Survey

As mentioned, the prime feature of factorial survey applications is the implementation of a multidimensional experimental design within a survey (Auspurg & Hinz, 2014). For this study, we implemented the experimental design described in the previous section in an online survey designed in Qualtrics and administered in English. Next, we describe the different elements of the online survey.

The survey consisted of two distinct parts. The first part of the survey asked twelve questions to gather information from participants regarding several personal characteristics.[19] We first questioned the participants about five demographic characteristics: (i) their gender (male, female, or 'X'), (ii) their age, (iii) the state they lived in, (iv) their highest obtained degree (university education, higher education outside university, secondary education, or lower than secondary education), and (v) their race or ethnicity (white, Hispanic or Latino, Black or African American, Asian, American Indian or Alaska Native, Middle Eastern or North African, Native Hawaiian or other Pacific islanders, or some other race or ethnicity).[20]

We also surveyed participants concerning their tendency to answer in a socially desirable way. To do so, we used the 13-item version of the Marlowe-Crowne Social Desirability Scale (MC-SDS) developed by Reynolds (1982), one of the instruments used most frequently to measure social desirability (Beretvas, Meyers, & Leite, 2002; Sârbescu, Costea, & Rusu, 2011; Baert, 2018b). The scale consists of 13 statements describing culturally sanctioned or approved behavior (e.g., 'There have been times when I was quite jealous of the good fortune of others.'). Participants were asked to indicate 'true' when the statement applied to them or 'false' when it did not. We coded the answers so that socially desirable responses corresponded to a score of 1, and socially undesirable ones corresponded to 0. Cronbach's

---

[19] Before we redirected the participants to the actual survey, we showed them a short introduction in which we explained the research aim and described the task they would have to perform. We described the study's aim broadly and vaguely to obscure the actual goal of the study from participants and avoid socially desirable answering. We also described the estimated survey time, ensured the anonymity and confidentiality of respondents' answers, and highlighted the value of their participation. After the introductory text, participants had to read and sign a consent form to participate in the survey.

[20] In addition to indicating their highest obtained degree, the participants had to check an additional box as an attention check.

alpha for this scale was 0.813 for our sample, which is sufficiently high. Summing the scores for all items yielded a total score between 0 and 13 for answering in a socially desirable manner. As in Van Borm, Burn, and Baert (2021), we divided this number by 13 to obtain a proportion between 0 and 1.

Lastly, we also asked participants to answer six questions about their experience evaluating job applicants. They were asked to indicate (i) whether they had experience in evaluating job candidates in their current or last job (yes or no), (ii) the frequency with which they were, in general, involved in the evaluation of job candidates in their current or last job (daily, weekly, biweekly, monthly, once per semester, once a year, or less frequently), (iii) the frequency with which they were involved in the evaluation of job candidates in the last year (0 times, 1 time, 2 times, 3 times, 4 times, or 5 times or more), (iv) the amount of time (in years) they had been involved in evaluating job candidates (less than one year, 1–5 years, more than 5 years, or not applicable), (v) whether they recruited in a specific sector (yes or no) and if so, which one (construction, technical sector, ICT sector, financial sector, administration, sales, human resources, transport or logistics, or other), and (vi) a description of their current or last occupation (manager, specialist in personnel and career development, employment agency employee, management assistant, general administrative assistant, or other).

The second part of the survey covered the online experiment discussed in Section 2.1. At the beginning of the online experiment, we informed the participants about their task: to evaluate five fictitious job applicants for a particular job vacancy at a hypothetical firm.[21] Subsequently, participants were shown a job opening for one of the following ten positions: (i) sewing machine operator, (ii) usher, (iii) stock clerk, (iv) data entry keyer, (v) farm labor contractor, (vi) judicial law clerk, (vii) gem and diamond worker, (viii) exercise physiologist, (ix) computer hardware engineer, and (x) management analyst. To improve the external validity of our study, we selected jobs from different sectors. We created the job descriptions for the vacancies based on specifications found on O*Net.[22] To avoid any effects of the job specifications on the results, we formulated them as uniformly as possible. We present an overview of the descriptions in Table A–1 in the Appendix. We randomly assigned the job vacancies to participants such that all ten vacancies were displayed with equal probability and did not correlate with the assigned deck of fictitious profiles.

Just after reading the assigned job vacancy, participants had to rate two statements that measured their

---

[21] We decided to ask participants to evaluate job candidates for a hypothetical firm instead of their own firm to ensure the internal validity of our experiment.

[22] O*Net is an online databank developed by the U.S. Department of Labor Employment and Training Administration that summarizes occupational information on thousands of jobs (National Center for O*NET Development, 2019).

experience in hiring job applicants for the presented vacancy and their feeling of competence to perform the task at hand. Participants had to rate both statements on an 11-point Likert scale ranging from 0 (strongly disagree) to 10 (strongly agree).

After participants completed the comprehension check, they were redirected to the scenario sketch constructed analogously to Sterkens, Baert, Rooman, and Derous (2021). In the scenario sketch, we explained to the participants that the selection process for the mentioned vacancy consisted of several rounds and that one of their colleagues had conducted the first interview with a selection of the candidates. We informed them that they had to give a second opinion on five of these candidates based on their colleague's interview notes, which were summarized in small tables presented on the following pages. The interview notes mentioned the candidate characteristics discussed in Section 2.1 and summarized in Table 1. We clarified that all candidates were formally eligible for the vacancy in terms of education level and work experience and highlighted that the mentioned periods of unemployment were between one and six months over the last five years. We also noted that all of the candidates lived in the United States and were American by nationality. We mentioned the latter to ensure that the participants in the experiment knew none of the job applicants were first-generation immigrants. Lastly, we told the participants to rate each candidate accurately and clarified that they could jump between the different candidates and adjust their ratings as desired.

When participants completed reading the scenario sketch, they were shown the tabulated summaries of the fictive job applicants' characteristics.[23] Based on the given information, participants had to assess each job candidate in terms of (i) their intention to invite the job applicant to the second round, (ii) their perceptions of the productivity of each candidate (based on the model of statistical discrimination and measured with 19 items), and (iii) their attitude toward collaborating with each candidate (based on the model of taste-based discrimination and measured with three items).

First, to measure participants' invitation intentions, we adopted one item from Sterkens, Baert, Rooman, and Derous (2021). The participants had to indicate their intention to invite a job candidate to the next round (hereafter called 'the invitation probability') on an 11-point Likert scale ranging from 0 (strongly disagree) to 10 (strongly agree). Second, to survey participants' perceptions of the job applicants, we developed 19 different statements, each corresponding to a particular productivity-

---

[23] We displayed the candidate features in the same order as they would appear on real résumés, which is also the order they are presented in Table 1. Moreover, to ensure that participants considered all elements of the information mentioned in the tables and avoid order effects, we fixed the position of each of the tables on the screen such that they remained visible to the participants while they assessed the candidates.

related stereotype about people who belong to one of the five ethnic minority groups, based on a thorough review of the relevant academic literature.[24] We developed items concerning job applicants' perceived (i) intellectual capacities, (ii) social capacities, (iii) physical capacities, (iv) technological knowledge and skills, (v) language skills, (vi) leadership skills, (vii) assertiveness, (viii) ambition, (ix) motivation, (x) autonomy, (xi) efficiency, (xii) organizational skills, (xiii) reliability, (xiv) trainability, (xv) flexibility, (xvi) accuracy, (xvii) openness to new people and experiences, (xviii) respect toward authority, and (xix) sense of responsibility. Participants had to rate all 19 items on an 11-point Likert scale ranging from 0 (strongly disagree) to 10 (strongly agree). Lastly, we included three items adapted from Van Borm, Burn, and Baert (2021) to measure potential taste-based discrimination induced by the employer, employees, or customers. Participants had to indicate their own, their employees', and their customers' willingness to collaborate with a particular candidate using the same 11-point Likert scale as before. We present an overview of the different measures and corresponding statements in Table 2.[25]

<Table 2 about here>

After the participants evaluated all five candidates, we presented them with a manipulation check to investigate whether they had received the intended experimental treatment (i.e., that they noticed the name of each job applicant and associated them with the intended gender and race). For the manipulation check, participants first had to select their assigned job candidate's name from a list with two possible options. Subsequently, they had to indicate the (i) gender (male or female), (ii) social class (low class, working class, middle class, or high class), and (iii) race-ethnicity group (white American, African American, Hispanic or Latino American, Arab American, Asian American, German American, or none of the above) with which they associated that name. At the end of the manipulation check, they also had to indicate on an 11-point Likert scale going from 0 (strongly disagree) to 10 (strongly agree) whether they found it hard to remember the names of the job applicants when asked.

As mentioned in Section 2.1, we conducted a pilot test on Prolific Academic (hereafter called 'Prolific') with 40 American Prolific users who were experienced in evaluating job applicants to check the overall

---

[24] Although, in this study, we focused on several individual stereotypes, Cuddy, Fiske, and Glick (2007) offer the stereotype content model, a framework that proposes potentially universal principles of societal stereotypes. The stereotype content model defines two universal dimensions of social perception, competence and warmth, to which all existing stereotypes regarding particular ethnic groups can be reduced. For research on the stereotype content model applied to ethnic minority groups, see Lee and Fiske (2006).

[25] To avoid order effects, we randomized the order of display of the 22 different statements (i.e., 19 related to statistical discrimination and three to taste-based discrimination) between the participants. However, within each deck, the order of display of the 22 statements was kept fixed.

quality of our online survey. The participants first had to complete the entire online survey. After completing the survey, they filled out an evaluation form in which they were asked to rate the clarity, relevance, ecological validity, and informativity of the assigned profiles, the presented job vacancy, and the overall survey. The results of this pilot with 200 (40 × 5) candidate evaluations (available upon request) indicated that the quality of the fictitious profiles, the fictive job vacancies, and the overall survey was sufficiently high.

# 3. Data

## 3.1. Data Collection

The online survey described in Section 2.2 was offered to participants via the online platform Prolific Academic. As previously mentioned, Prolific is one of the many online crowdsourcing platforms on which scholars can recruit participants to perform particular tasks in return for financial compensation (Palan & Schitter, 2018). Prolific is particularly attractive because it is specifically designed for researchers and their needs and has been found especially suitable for social and economic science experiments (Palan & Schitter, 2018; Peer, Brandimarte, Samat, & Acquisti, 2017). Moreover, the data quality derived from Prolific has been found to be superior to that from other platforms, such as the widely used platform Amazon Mechanical Turk (Peer et al., 2017; Peer, Rothschild, Evernden, Gordon, & Damer, 2021). Unsurprisingly, a vast number of scholars from various disciplines, such as economics (Frimpong, Shuridah, Wilson, & Sarpong, 2020; Lucas, Berry, Giurge, & Chugh, 2021; Marreiros, Tonin, Vlassopoulos, & Schraefel, 2017; Singh, Crisafulli, Quamina, & Kottasz, 2020), psychology (Anwyl-Irvine, Massonnié, Flitton, Kirkhal, & Evershed, 2020; Callan, Kim, Gheorghiu, & Matthews, 2017; Costin & Vignoles, 2020; Dutt & Kohfeldt, 2019; Roster & Ferrari, 2020), and political sciences (Kaufmann, 2019; Peitz, Dhont, & Seyd, 2018), have used the platform to recruit participants for their studies.

To participate in our study, participants had to meet two criteria. First, they had to be born and currently reside in the United States. Second, they had to have experience in evaluating job applicants. To reach our target group of participants, we employed the pre-screening filters offered by Prolific. Only people who met these two screening criteria were invited to participate in the study. At the beginning of the survey, we included four screening questions that were identical to the pre-screening filters to confirm

that the filters worked properly and were still up to date.[26] Participants who provided answers to the screening questions that deviated from their answers to the pre-screening filters were redirected out of the survey.

During data collection, we subjected all completed surveys to quality control, taking into account (i) the quality of the answers, (ii) the time spent taking the survey, (iii) whether the respondent passed the attention check (see footnote 20), and (iv) whether the respondent was able to provide the correct completion code.[27] Only participants who provided high-quality data were included in our final sample and compensated for participating in the study. Between March 18 and May 7, 2021, 648 participants filled in the survey completely and accurately, resulting in a preliminary sample of 3,240 observations because each participant had to assess five fictitious job applicants.

However, this preliminary sample also contained participants who did not receive the experimental stimuli correctly. That is, not all participants in this sample categorized the job applicants in the intended gender, race-ethnicity, or social class in the manipulation check (see Section 2.2). To ensure that our results would not be biased by noisy name signals, we restricted our sample by considering only those observations from participants who categorized the fictitious job applicants correctly in the manipulation check. Our final sample consisted of 2,362 observations. We give a short description of our data in the next section.

## 3.2. Data Description

In Table 3, we present descriptive statistics regarding various participant characteristics (Panel A) and the ten different job vacancies (Panel B) for both our total sample (column 1) and six subsamples of participants: those who evaluated (i) white Anglo-Saxon Americans, (ii) German Americans, (iii) Asian Americans, (iv) African Americans, (v) Hispanic Americans, and (vi) Arab Americans (columns 2 to 7). We ran multiple one-way ANOVA estimations to check whether the randomization of the job applicants' race-ethnicity over the participants had gone as planned and the job vacancies were evaluated with about the same frequency for each race-ethnicity group of job applicants. We present the results of

---

[26] The screening questions consisted of answering 'United States' to the questions (i) 'What is your nationality?', (ii) 'What is your country of birth?', and (iii) 'In what country do you currently reside?' and answering 'yes' to the question, 'Do you have any experience in making hiring decisions (i.e., have you been responsible for hiring job candidates)?'

[27] When participants finished the survey, they were assigned a completion code to submit to Prolific. Only when participants submitted the correct completion code were they included in the final sample and paid the predetermined financial compensation.

these estimations in column 8 of Table 3.

<center>&lt;Table 3 about here&gt;</center>

From the descriptive statistics in Table 3, it becomes clear that our total sample matches the intended target population of American people with experience evaluating job applicants (Panel A, column 1). Almost 70% of the participants in our sample assess job applicants at least once per semester in their current jobs, about half of them have been evaluating job candidates for more than five years, and 92.20% have been assessing job applicants for at least one year. Moreover, our sample of participants is 47.40% female, 58.10% younger than 40, and 81.80% white Anglo-Saxon Americans. Lastly, the majority of our participants have a university degree (75.90%) and are full-time employed (78.90%). Panel B of column 1 shows that participants were assigned each of the ten job vacancies with about the same frequency.

Table 3 (Panel A, columns 2 to 8) also shows that the randomization of the ethnic job candidates across the different participants was successful. The participants who evaluated white Anglo-Saxon Americans and those who assessed candidates belonging to one of the five race-ethnicity minority groups are similar in terms of their measured characteristics. The same is generally true for the randomization of the candidates' race-ethnicity across the ten different job vacancies (Panel B, columns 2 to 8). Participants evaluated roughly the same number of job applicants from each of the six race-ethnicity groups for every job vacancy, except for the job of exercise physiologist. African, Hispanic, and Arab American candidates were assessed less often for that job.

To check whether our sample represents real-world American HR professionals, we compare some summarized statistics from our sample with a sample of HR professionals in the American Community Survey (ACS) (see Table A–2 in the Appendix). In light of this, we conducted different binomial tests (for the binary variables) and one one-sample $t$-test (to compare the mean age between the two samples). The results show that our sample is not completely representative of real-world American HR professionals. Our sample contains more males than the ACS sample (47.39% females versus 67.00% females, $p = 0.000$), is slightly younger (39.632 years old versus 45.363 years old, $p = 0.000$), and slightly lower educated (82.89% with a tertiary degree versus 86.00%, $p = 0.000$). However, our sample is similar to the ACS sample of recruiters in terms of race-ethnicity.

# 4. Results

## 4.1. Hiring Chances of the Six Race-Ethnicity Groups

Although the main goal of this study was not to measure hiring discrimination, we nonetheless briefly examine whether belonging to a race-ethnicity minority group but not to a lower social class negatively influences one's hiring chances in the context of our experiment.

To this end, we ran a linear regression in which we regressed the standardized version of the invitation probability on the candidates' race-ethnicity, taking white Anglo-Saxon Americans as the reference category and including the other candidate characteristics, participant characteristics, and job vacancies as control variables.[28, 29] The standard errors are corrected for the clustering of observations at the participant level. We can then run an additional regression similar to the one described above, however, instead of including candidates' race-ethnicity in the regression using five dummy variables corresponding to each of the five studied minority groups, we use ten different dummy variables, one per minority group corresponding to Anglo-sounding first names and one to ethnic-sounding first names. We present the results of these analyses in Table 4 (Panel A). In Table A–3 (Panel A) in the Appendix, we present the results of these analyses conducted with a sample of participants with low scores on the MC-SDS (lower than the sample mean plus one standard deviation: 0.736).

<Table 4 about here>

As Table 4 (Panel A, columns 1, 4, 7, 10, and 13) shows, we do not find any negative hiring discrimination toward the different race-ethnicity minority groups when we run the analyses with our total sample. We even find quite the opposite. While both African and Hispanic Americans have invitation chances that are no different from the majority group of white Anglo-Saxon Americans, German Americans, Asian Americans, and Arab Americans have a higher probability to be invited to the second job interview than the majority group (significant at the 5% significance level). If we consider the results of the sample of participants with low scores on the MC-SDS, the results remain generally the same, albeit the positive effects of being a German American or an Arab American on one's hiring chances then become weakly significant at the 10% significance level (Table A–3, Panel A, columns 1, 4, and 13).

Although these overall positive results are initially unexpected and contradictory to the results reported

---

[28] All statistical analyses were performed in Stata and the code used for the different analyses is available upon request.

[29] More complex analyses are unnecessary as our independent variable is experimentally manipulated so that there is no correlation with unobserved characteristics.

in previous studies, there are different explanations for them linked to our research design. First, we focus on second- or higher-generation immigrants (see Section 2.2). In doing so, we account for potentially assumed differences in human capital (e.g., language skills or institutions and quality of education) between majority and minority job applicants (Bisschop, ter Weel, & Zwetsloot, 2020; Jacobs, Rycx, & Volral, in press; Piton & Rycx, 2021). Moreover, second- or higher-generation immigrants might also be perceived as more assimilated to the host country, decreasing the perceived sociocultural distance between them and the majority group. This could explain why the employers in our experiment evaluated minority group job applicants like majority group job candidates. Although previous research has found that, overall, ethnic minorities are penalized in the hiring process (Zschirnt & Ruedin, 2016), studies that particularly diversify between domestic-born (i.e., second or higher generation) and foreign-born immigrants in their correspondence tests find results that align with ours. That is, the hiring discrimination either decreases remarkably when the minority job applicant is born in the host society (Busetta, Campolo, & Panarello, 2018; Carlsson, 2010; Drydakis, 2010; Oreopoulos, 2011; Veit & Thijssen, 2021) or completely disappears (Veit & Thijssen, 2021, found no discrimination toward domestic-born race-ethnicity minorities in Germany and Spain). Furthermore, previous experimental research in which job applicants' levels of assimilation were experimentally manipulated found that signals of higher assimilation (e.g., the type of first name or the extracurricular activities mentioned in a résumé) increased the hiring chances of race-ethnicity minority job applicants (Baert & Vujić, 2016; Derous, Nguyen, and Ryan, 2009; Fossati, Liechti, & Auer, 2020; Kang, DeCelles, Tilcsik, & Jun, 2016).

Another explanation for the overall positive hiring chances found in our study is that we also accounted for the potential unintended name signals of social class. As Gaddis (2017a, 2017b, 2019) has argued, previous research on ethnic discrimination often failed to take these possible additional name signals into account during data collection, which could have led researchers to overestimate the discrimination based on race or ethnicity that they found because some of it could be attributed to the perceived lower social class of the race-ethnicity minority job applicants.[30] Our results suggest that Gaddis's (2017a, 2017b, 2019) criticism is justified. That is, according to our findings, hiring discrimination based on race-ethnicity seems to disappear completely and even reverse for some groups when we keep the social class associated with names constant across the experimental stimuli. This finding is further supported

---

[30] This argument seems valid as previous research in the United States has found that having a lower social class negatively affects someone's hiring chances (Rivera & Tilcsik, 2016) and that people typically rate individuals who appear to be from higher-class backgrounds as more competent and more desirable workers than those from lower-class backgrounds (Fiske, Cuddy, Glick, & Xu, 2002; Neckerman & Kirschenman, 1991; Rivera, 2015; Rivera & Tilcsik, 2016). At the same time, it is widely known that in the United States, many ethnic minority groups, and especially African Americans and Hispanic Americans, are typically associated with a lower social class (Ridgeway & Kricheli-Katz, 2013).

by Darolia et al.'s (2016) correspondence study findings, which also found no discrimination toward Black and Hispanic job applicants in the United States after accounting for the social class signals of the names used in the experiment.

The last explanation for our overall positive results is social desirability bias. Although we designed our vignette experiment to minimize social desirability bias as much as possible (see Sections 2.1 and 2.2) and our results do not change drastically when we rerun our analyses with a subsample of participants with low scores on the MC-SDS, we cannot ensure that our results were completely safeguarded from socially desirable answering. Therefore, we may have somewhat underestimated the ethnic discrimination in hiring toward the five minority groups studied in this research.[31, 32]

If we then consider the types of first names revealed in a résumé (Table 4, Panel A, columns 2, 3, 5, 6, 14, and 15), we see that the advantage that German Americans have over white Anglo-Saxon Americans is present for both job applicants with English-sounding and ethnic-sounding first names, while the advantages that Asian Americans and Arab Americans have mostly exist for job applicants with Anglo-sounding first names. Again, we obtain similar results if we run the analysis with a sample of participants who are less inclined to answer in socially desirable ways (Table A–3, Panel A, columns 2, 3, 5, 6, 14, and 15), albeit the results obtained by doing so are only weakly significant at the 10% significance level, except for the positive effect of having an Anglo-sounding Arab first name on one's invitation probability ($p$ = 0.009).

These results align extremely well with the abovementioned idea that immigrants who signal high levels of assimilation in the host country have higher hiring chances than those who are perceived to be less assimilated (higher-generation immigrants have better chances) (Fossati, Liechti, & Auer, 2020). As Kang

---

[31] Note also that his research was conducted after the murder of George Floyd in the United States, which led to massive protests and the foundation of the Black Lives Matter movement. The incident and the protests that followed could have increased employers' awareness of structural racism and discrimination toward ethnic minorities in the United States, which could have lowered possible unintentional ethnic discrimination. However, the incident could also have increased employers' tendency to answer in socially desirable ways because the subjects became more racially sensitive.

[32] Another potential explanation for the null findings is that our data may lack statistical power. However, we believe that this explanation is invalid because (i) ex ante, we ensured sufficient levels of statistical power by working with a D-efficient resolution V design and conducting the experiment with a relatively large sample of participants, (ii) ex post, we can see that the coefficients of the null findings are positive and not insignificantly negative, and (iii) ex post, we can see that the standard errors of non-significant coefficients are rather small (1.96 times the standard error is still a small number compared to the standard deviation of the outcome variable). In other words, our null hypothesis of no effect could have been rejected even with a small coefficient.

et al. (2016) demonstrated, having an Anglo-sounding first name could signal higher levels of assimilation in the host country, which may cause employers to prefer job applicants with Anglo-sounding first names compared with those with ethnic-sounding first names.

Although we did not find any hiring discrimination toward ethnic minority job applicants, this does not necessarily mean that no stigma toward race-ethnicity minority applicants exists. Both positive and negative stereotypes and attitudes about ethnic minority job applicants may be present but counteracting each other. In the following section, we investigate whether the stereotypes and attitudes identified in previous studies are also identifiable in the context of the hiring process and might give us further insights into why we did not find any hiring discrimination toward some ethnic minority groups and a hiring premium for others.

### 4.2. Positive and Negative Stereotypes and Attitudes: Do They Counteract Each Other?

To investigate the presence of compensating positive or negative stereotypes and attitudes toward the ethnic minority job applicants in our experiment, we ran regressions with the same specifications as those discussed in Section 4.1. We now use the 22 scales related to the different ethnic minority stereotypes and attitudes discussed in Section 2.2 and summarized in Table 2 as dependent variables in each of the individual regressions. We present the results of the regressions concerning the total sample in Table 4 (Panel B) and the results concerning the sample of participants with low scores on the MC-SDS scale in Table A–3 (Panel B) in the Appendix. In both tables, each row in Panel B contains the results of one of the 22 individual regression analyses.

For our total sample (Table, Panel B, columns 1, 4, 7, 10, and 13), we find that employers perceive most of the minority job applicants similarly to the majority applicants. German Americans, African Americans, Hispanic Americans, and Arab Americans are not rated very differently from white Anglo-Saxon American majority applicants on the 22 employer perception and attitude scales. Only African Americans are perceived as less responsible ($p$ = 0.033) and Hispanic Americans are perceived as more ambitious ($p$ = 0.042) than majority group job applicants.[33] However, these two effects do not hold if we rerun the analyses with the MC-SDS sample—the positive perception that Hispanic Americans are

---

[33] We also find some weakly significant effects. German Americans are rated more positively than white Anglo-Saxon job applicants on perceived flexibility and Hispanic Americans more positively on perceived motivation, organizational skills, respect toward authority, and sense of responsibility. Moreover, African American job candidates were evaluated less positively than the majority job candidates about their perceived leadership skills, and Arab Americans were evaluated less positively about customers' willingness to collaborate with them as perceived by employers.

more ambitious becomes insignificant and the negative perception that African American applicants are less responsible becomes significant only at the 10% significance level. Moreover, if we run our analyses with the MC-SDS sample, the weakly significant effect about German Americans' perceived flexibility found for the total sample (see footnote 38) becomes more significant ($p$ = 0.021). [34]

Similar to the results found concerning job candidates' hiring chances, we again found that employers did not rate German, African, Hispanic, or Arab Americans very differently from the majority group. Although this might seem logical for the German American job applicants because they are very socioculturally close to the majority group (Bogardus, 1925; Fiske & Lee, 2006), the results about African, Hispanic, and Arab American job applicants contrast with those from previous stereotype content research (Federal Glass Ceiling Commission, 1995; Ghavami & Peplau, 2012; Lee & Fiske, 2006; Schulz & Haerle, 1995; Zou & Cheryan, 2017). The idea that compensating positive and negative stereotypes about these four race-ethnicity minority groups, which could explain the null effects found in Section 4.1, does not seem to apply here.

However, similar explanations to those mentioned in Section 4.1 might apply here as well. Because we focus on second-or higher-generation immigrants in our study, the race-ethnicity minority job applicants could be perceived as being more assimilated into the host society than first-generation immigrants, which could have increased the perceived sociocultural closeness of these minority groups to the majority group. This may have led to the race-ethnicity minority applicants being perceived as less prototypical for their race-ethnicity groups and, thus, being evaluated less in stereotypical terms than less-assimilated minority group applicants.[35] This statement is supported by Lee and Fiske's (2006) results, who found that third-generation immigrants are rated similarly to white Anglo-Saxon Americans in terms of competence (e.g., intellectual abilities) and warmth (e.g., openness). Moreover, employers might have rated the minority and majority job applicants similarly because we kept the social class constant across the job applicants, further increasing the sociocultural closeness of our race-ethnicity minority groups to the majority group and decreasing the degree to which job applicants belonging to these minority groups conform to the prototypical image of their race-ethnicity group. Moreover, Lee

---

[34] Running the analyses with the MC-SDS sample, some weakly significant effects identified for the total sample disappear while other weakly significant effects appear. All of the weakly significant effects found regarding Hispanic and Arab job applicants disappear (see footnote 38). Meanwhile, weakly significant positive effects appear concerning German Americans' perceived openness and respect toward authority and weakly significant negative effects appear regarding African Americans' perceived social abilities, ambition, and trainability.

[35] The race-ethnicity prototypicality of a person refers to the degree to which a person conforms to the prototypical (or stereotypical) image of a certain race-ethnicity group in society (Ridgeway & Kricheli-Katz, 2013; Van Borm & Baert, 2022).

and Fiske (2006) found that most immigrant groups' stereotypes are associated with their social class. The perceived socioeconomic status of immigrant groups seems to influence the content of the stereotypes about immigrant groups. Our choice to keep the different job candidates' associated social class constant might further explain the low variation in employers' perceptions and attitudes toward the job applicants of the different race-ethnicity groups.

Although employers' perceptions and attitudes toward job applicants do not explain the lack of race-ethnicity effects on German, African, Hispanic, and Arab Americans' hiring chances, they do seem to explain the hiring premium for Asian Americans found in Section 4.1. In the context of our experiment, we do find that employers perceive Asian Americans (i) to have more intellectual abilities, (ii) to have better organizational skills, and (iii) to be more ambitious, (iv) motivated, (v) efficient, and (vi) open to new people and experiences (Table 4, Panel B, column 4). All of these effects are significant at the 5% or 1% significance level.[36] These results remain the same if we run the analyses again with the MC-SDS sample, except for the highly significant positive effect concerning Asian Americans' efficiency, which becomes significant only at the 10% significance level.[37, 38] As employers perceive Asian American job applicants more positively than white Anglo-Saxon Americans, it is unsurprising that their hiring chances are also higher than those of white Anglo-Saxon Americans (see Section 4.1).

If we distinguish the minority job applicants by the type of first name mentioned in their résumés, we find further differences in employers' candidate evaluations. Concerning German Americans (Table 4, Panel B, columns 2–3), we find that employers assume their coworkers will be less willing to collaborate with German Americans with Anglo-sounding first names ($p = 0.036$). However, this result does not hold if we run the analyses with our MC-SDS sample (Table A–3, Panel B, column 2). Additionally, for our total sample, we also find weakly significant positive effects about the perceived social abilities and flexibility of German Americans with ethnic-sounding first names. These two effects become significant at the 5%

---

[36] We also find two weakly significant positive effects—employers perceive Asian American job candidates as having more technological knowledge and skills and a greater sense of responsibility than white Anglo-Saxon job applicants.

[37] When running the analyses with the MC-SDS sample, the weakly significant positive effect related to Asian Americans' technological knowledge and skills that was identified using the total sample disappears, while other weakly significant effects appear. A weakly significant positive effect appears regarding Asian American job applicants' perceived (i) reliability, (ii) flexibility, and (iii) respect toward authority.

[38] These positive perceptions, along with the hiring premium found in Section 4.1., toward Asian American job applicants partially align with the notion of Asian Americans as the 'model minority,' or the idea that Asian Americans are a relatively successful minority group who are highly competent and competitive but also less warm, which was identified in previous studies (Federal Glass Ceiling Commission,1995; Ghavami & Peplau, 2012; Kitano & Sue, 1973; Lee & Fiske, 2006; Lin, Kwan, Cheung & Fiske, 2005; Reyna, Wetherell, & Dobria, 2013; Zou & Cheryan, 2017).

significance level if we run our regressions with the MC-SDS sample (Table A–3, Panel B, column 3). The positive perception that German Americans as a whole are more flexible than white Anglo-Saxon candidates (Table A–3, Panel B, column 1) found when using the MC-SDS sample, therefore, seems to be driven mainly by German Americans with an ethnic-sounding first name.

Furthermore, we find that the positive stereotypes about Asian Americans mostly apply to Asian Americans with ethnic-sounding first names (Table 4, Panel B, columns 5–6). The participants in our experiment perceived Asian Americans with ethnic-sounding first names as (i) having more intellectual abilities ($p = 0.002$), and (ii) being more efficient ($p = 0.021$), and to a lesser extent, also as (iii) having more technological knowledge and skills ($p = 0.051$), (iv) being more motivated ($p = 0.057$), (v) having better organizational skills ($p = 0.067$), (vi) being easier to train ($p = 0.098$), and (vii) being more open ($p = 0.057$) (Table 4, Panel B, columns 6). Alternatively, we only find two significant positive effects of being an Asian American with an Anglo-sounding first name on employers' perception of Asian American job applicants: ambition ($p = 0.030$) and openness ($p = 0.053$) (Table 4, Panel B, columns 5). These results remain generally the same when we rerun the analyses with our MC-SDS sample (Table A–3, Panel B, columns 5–6), albeit some of the weakly significant effects (significant at the 10% level) concerning Asian Americans with ethnic-sounding first names become significant at the 5% significance level (those concerning their perceived motivation, trainability, and openness) and one additional effect occurs: employers in the MC-SDS sample perceive Asian Americans with ethnic-sounding first names as more flexible ($p = 0.036$) than white Anglo-Saxon Americans.

Also, although we found that African Americans were not rated very differently from the majority group overall, we did find that employers perceive African Americans with ethnic-sounding first names to be (i) less trainable and (ii) less responsible, and (iii) to have less social abilities than majority group job candidates (Table 4, Panel B, columns 8–9). These results are also present in the MC-SDS sample, except for the negative effect concerning African Americans with ethnic-sounding first names' sense of responsibility, which becomes insignificant (Table A–3, Panel B, columns 8–9). Our results, therefore, indicate that some of the negative stereotypes about African American people are triggered more strongly when the African American job applicant has an ethnic-sounding first name than if they have an Anglo-sounding first name.

Lastly, we also find that Arab Americans with English-sounding first names are rated more negatively by employers in terms of perceived assertiveness ($p = 0.014$) and customers' willingness to interact with them as employees ($p = 0.045$) (Table 4, Panel B, columns 14–15). While employers' perceptions regarding Arab Americans with Anglo-sounding first names persist and even become more significant ($p = 0.009$) in the MC-SDS sample, employers' beliefs that their customers would be less willing to

interact with these employees become significant at the 10% level ($p$ = 0.097) (Table A–3, Panel B, columns 14–15).[39]

Overall, we can conclude that job candidates with ethnic-sounding first names are evaluated in more stereotypical ways than job applicants who go by Anglo-sounding first names, except for Arab Americans. This is again very much in line with the idea of prototypicality introduced above and often mentioned in intersectionality research (Ridgeway & Kricheli-Katz, 2013; Van Borm & Baert, 2022). Race-ethnicity job applicants who are perceived as less assimilated to the host country—for example, because they have a first name typically associated with a particular race-ethnicity minority group—might be perceived as more prototypical for their race-ethnicity than job applicants who are perceived as more assimilated, such as those without ethnic-sounding first names. The stereotypes existing about a certain race-ethnicity group might be triggered more strongly for job applicants who conform to the prototypical image of their race-ethnicity group.

# 5. Conclusion

To gain insights into the drivers of ethnic discrimination in hiring toward multiple race-ethnicity groups in the United States, we conducted a state-of-the-art scenario experiment. We asked 473 people with genuine experience in hiring to evaluate five fictitious job applicants belonging to different race-ethnicity groups (German American, Asian American, African American, Hispanic American, and Arab American). An important improvement on earlier field and lab experiments on ethnic discrimination was that we ensured that applicant names signaled the same social class. We experimentally varied candidates' sociocultural closeness to the majority group of white Anglo-Saxon Americans by alternating between candidates with Anglo-sounding first names and candidates with ethnic-sounding first names. The recruiters in our experiment first had to indicate their hiring intentions regarding five job applicants. Subsequently, they assessed those job candidates using 22 statements related to the dominant explanations for ethnic discrimination in hiring found in the models of taste-based and statistical discrimination. To address potential social desirability bias, we reran the analyses using a subset of recruiters with low scores on a social desirability scale.

Our results indicate that, after controlling for social class, the hiring chances of job applicants belonging

---

[39] For our total sample, we also find two significant positive effects regarding Hispanic American job applicants with Anglo-sounding first names: ambition and sense of responsibility. However, as they do for the group of Hispanic Americans as a whole, these effects become insignificant when we run the analyses with the MC-SDS sample.

to race-ethnicity minority groups are no lower than those of the majority group of white Anglo-Saxon Americans. On the contrary, a substantial hiring premium for Asian American job applicants was found, aligning with the idea that Asian Americans are perceived as a relatively successful model minority (Kitano & Sue, 1973; Lee & Fiske, 2006). In addition, the hiring chances of German and Arab Americans tend to be higher in some respects.

These results suggest that Gaddis's (2017a, 2017b, 2019) criticism that the hiring discrimination found in previous experimental research might be overestimated because part of the hiring penalty could be attributed to other aspects than race-ethnicity (e.g., social class) is justified. According to our findings, hiring discrimination based on race-ethnicity seems to disappear completely and even reverse, for some groups, when we keep the social class associated with job applicants' full names constant across the experimental stimuli. To further explore Gaddis's (2017a, 2017b, 2019) claim, it would be interesting for future research to conduct field correspondence tests in which, in addition to race-ethnicity and gender, the signaled social class of job applicants was experimentally manipulated. Doing so would offer better insights into how much of the hiring penalty for race-ethnicity minority groups reported in previous correspondence tests is attributable to someone's race-ethnicity and how much can be attributed to other aspects.

Our results also show that the hiring premium found for some of the race-ethnicity minority groups is mainly present for job applicants with Anglo-sounding first names. This result suggests that mentioning signals of higher assimilation (such as first-name types) in their résumés could potentially increase the hiring chances of job applicants belonging to race-ethnicity minority groups. It would be interesting for future research to investigate this claim further through experiments like Baert and Vujić's (2016) and Kang, DeCelles, Tilcsik, and Jun's (2016), in which signals of assimilation (such as the extracurricular activities mentioned in a résumé) are experimentally manipulated. Doing so would permit examining whether these assimilation signals ameliorate the hiring chances of different race-ethnicity minority groups and whether the effect of assimilation varies across race-ethnicity groups.

Furthermore, we did not find much evidence for the presence of racial or ethnic stereotypes about most of the race-ethnicity groups examined in the context of our experiment after controlling for social class. The exception is, again, Asian American job applicants. In accordance with the aforementioned model minority stereotype, employers perceived Asian Americans (i) to have more intellectual abilities and (ii) to have better organizational skills, and (iii) to be more ambitious, (iv) motivated, (v) efficient, and (vi) open to new people and experiences.

We suggest the notion of prototypicality to explain our null results. We argue that by keeping the social

class of our job applicants constant across different race-ethnicity groups, we signaled higher levels of assimilation. Thus, our minority group job applicants might have been perceived as less prototypical for their race-ethnicity groups and, therefore, might have been evaluated in less stereotypical terms than job applicants who were perceived as more prototypical for their race-ethnicity groups. Our results concerning employers' perceptions and attitudes toward job applicants with ethnic-sounding first names versus Anglo-sounding first names seem to confirm this line of thought.

# References

Acquisti, A., & Fong, C. (2020). An experiment in hiring discrimination via online social networks. Management Science, 66(3), 1005–1024.

Aeberhardt, R., Coudin, É., & Rathelot, R. (2016). The heterogeneity of ethnic employment gaps. Journal of Population Economics, 30(1), 307–337. doi: 10.1007/s00148- 016-0602-3

Agerström, J., Björklund, F., Carlsson, R., & Rooth, D. O. (2012). Warm and competent Hassan= cold and incompetent Eric: A harsh equation of real-life hiring discrimination. Basic and Applied Social Psychology, 34(4), 359–366.

Allport, G. W. (1979). The nature of prejudice. Cambridge, MA: Perseus Books.

Ambuehl, S., & Ockenfels, A. (2017). The ethics of incentivizing the uninformed: A vignette study. American Economic Review, 107, 91–95.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. Behavior Research Methods, 52(1), 388–407.

Arrow, K. (1973). The theory of discrimination. Discrimination in Labor Markets, 3(10), 3–33.

Åslund, O., & Rooth, D.-O. (2005). Shifts in attitudes and labor market discrimination: Swedish experiences after 9-11. Journal of Population Economics, 18(4), 603–629. doi: 10.1007/s00148-005-0036-9

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. Methodology, 6(3), 128–138.

Auer, D., Bonoli, G., Fossati, F., & Liechti, F. (2019). The matching hierarchies model: evidence from a survey experiment on employers' hiring intent regarding immigrant applicants. International Migration Review, 53(1), 90–121.

Auspurg, K., & Hinz, T. (2014). Factorial survey experiments. Thousand Oaks: Sage.

Auspurg, K., Hinz, T., & Sauer, C. (2017). Why should women get less? Evidence on the gender pay gap from multifactorial survey experiments. American Sociological Review, 82, 179–210.

Baert, S. (2018a). Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. In Audit Studies: Behind the Scenes with Theory, Method, and Nuance (pp. 63-77). Springer, Cham.

Baert, S. (2018b). Hiring a gay man, taking a risk?: A lab experiment on employment discrimination and risk aversion. Journal of Homosexuality, 65(8), 1015–1031.

Baert, S., & De Pauw, A. S. (2014). Is ethnic discrimination due to distaste or statistics? Economics Letters, 125(2), 270–273.

Baert, S., Lippens, L., & Van Borm, H. (2022). Selecting names for experiments on ethnic discrimination. IZA Discussion Paper Series, 15524.

Baert, S., & Verhaest, D. (2019). Unemployment or Overeducation: Which is a Worse Signal to Employers? De Economist, 167, 1–21.

Baert, S., & Vujić, S. (2016). Immigrant volunteering: a way out of labour market discrimination? Economics Letters, 146, 95–98.

Bartkoski, T., Lynch, E., Witt, C., & Rudolph, C. (2018). A meta-analysis of hiring discrimination against Muslims and Arabs. Personnel Assessment and Decisions, 4(2), 1.

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

Becker, G. S. (1957). The economics of discrimination. Chicago, IL: University of Chicago Press.

Beretvas, S .N., Meyers, J .L., & Leite, W. L. (2002). A reliability generalization study of the Marlowe-Crowne social desirability scale. Educational and Psychological Measurements, 62, 570–589.

Berger, J., Cohen, B., & Zelditch, M. (1966). Status characteristics and expectation states. In Sociological Theories in Progress (pp. 29–46). New York: Houghton Mifflin.

Bertrand, M., & Duflo, E. (2017). Field experiments on discrimination. In Handbook of Economic Field Experiments (Vol. 1, pp. 309–393). Amsterdam: Elsevier.

Bertrand, M., & Mullainathan, S. (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.

Bessudnov, A., & Shcherbak, A. (2019). Ethnic Discrimination in Multi-ethnic Societies: Evidence from Russia. European Sociological Review, 1–17.

Bird, S. R. (1996). Welcome to the men's club: Homosociality and the maintenance of hegemonic masculinity. Gender & Society, 10(2), 120–132.

Bisschop, P., ter Weel, B. & Zwetsloot, J. (2020). Ethnic Employment Gaps of Graduates in the Netherlands. De Economist 168, 577–598.

Blais, A.-R., & Weber, E. (2006). A domain-specific risk-taking (DOSPERT) scale for adult populations. Judgment and Decision Making, 1, 33–47.

Bogardus, E. S. (1925). Social distance and its origins. Journal of Applied Sociology, 9(2), 216–226.

Bonoli, G., & Fossati, F. (2020). More than noise? Explaining instances of minority preference in correspondence studies of recruitment. Journal of Ethnic and Migration Studies, 46(9), 1886–1902.

Booth, A. L., Leigh, A., & Varganova, E. (2012). Does ethnic discrimination vary across minority groups? Evidence from a field experiment. Oxford Bulletin of Economics and Statistics, 74(4), 547–573.

Boyd-Swan, C., & Herbst, C. M. (2019). Racial and ethnic discrimination in the labor market for child care teachers. Educational Researcher, 48(7), 394–406.

Busetta, G., Campolo, M. G., & Panarello, D. (2018). Immigrants and Italian labor market: statistical or taste-based discrimination? Genus, 74(1), 1–20.

Callan, M. J., Kim, H., Gheorghiu, A. I., & Matthews, W. J. (2017). The interrelations between social class, personal relative deprivation, and prosociality. Social Psychological and Personality Science, 8(6), 660–669.

Carlsson, M. (2010). Experimental evidence of discrimination in the hiring of first-and second-generation immigrants. Labour, 24(3), 263–278.

Carlsson, M., Reshid, A. A., & Rooth, D. O. (2018). Neighborhood signaling effects, commuting time, and employment: Evidence from a field experiment. International Journal of Manpower, 39(4), 534–549.

Charness, G., Gneezy, U., & Kuhn, M. A. (2013). Experimental methods: Extra-laboratory experiments-extending the reach of experimental economics. Journal of Economic Behavior & Organization, 91, 93–100.

Colquitt, J.A. (2008). From the editors publishing laboratory research in AMJ: A question of when, not if. Academy of Management Journal, 51(2008), 616–620.

Combes, P.-P., Decreuse, B., Laouénan, M., & Trannoy, A. (2016). Customer discrimination and employment outcomes: Theory and evidence from the French labor market. Journal of Labor Economics, 34(1), 107–160. doi: 10.1086/682332

Coppock, A. (2019). Avoiding post-treatment bias in audit experiments. Journal of Experimental Political Science, 6(1), 1–4.

Costin, V., & Vignoles, V. L. (2020). Meaning is about mattering: Evaluating coherence, purpose, and existential mattering as precursors of meaning in life judgments. Journal of Personality and Social Psychology, 118(4), 864.

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: behaviors from intergroup affect and stereotypes. Journal of Personality and Social Psychology, 92(4), 631.

Damelang, A., & Abraham, M. (2016). You can take some of it with you!: A vignette study on the acceptance of foreign vocational certificates and ethnic inequality in the German labor market. Zeitschrift für Soziologie, 45(2), 91–106.

Damelang, A., Abraham, M., Ebensperger, S., & Stumpf, F. (2019). The hiring prospects of foreign-educated immigrants: A factorial survey among German employers. Work, Employment and Society, 33(5), 739–758.

Darolia, R., Koedel, C., Martorell, P., Wilson, K., & Perez-Arce, F. (2016). Race and gender effects on employer interest in job applicants: new evidence from a resume field experiment. Applied Economics Letters, 23(12), 853–856.

Decker, S. H., Ortiz, N., Spohn, C., & Hedberg, E. (2015). Criminal stigma, race, and ethnicity: The consequences of imprisonment for employment. Journal of Criminal Justice, 43(2), 108–121.

Derous, E., Nguyen, H. H., & Ryan, A. M. (2009). Hiring discrimination against Arab minorities: Interactions between prejudice and job characteristics. Human Performance, 22(4), 297–320.

Derous, E., Ryan, A. M., & Nguyen, H. H. D. (2012). Multiple categorization in resume screening: Examining effects on hiring discrimination against Arab applicants in field and lab settings. Journal of Organizational Behavior, 33(4), 544–570.

Di Stasio, V. (2014). Education as a signal of trainability: Results from a vignette study with Italian employers. European Sociological Review, 30(6), 796–809.

Drydakis, N. (2010). Religious affiliation and employment bias in the labor market. Journal for the Scientific Study of Religion, 49(3), 477–493.

Dülmer, H. (2007). Experimental plans in factorial surveys: random or quota design? Sociological Methods & Research, 35(3), 382–409.

Dutt, A., & Kohfeldt, D. (2019). Assessing the relationship between neoliberal ideology and reactions to Central American asylum seekers in the United States. Journal of Social Issues, 75(1), 134–152.

Edo, A., Jacquemet, N., & Yannelis, C. (2019). Language skills and homophilous hiring discrimination: Evidence from gender and racially differentiated applications. Review of Economics of the Household, 17(1), 349–376.

Eriksson, T., & Kristensen, N. (2014). Wages or fringes? Some evidence on trade-offs and sorting. Journal of Labor Economics, 32(4), 899–928.

Federal Glass Ceiling Commission. (1995). A solid investment: Making full use of the nation's human capital: Recommendations of the Federal Glass Ceiling Commission. The Federal Glass Ceiling Commission. United States.

Flake, D. F. (2019). Do ban-the-box laws really work. Iowa Law Review, 104(3), 1079–1128.

Fossati, F., Liechti, F., & Auer, D. (2020). Can signaling assimilation mitigate hiring discrimination? Evidence from a survey experiment. Research in Social Stratification and Mobility, 65, 100462.

Fouarge, D. (2018). Impact of employer characteristics and beliefs on propensity to hire older job applicants. Paper presented at Netspar International Pension Workshop, Leiden, The Netherlands.

Frimpong, K., Shuridah, O., Wilson, A., & Sarpong, F. (2020). A cross-national investigation of trait antecedents of mobile-banking adoption. Thunderbird International Business Review, 62(4), 411–424.

Gaddis, S. M. (2017a). How black are Lakisha and Jamal? Racial perceptions from names used in correspondence audit studies. Sociological Science, 4, 469–489.

Gaddis, S. M. (2017b). Racial/ethnic perceptions from Hispanic names: Selecting names to test for discrimination. Socius, 3, 2378023117737193.

Gaddis, S. M. (2019). Signaling class: An experiment examining social class perceptions from names used in correspondence audit studies. SSRN Electronic Journal, 3350739.

Ghavami, N., & Peplau, L. A. (2013). An intersectional analysis of gender and ethnic stereotypes: Testing three hypotheses. Psychology of Women Quarterly, 37(1), 113–127.

Gorsuch, M., & Rho, D. (2018). Customer Prejudice & Salience: The Effect of the 2016 Election on Employment Discrimination. Available at SSRN Electronic Journal, 3184984.

Gutfleisch, T., Samuel, R., & Sacchi, S. (2020). The application of factorial surveys to study recruiters' hiring intentions: comparing designs based on hypothetical and real vacancies. Quality & Quantity, 55, 775–804.

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. Proceedings of the National Academy of Sciences, 112(8), 2395–2400.

Hall, R. E. (2001). The ball curve: Calculated racism and the stereotype of African American men. Journal of Black Studies, 32(1), 104–119.

Harrison, L. (2001). Understanding the influence of stereotypes: Implications for the African American in sport and physical activity. Quest, 53(1), 97–114.

Heath, A. F., & Di Stasio, V. (2019). Racial discrimination in Britain, 1969–2017: a meta-analysis of field experiments on racial discrimination in the British labour market. The British Journal of Sociology, 70(5), 1774–1798.

Horvath, T., & Huber, P. (2018). Regional ethnic diversity and the employment prospects of immigrants. Regional Studies, 53(2), 272–282.

Hosoda, M., Stone, D. L., & Stone-Romero, E. F. (2003). The interactive effects of race, gender, and job type on job suitability ratings and selection decisions. Journal of Applied Social Psychology, 33(1), 145–178.

Jacobs, V., Rycx, F. & Volral, M. (2022). Wage Effects of Educational Mismatch According to Workers' Origin: The Role of Demographics and Firm Characteristics. De Economist. doi: 10.1007/s10645-022-09413-9.

Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments. Sociological Methods and Research, 34(3), 334–423.

Kaas, L., & Manger, C. (2012). Ethnic discrimination in Germany's labour market: A field experiment. German Economic Review, 13(1), 1–20.

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened résumés: Race and self-presentation in the labor market. Administrative Science Quarterly, 61(3), 469–502.

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened résumés: Race and self-presentation in the labor market. Administrative Science Quarterly, 61(3), 469–502.

Katz, D., & Braly, K. (1933). Racial stereotypes of one hundred college students. Journal of Abnormal and Social Psychology, 28, 280–290.

Kaufmann, E. (2019). Ethno-traditional nationalism and the challenge of immigration. Nations and

Nationalism, 25(2), 435–448.

Kirschenman, J., & Neckerman, K. M. (1991). We'd love to hire them, but…': The meaning of race for employers. The Urban Underclass, 203, 203–32.

Kitano, H. H., & Sue, S. (1973). The model minorities. Journal of Social Issues, 29(2), 1–9.

Koopmans, R., Veit, S., & Yemane, R. (2019). Taste or statistics? A correspondence study of ethnic, racial and religious labour market discrimination in Germany. Ethnic and Racial Studies, 42(16), 233–252.

Kuhfeld, W. F. (1997). Efficient experimental designs using computerized searches. Sawtooth Software Research Paper Series.

Kuhfeld, W. F. (2010). Statistical graphics in SAS: An introduction to the graph template language and the statistical graphics procedures. SAS Publishing.

Kuhfeld, W. F., Tobias, R. D., & Garratt, M. (1994). Efficient experimental design with marketing research applications. Journal of Marketing Research, 31(4), 545–557.

Lahey, J. N. (2008). Age, women, and hiring an experimental study. Journal of Human Resources, 43(1), 30–35.

Lambert, J. R., & Akinlade, E. Y. (2019). Immigrant stereotypes and differential screening. Personnel Review, 49(4), 921–938.

Lancee, B. (2019). Ethnic discrimination in hiring: comparing groups across contexts. Results from a cross-national field experiment. Journal of Ethnic and Migration Studies, 47(6), 1181–1200.

Lane, T. (2016). Discrimination in the laboratory: A meta-analysis of economics experiments. European Economic Review, 90, 375–402.

Lang, K., & Lehmann, J. Y. K. (2012). Racial discrimination in the labor market: Theory and empirics. Journal of Economic Literature, 50(4), 959–1006.

Laouénan, M. (2017). 'Hate at first sight': Evidence of consumer discrimination against African-Americans in the US. Labour Economics, 46, 94–109.

Leasure, P., & Andersen, T. S. (2020). Race, criminal records, and certificates of relief: An experimental pilot study. Deviant Behavior, 41(9), 1065–1083.

Lee, T. L., & Fiske, S. T. (2006). Not an outgroup, not yet an ingroup: Immigrants in the stereotype content model. International Journal of Intercultural Relations, 30(6), 751–768.

Liebe, U., Moumouni, I. M., Bigler, C., Ingabire, C., & Bieri, S. (2020). Using factorial survey experiments

to measure attitudes, social norms, and fairness concerns in developing countries. Sociological Methods & Research, 49(1), 161–192.

Lin, M. H., Kwan, V. S., Cheung, A., & Fiske, S. T. (2005). Stereotype content model explains prejudice for an envied outgroup: Scale of anti-Asian American stereotypes. Personality and Social Psychology Bulletin, 31(1), 34–47.

Lippens, L., Baert, S., Ghekiere, A., Verhaeghe, P. P., & Derous, E. (2022). Is labour market discrimination against ethnic minorities better explained by taste or statistics? A systematic review of the empirical evidence. Journal of Ethnic and Migration Studies, 48(17), 4243–4276.

Lippens, L., Vermeiren, S., & Baert, S. (2023). The state of hiring discrimination: A meta-analysis of (almost) all recent correspondence experiments. European Economic Review, 151, 104315.

Longley, N. (2003). Measuring employer-based discrimination versus customer-based discrimination: The case of French Canadians in the National Hockey League. American Journal of Economics and Sociology, 62(2), 365–381.

Lucas, B. J., Berry, Z., Giurge, L. M., & Chugh, D. (2021). A longer shortlist increases the consideration of female candidates in male-dominant domains. Nature Human Behaviour, 1–7.

Marreiros, H., Tonin, M., Vlassopoulos, M., & Schraefel, M. C. (2017). 'Now that you mention it': A survey experiment on information, inattention and online privacy. Journal of Economic Behavior & Organization, 140, 1–17.

Mathew, S. (2017). How the second-order free rider problem is solved in a small-scale society. American Economic Review, 107, 578–581.

McDonald, P. (2019). How factorial survey analysis improves our understanding of employer preferences. Swiss Journal of Sociology, 45(2), 237–260.

McPherson, M., Smith-Lovin, L., & Cook, J. (2001). Birds of a feather: Homophily in social networks. Annual Review of Sociology, 27, 415–444.

Migration Policy Institute. (2021). Countries of Birth for U.S. Immigrants, 1960-Present. Retrieved from: https://www.migrationpolicy.org/programs/data-hub/charts/immigrants-countries-birth-over-time

Mobasseri, S. (2019). Race, place, and crime: How violent crime events affect employment discrimination. American Journal of Sociology, 125(1), 63–104.

Montgomery, J. M., Nyhan, B., & Torres, M. (2018). How conditioning on posttreatment variables can ruin your experiment and what to do about it. American Journal of Political Science, 62(3), 760–775.

Moss, P., & Tilly, C. (2001). Stories employers tell: Race, skill, and hiring in America. New York: Russell Sage Foundation.

National Center for O*NET Development. O*NET OnLine. Retrieved from: https://www.onetonline.org/.

Neckerman, K. M., & Kirschenman, J. (1991). Hiring strategies, racial bias, and inner-city workers. Social Problems, 38(4), 433–447.

Neumark (2018): Experimental research on labour market discrimination. Journal of Economic Literature, 56(3), 799–866.

Neyt, B., Baert, S. & Vynckier, J. (in press). Job Prestige and Mobile Dating Success: A Field Experiment. De Economist. doi: 10.1007/s10645-022-09412-w

Nuijten, M. P., Poell, R. F., & Alfes, K. (2017). Extracurricular activities of Dutch university students and their effect on employment opportunities as perceived by both students and organizations. International Journal of Selection and Assessment, 25(4), 360–370.

Olian, J. D., Schwab, D. P., & Haberfeld, Y. (1988). The impact of applicant gender compared to qualifications on hiring recommendations: A meta-analysis of experimental studies. Organizational Behavior and Human Decision Processes, 41(2), 180–195.

Oreopoulos, P. (2011). Why do skilled immigrants struggle in the labor market? A field experiment with thirteen thousand resumes. American Economic Journal: Economic Policy, 3(4), 148–71.

Pager, D., & Karafin, D. (2009). Bayesian bigot? Statistical discrimination, stereotypes, and employer decision making. The Annals of the American Academy of Political and Social Science, 621(1), 70–93.

Palan, S., & Schitter, C. (2018). Prolific.ac—A subject pool for online experiments. Journal of Behavioral and Experimental Finance, 17, 22–27.

Pedulla, D. S. (2018). How race and unemployment shape labor market opportunities: Additive, amplified, or muted effects? Social Forces, 96(4), 1477–1506.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. Journal of Experimental Social Psychology, 70, 153–163.

Peer, E., Rothschild, D. M., Evernden, Z., Gordon, A., & Damer, E. (2021). Mturk, Prolific or panels? Choosing the right audience for online research. Choosing the right audience for online research. SSRN Electronic Journal, 3765448.

Peitz, L., Dhont, K., & Seyd, B. (2018). The psychology of supranationalism: Its ideological correlates and implications for EU attitudes and post-Brexit preferences. Political Psychology, 39(6), 1305–1322.

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

Piton, C., & Rycx, F. (2021). A Broken Social Elevator? Employment Outcomes of First- and Second-Generation Immigrants in Belgium. *De Economist*, 169, 319–365.

Quillian, L., Heath, A., Pager, D., Midtbøen, A. H., Fleischmann, F., & Hexel, O. (2019). Do some countries discriminate more than others? Evidence from 97 field experiments of racial discrimination in hiring. Sociological Science, 6, 467–496.

Quillian, L., & Midtbøen, A. H. (2021). Comparative perspectives on racial discrimination in hiring: The rise of field experiments. Annual Review of Sociology, 47, 391–415

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. Proceedings of the National Academy of Sciences, 114(41), 10870–10875.

Reyna, C., Dobria, O., & Wetherell, G. (2013). The complexity and ambivalence of immigration attitudes: Ambivalent stereotypes predict conflicting attitudes toward immigration policies. Cultural Diversity and Ethnic Minority Psychology, 19(3), 342.

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. Journal of Clinical Psychology, 38(1), 119–125.

Riach, P., & Rich, J. (2004). Fishing for discrimination. Review of Social Economy, 62, 465–486.

Rich, J. (2014). What do field experiments of discrimination in markets tell us? A meta-analysis of studies conducted since 2000. IZA Discussion Paper Series, 8584.

Ridgeway, C.L., & Fisk, S. (2012). Class rules, status dynamics, and 'gateway' interactions. In Facing Social Class: How Societal Rank Influences Interaction (pp. 131–51). Russell Sage, New York.

Ridgeway, C. L., & Kricheli-Katz, T. (2013). Intersecting cultural beliefs in social relations: Gender, race, and class binds and freedoms. Gender & Society, 27(3), 294–318.

Ritter, J. A., & Taylor, L. J. (2011). Racial disparity in unemployment. Review of Economics and Statistics, 93(1), 30–42.

Rivera, L. A. (2015). Pedigree: How elite students get elite jobs. Princeton, NJ: Princeton University Press.

Rivera, L. A., & Tilcsik, A. (2016). Class advantage, commitment penalty: The gendered effect of social class signals in an elite labor market. American Sociological Review, 81, 1097–1131.

Rossi, P. H. (1951). The Application of Latent Structure Analysis to the Study of Social Stratification. Ph.D.

dissertation, Columbia University.

Rossi, P. H., & Nock, S. L. (1982). Measuring social judgements: The factorial survey approach. Thousand Oaks, CA: Sage.

Roster, C. A., & Ferrari, J. R. (2020). Does work stress lead to office clutter, and how? Mediating influences of emotional exhaustion and indecision. Environment and Behavior, 52(9), 923–944.

Sârbescu, P., Costea, I., & Rusu, S. (2011). Psychometric properties of the Marlowe-Crowne Social Desirability Scale in a Romanian sample. Procedia—Social and Behavioral Sciences, 33(1), 707–711.

Sauer, C., Auspurg, K., & Hinz, T. (2020). Designing multi-factorial survey experiments: Effects of presentation style (text or table), answering scales, and vignette order. Methods, Data, Analyses: MDA, 14(2), 195–214.

Sauer, C., Auspurg, K., Hinz, T., & Liebig, S. (2011). The application of factorial surveys in general population samples: The effects of respondent age and education on response times and response consistency. Survey Research Methods, 5, 89–102.

Schulz, R. A., & Haerle, B. M. (1995). Beer, fast cars, and…: Stereotypes held by US college-level students of German. Die Unterrichtspraxis/Teaching German, 29–39.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston: Houghton Mifflin Company.

Shamon, H., Dülmer, H., & Giza, A. (2022). The factorial survey: the impact of the presentation format of vignettes on answer behavior and processing time. Sociological Methods & Research, 51(1), 396–438

Singh, J., Crisafulli, B., Quamina, L. T., & Kottasz, R. (2020). The role of brand equity and crisis type on corporate brand alliances in crises. European Management Review, 17(4), 821–834.

Social Security Administration. (2021). Popular Baby Names. Retrieved from: https://www.ssa.gov/OACT/babynames/

Sterkens, P., Baert, S., Rooman, C., & Derous, E. (2021). As if it weren't hard enough already: Breaking down hiring discrimination following burnout. Economics & Human Biology, 43, 101050.

Stone, J., Lynch, C. I., Sjomeling, M., & Darley, J. M. (1999). Stereotype threat effects on Black and White athletic performance. Journal of Personality and Social Psychology, 77(6), 1213.

Stone, J., Perry, W., & Darley, J. M. (1997). 'White men can't jump': Evidence for the perceptual

confirmation of racial stereotypes following a basketball game. Basic and Applied Social Psychology, 19(3), 291–306.

Thijssen, L., Coenders, M., & Lancee, B. (2020). Ethnic discrimination in the Dutch labor market: Differences between ethnic minority groups and the role of personal information about job applicants—evidence from a field experiment. Journal of International Migration and Integration, 1–26.

Thijssen, L., Coenders, M., & Lancee, B. (2021). Is there evidence for statistical discrimination against ethnic minorities in hiring? Evidence from a cross-national field experiment. Social Science Research, 93, 102482.

Thijssen, L., Lancee, B., Veit, S., & Yemane, R. (2019). Discrimination against Turkish minorities in Germany and the Netherlands: field experimental evidence on the effect of diagnostic information on labour market outcomes. Journal of Ethnic and Migration Studies, 1–18.

Tajfel, H. (1974). Social identity and intergroup behaviour. Information (International Social Science Council), 13(2), 65–93.

U.S. Census Bureau (2021a). Frequently Occurring Surnames from the 2010 Census. Retrieved from: https://www.census.gov/topics/population/genealogy/data/2010_surnames.html

U.S. Census Bureau (2021b). QuickFacts United States. Retrieved from: https://www.census.gov/quickfacts/fact/table/US/PST045219

Valfort, M. A. (2020). Anti-Muslim discrimination in France: Evidence from a field experiment. World Development, 135, 105022.

Van Belle, E., Caers, R., Cuypers, L., De Couck, M., Neyt, B., Van Borm, H., & Baert, S. (2020). What do student jobs on graduate CVs signal to employers? Economics of Education Review, 75, 101979.

Van Belle, E., Caers, R., De Couck, M., Di Stasio, V., & Baert, S. (2019). The signal of applying for a job under a vacancy referral scheme. Industrial Relations, 58(2), 251–274.

Van Belle, E., Di Stasio, V., Caers, R., De Couck, M., & Baert, S. (2018). Why are employers put off by long spells of unemployment? European Sociological Review, 34(6), 694–710.

Van Borm, H. & Baert, S. (2022). Diving in the Minds of Recruiters: When do Stereotypes Affect the Hiring Chances of Women? IZA Discussion Paper Series, 15261.

Van Borm, H., Burn, I., & Baert, S. (2021). What does a job candidate's age signal to employers? Labour Economics, 102003.

Van Borm, H., Dhoop, M., Van Acker, A., & Baert, S. (2020). What does someone's gender identity signal

to employers? International Journal of Manpower, 41(6), 753–777.

Van Hoye, G., & Lievens, F. (2003). The effects of sexual orientation or hirability ratings: An experimental study. Journal of Business and Psychology, 18(1), 15–30.

Veit, S., Arnu, H., Di Stasio, V., Yemane, R., & Coenders, M. (2022). The 'big two' in hiring discrimination: Evidence from a cross-national field experiment. Personality and Social Psychology Bulletin, 48(2), 167–182.

Veit, S., & Thijsen, L. (2021). Almost identical but still treated differently: Hiring discrimination against foreign-born and domestic-born minorities. Journal of Ethnic and Migration Studies, 47(6), 1285–1304.

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. Science Research, 38, 505–520.

Webster, S. W., O'Toole, R., O'Toole, A. W., & Lucal, B. (2005). Overreporting and underreporting of child abuse: Teachers' use of professional discretion. Child Abuse & Neglect, 29(11), 1281–1296.

Widner, D., & Chicoine, S. (2011). It's all in the name: Employment discrimination against Arab Americans 1. In Sociological Forum (Vol. 26, No. 4, pp. 806–823). Blackwell Publishing Ltd., Oxford, UK.

Yemane, R., & Fernández-Reino, M. (2019). Latinos in the United States and in Spain: the impact of ethnic group stereotypes on labour market outcomes. Journal of Ethnic and Migration Studies, 1–21.

Yemane, R. (2020). Cumulative disadvantage? The role of race compared to ethnicity, religion, and non-white phenotype in explaining hiring discrimination in the US labour market. Research in Social Stratification and Mobility, 69, 100552.

Zou, L. X., & Cheryan, S. (2017). Two axes of subordination: A new model of racial position. Journal of Personality and Social Psychology, 112(5), 696.

Zschirnt, E., & Ruedin, D. (2016). Ethnic discrimination in hiring decisions: a meta-analysis of correspondence tests 1990–2015. Journal of Ethnic and Migration Studies, 42(7), 1115–1134.

# Appendix

<Table A–1 about here>

<Table A–2 about here>

<Table A–3 about here>

*Table 1. Vignette Factors and Corresponding Levels Used in the Experimental Materials*

| Vignette factors | Vignette levels |
|---|---|
| Name | {7x Name of white Anglo-Saxon man, Name of German man, Name of Asian man, Name of Black man, Name of Hispanic man, Name of Arab man,7x Name of white Anglo-Saxon woman, Name of German woman, Name of Asian woman, Name of Black woman, Name of Hispanic woman, Name of Arab woman} |
| Commuting distance | {0–5 miles, 5–10 miles, 10–50 miles, More than 50 miles} |
| Experience in the occupation | {None, About 2 years, About 5 years, About 10 years} |
| Recent period of unemployment | {Yes, No} |
| Extracurricular activities | {None, Volunteering, Sport activities, Cultural activities} |

Note: The factorial product of the vignette levels (i.e., 24 x 4 x 4 x 2 x 4) generated 3,072 possible combinations. In total, 96 sets of five vignettes were drawn from this vignette universe using a D-efficient resolution V design (D-efficiency: 92.13; Auspurg & Hinz, 2014) and distributed at random to the participants as described in Subsection 2.1.

*Table 2. Statements Used in the Experimental Materials*

| Signals and evaluation outcomes | Statements |
| --- | --- |
| Perceived intellectual abilities | 'People with this profile usually have enough intellectual capacities to perform this job well.' |
| Perceived social abilities | 'People with this profile usually have enough social capacities to perform this job well.' |
| Perceived physical abilities | 'People with this profile usually have enough physical capacities to perform this job well.' |
| Perceived technological knowledge and skills | 'People with this profile usually have enough technological knowledge and skills to perform this job well.' |
| Perceived language skills | 'People with this profile usually have enough language skills to perform this job well.' |
| Perceived leadership skills | 'People with this profile usually have enough leadership skills to perform this job well.' |
| Perceived assertiveness | 'People with this profile usually are assertive enough to perform this job well.' |
| Perceived ambition | 'People with this profile usually are ambitious enough to perform this job well.' |
| Perceived motivation | 'People with this profile usually are motivated enough to perform this job well.' |
| Perceived autonomy | 'People with this profile usually are autonomous enough to perform this job well.' |
| Perceived efficiency | 'People with this profile usually are efficient enough to perform this job well.' |
| Perceived organizational skills | 'People with this profile usually are organized enough to perform this job well.' |
| Perceived reliability | 'People with this profile usually are reliable enough to perform this job well.' |
| Perceived trainability | 'People with this profile usually are trainable enough to perform this job well.' |
| Perceived flexibility | 'People with this profile usually are flexible enough to perform this job well.' |
| Perceived accuracy | 'People with this profile usually are accurate enough to perform this job well.' |
| Perceived openness to new people and experiences | 'People with this profile usually are open enough to new people and experiences to perform this job well.' |
| Perceived respect toward authority | 'People with this profile usually have enough respect towards authority to perform this job well.' |
| Perceived sense of responsibility | 'People with this profile usually have a large enough sense of responsibility to perform this job well.' |
| Employer's attitude toward collaboration | 'I think I usually would enjoy collaborating with people like this person.' |
| Other employees' predicted attitude toward collaboration | 'I think other employees usually would enjoy collaborating with people like this person.' |
| Customers' predicted attitude toward collaboration | 'I think customers usually would enjoy collaborating with people like this person.' |
| Invitation probability | 'I advise to invite this candidate for the second phase of the application process.' |

Note: In this table, we present the potential signals of the five minority groups, the evaluation outcome, and their corresponding statements as they were presented in the online survey experiment. The participants evaluated each statement on an 11-point Likert scale ranging from 0 (strongly disagree) to 10 (strongly agree).

*Table 3. Data Description by Fictitious Candidates' Ethnic Backgrounds*

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Total sample [$N$ = 2,362] | White Anglo-Saxon Americans [$N$ = 1,607] | German Americans [$N$ = 67] | Asian Americans [$N$ = 190] | African Americans [$N$ = 131] | Hispanic Americans [$N$ = 204] | Arab Americans [$N$ = 163] | $F$-statistic (df = 5) |
| **A. PARTICIPANT CHARACTERISTICS** | | | | | | | | |
| Gender: female | 0.474 | 0.477 | 0.388 | 0.519 | 0.458 | 0.502 | 0.401 | 1.560 |
| Age: <40 years old | 0.581 | 0.576 | 0.552 | 0.626 | 0.626 | 0.5833 | 0.558 | 0.690 |
| Race-ethnicity: white Anglo-Saxon | 0.818 | 0.817 | 0.821 | 0.842 | 0.809 | 0.779 | 0.853 | 0.840 |
| Highest educational degree: university | 0.759 | 0.761 | 0.746 | 0.737 | 0.756 | 0.745 | 0.785 | 0.290 |
| Employment status: full time employment | 0.789 | 0.801 | 0.716 | 0.768 | 0.779 | 0.779 | 0.748 | 1.150 |
| Frequency of hiring: ≥once per semester | 0.697 | 0.706 | 0.582 | 0.668 | 0.687 | 0.701 | 0.693 | 1.120 |
| Experience as HR professional: >1 year | 0.922 | 0.918 | 0.955 | 0.953 | 0.924 | 0.902 | 0.932 | 1.050 |
| Experience as HR professional: >5 years | 0.497 | 0.488 | 0.567 | 0.505 | 0.496 | 0.500 | 0.534 | 0.540 |
| **B. JOB CHARACTERISTICS** | | | | | | | | |
| Sewing machine operator | 0.099 | 0.092 | 0.119 | 0.110 | 0.061 | 0.113 | 0.153 | 1.890* |
| Usher | 0.113 | 0.111 | 0.164 | 0.100 | 0.145 | 0.098 | 0.117 | 0.790 |
| Stock clerk | 0.100 | 0.103 | 0.149 | 0.089 | 0.084 | 0.098 | 0.074 | 0.780 |
| Data entry keyer | 0.101 | 0.100 | 0.045 | 0.121 | 0.076 | 0.113 | 0.110 | 0.910 |
| Farm labor contractor | 0.094 | 0.100 | 0.104 | 0.058 | 0.099 | 0.083 | 0.074 | 0.970 |
| Judicial law clerk | 0.106 | 0.108 | 0.075 | 0.100 | 0.092 | 0.103 | 0.123 | 0.330 |
| Gem and diamond worker | 0.097 | 0.088 | 0.104 | 0.105 | 0.114 | 0.113 | 0.141 | 1.240 |
| Exercise physiologist | 0.096 | 0.110 | 0.104 | 0.074 | 0.061 | 0.069 | 0.043 | 2.750** |
| Computer hardware engineer | 0.095 | 0.092 | 0.044 | 0.100 | 0.145 | 0.118 | 0.074 | 1.610 |
| Management analyst | 0.099 | 0.095 | 0.089 | 0.142 | 0.122 | 0.093 | 0.092 | 1.060 |

Note: We present the descriptive statistics for our full sample as well as for six subsamples of applicants classified by their ethnic backgrounds. The statistics in column 8 are F-values from one-way ANOVA estimations. *, ** and *** indicate significance at the 10%, 5% and 1% levels.

*Table 4. Relationship between Candidates' Ethnic Backgrounds and Employers' Candidate Evaluations (Total Sample)*

| Reference: White Anglo-Saxon Americans | German Americans | | | Asian Americans | | | African Americans | | | Hispanic Americans | | | Arab Americans | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *N* = 2,259 | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) |
| | Regression coefficients | | | Regression coefficients | | | Regression coefficients | | | Regression coefficients | | | Regression coefficients | | |
| **A. HIRING INTENTIONS** | | | | | | | | | | | | | | | |
| Invitation probability | 0.232** (0.093) | 0.217* (0.114) | 0.238* (0.131) | 0.147** (0.058) | 0.163** (0.073) | 0.123 (0.091) | 0.043 (0.056) | −0.009 (0.082) | 0.075 (0.072) | 0.050 (0.057) | 0.109 (0.075) | −0.028 (0.083) | 0.128** (0.057) | 0.250*** (0.082) | 0.022 (0.075) |
| **B. SIGNALS** | | | | | | | | | | | | | | | |
| 1 Perceived intellectual abilities | 0.097 (0.115) | 0.148 (0.164) | 0.067 (0.157) | 0.183*** (0.063) | 0.090 (0.079) | 0.327*** (0.104) | 0.000 (0.066) | 0.095 (0.112) | −0.055 (0.081) | 0.017 (0.060) | 0.149* (0.076) | −0.158 (0.097) | 0.084 (0.071) | −0.037 (0.098) | 0.188* (0.099) |
| 2 Perceived social abilities | 0.081 (0.094) | −0.134 (0.139) | 0.208* (0.119) | −0.024 (0.075) | −0.033 (0.092) | −0.010 (0.128) | −0.102 (0.075) | 0.098 (0.097) | −0.220** (0.101) | 0.062 (0.059) | 0.103 (0.072) | 0.010 (0.095) | −0.043 (0.076) | 0.000 (0.097) | −0.081 (0.112) |
| 3 Perceived physical abilities | −0.040 (0.112) | 0.136 (0.151) | −0.146 (0.152) | 0.006 (0.068) | −0.010 (0.087) | 0.033 (0.109) | −0.009 (0.078) | 0.123 (0.123) | −0.086 (0.098) | −0.029 (0.060) | 0.004 (0.079) | −0.072 (0.091) | 0.066 (0.073) | 0.115 (0.095) | 0.023 (0.107) |
| 4 Perceived technological knowledge and skills | 0.032 (0.111) | −0.051 (0.167) | 0.080 (0.145) | 0.104* (0.061) | 0.059 (0.080) | 0.175* (0.089) | −0.057 (0.069) | 0.024 (0.107) | −0.105 (0.090) | 0.000 (0.061) | 0.067 (0.077) | −0.088 (0.093) | 0.071 (0.064) | 0.101 (0.076) | 0.044 (0.096) |
| 5 Perceived language skills | 0.060 (0.108) | 0.010 (0.152) | 0.089 (0.147) | 0.013 (0.074) | −0.013 (0.094) | 0.051 (0.119) | 0.046 (0.078) | 0.214* (0.120) | −0.053 (0.099) | 0.057 (0.062) | 0.110 (0.083) | −0.012 (0.096) | −0.041 (0.073) | −0.050 (0.094) | −0.034 (0.105) |
| 6 Perceived leadership skills | 0.002 (0.105) | −0.141 (0.153) | 0.087 (0.139) | 0.093 (0.068) | 0.085 (0.09) | 0.104 (0.103) | −0.121* (0.072) | −0.054 (0.105) | −0.160* (0.095) | 0.041 (0.059) | 0.068 (0.077) | 0.005 (0.09) | −0.049 (0.067) | −0.113 (0.079) | 0.006 (0.103) |
| 7 Perceived assertiveness | −0.009 (0.123) | −0.137 (0.181) | 0.067 (0.164) | 0.005 (0.066) | 0.026 (0.079) | −0.028 (0.118) | 0.007 (0.073) | 0.027 (0.105) | −0.005 (0.097) | −0.023 (0.061) | 0.051 (0.078) | −0.121 (0.095) | −0.087 (0.071) | −0.217** (0.089) | 0.025 (0.103) |
| 8 Perceived ambition | 0.126 (0.106) | 0.135 (0.155) | 0.121 (0.140) | 0.177*** (0.066) | 0.171** (0.078) | 0.187 (0.116) | −0.105 (0.067) | −0.041 (0.109) | −0.142* (0.081) | 0.129** (0.063) | 0.195** (0.082) | 0.041 (0.100) | 0.033 (0.066) | −0.109 (0.093) | 0.156* (0.086) |
| 9 Perceived motivation | 0.157 (0.103) | 0.134 (0.172) | 0.172 (0.128) | 0.138** (0.068) | 0.094 (0.086) | 0.205* (0.108) | −0.054 (0.069) | 0.078 (0.083) | −0.132 (0.095) | 0.107* (0.062) | 0.109 (0.079) | 0.104 (0.100) | 0.101 (0.069) | 0.032 (0.097) | 0.162* (0.093) |
| 10 Perceived autonomy | 0.079 (0.098) | 0.141 (0.129) | 0.041 (0.137) | 0.060 (0.068) | −0.005 (0.081) | 0.162 (0.118) | −0.100 (0.075) | −0.017 (0.104) | −0.148 (0.102) | 0.013 (0.060) | 0.106 (0.074) | −0.109 (0.099) | −0.076 (0.068) | −0.050 (0.090) | −0.100 (0.099) |
| 11 Perceived efficiency | 0.140 (0.104) | 0.258 (0.162) | 0.071 (0.136) | 0.126** (0.064) | 0.057 (0.080) | 0.234** (0.101) | −0.017 (0.071) | −0.012 (0.101) | −0.020 (0.096) | 0.039 (0.062) | 0.114 (0.081) | −0.061 (0.095) | 0.016 (0.069) | −0.038 (0.094) | 0.062 (0.096) |
| 12 Perceived organizational skills | 0.128 (0.106) | 0.134 (0.139) | 0.125 (0.150) | 0.152** (0.069) | 0.107 (0.081) | 0.224* (0.122) | −0.060 (0.069) | −0.093 (0.092) | −0.040 (0.094) | 0.106* (0.062) | 0.148* (0.075) | 0.051 (0.101) | −0.025 (0.070) | −0.092 (0.094) | 0.032 (0.099) |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 13 Perceived reliability | 0.086 | 0.136 | 0.056 | 0.107 | 0.084 | 0.144 | −0.096 | −0.067 | −0.112 | 0.055 | 0.110 | −0.018 | 0.005 | −0.038 | 0.043 |
| | (0.105) | (0.168) | (0.135) | (0.069) | (0.086) | (0.115) | (0.071) | (0.090) | (0.098) | (0.062) | (0.078) | (0.098) | (0.065) | (0.082) | (0.095) |
| 14 Perceived trainability | 0.150 | 0.178 | 0.133 | 0.111 | 0.062 | 0.186* | −0.102 | 0.101 | **−0.221**\*\* | −0.014 | 0.052 | −0.101 | 0.035 | −0.037 | 0.098 |
| | (0.092) | (0.125) | (0.124) | (0.071) | (0.090) | (0.112) | (0.071) | (0.102) | (0.090) | (0.065) | (0.082) | (0.107) | (0.071) | (0.093) | (0.100) |
| 15 Perceived flexibility | 0.181* | 0.101 | 0.230* | 0.090 | 0.040 | 0.167 | −0.021 | −0.044 | −0.009 | 0.085 | 0.124 | 0.032 | −0.035 | −0.103 | 0.023 |
| | (0.101) | (0.141) | (0.135) | (0.072) | (0.094) | (0.103) | (0.079) | (0.108) | (0.106) | (0.064) | (0.083) | (0.100) | (0.069) | (0.086) | (0.102) |
| 16 Perceived accuracy | 0.143 | 0.183 | 0.120 | 0.104 | 0.079 | 0.143 | −0.016 | 0.076 | −0.070 | 0.069 | 0.126* | −0.006 | −0.016 | −0.126 | 0.080 |
| | (0.101) | (0.167) | (0.128) | (0.067) | (0.088) | (0.107) | (0.070) | (0.099) | (0.094) | (0.058) | (0.076) | (0.089) | (0.065) | (0.086) | (0.094) |
| 17 Perceived openness to new people and experiences | 0.134 | 0.097 | 0.155 | **0.201**\*\*\* | 0.188* | 0.221* | −0.002 | 0.137 | −0.083 | 0.062 | 0.122 | −0.016 | −0.039 | −0.069 | −0.013 |
| | (0.108) | (0.137) | (0.150) | (0.075) | (0.097) | (0.116) | (0.071) | (0.107) | (0.091) | (0.062) | (0.080) | (0.098) | (0.072) | (0.097) | (0.102) |
| 18 Perceived respect toward authority | 0.112 | 0.068 | 0.135 | 0.105 | 0.142 | 0.049 | −0.009 | 0.107 | −0.077 | 0.114* | 0.147* | 0.072 | −0.012 | 0.009 | −0.029 |
| | (0.100) | (0.149) | (0.135) | (0.078) | (0.094) | (0.135) | (0.073) | (0.105) | (0.097) | (0.065) | (0.083) | (0.104) | (0.080) | (0.107) | (0.116) |
| 19 Perceived sense of responsibility | 0.083 | 0.077 | 0.087 | 0.109* | 0.088 | 0.142 | **−0.134**\*\* | −0.077 | **−0.168**\*\* | 0.103* | **0.185**\*\* | −0.007 | 0.020 | −0.069 | 0.096 |
| | (0.098) | (0.111) | (0.142) | (0.066) | (0.081) | (0.113) | (0.063) | (0.088) | (0.082) | (0.057) | (0.074) | (0.089) | (0.068) | (0.089) | (0.099) |
| 1 Employer's attitude toward collaboration | −0.045 | −0.145 | 0.015 | 0.076 | 0.061 | 0.099 | 0.041 | 0.113 | −0.001 | 0.078 | 0.143* | −0.007 | 0.022 | −0.115 | 0.140 |
| | (0.101) | (0.143) | (0.137) | (0.065) | (0.083) | (0.106) | (0.066) | (0.092) | (0.087) | (0.062) | (0.081) | (0.093) | (0.070) | (0.087) | (0.102) |
| 2 Employees' predicted attitude toward collaboration | 0.010 | **−0.298**\*\* | 0.193 | −0.027 | −0.054 | 0.013 | −0.001 | 0.125 | −0.074 | 0.025 | 0.098 | −0.071 | −0.099 | −0.127 | −0.075 |
| | (0.109) | (0.142) | (0.142) | (0.068) | (0.087) | (0.105) | (0.068) | (0.093) | (0.089) | (0.061) | (0.076) | (0.096) | (0.070) | (0.087) | (0.105) |
| 3 Customers' predicted attitude toward collaboration | 0.025 | −0.146 | 0.126 | 0.048 | 0.019 | 0.091 | −0.058 | 0.072 | −0.135 | 0.019 | 0.078 | −0.059 | −0.133* | **−0.192**\*\* | −0.081 |
| | (0.095) | (0.146) | (0.121) | (0.067) | (0.087) | (0.103) | (0.068) | (0.087) | (0.094) | (0.065) | (0.085) | (0.100) | (0.072) | (0.096) | (0.103) |

Note: We present coefficient estimates for the 23 distinct linear regression models outlined in Sections 4.1 and 4.2 in which we regress the standardized version of (i) the invitation probability (Panel A) and (ii) the stereotypes and attitudes toward race-ethnicity minority groups (Panel B) based on each candidate's ethnic background. Each row of this table, thus, contains the results from one regression analysis. Standard errors are presented in parentheses and corrected for the clustering of observations at the participant level. *, ** and *** indicate significance at the 10%, 5% and 1% levels. The coefficients related to *p*-values below 5% are in bold.

*Table A–1. Job Descriptions Used in the Experimental Materials*

| Job function | Job description |
|---|---|
| Sewing machine operator | 'This employee will be responsible for operating or tending sewing machines to join, reinforce, decorate, or perform related sewing operations in the manufacture of garment and non-garment products.' |
| Usher | 'This employee will be responsible for assisting patrons at entertainment events. His or her duties include collecting admission tickets and passes from patrons, assisting patrons in finding seats, searching for lost articles, and helping patrons with locating such facilities as restrooms and telephones.' |
| Stock clerk | 'This employee will be responsible for receiving, storing, and issuing merchandise, materials, equipment, and other items from the warehouse or stockroom. Additionally, he or she is responsible for filling shelves, racks, tables, or customers' orders. He or she must operate power equipment to fill orders, mark prices on merchandise, and set up sales displays.' |
| Data entry keyer | 'This employee will be responsible for operating data entry devices, such as keyboards or photo composing perforators. His or her duties include verifying data and preparing materials for printing.' |
| Farm labor contractor | 'This employee will be responsible for recruiting and hiring seasonal or temporary agricultural laborers.' |
| Judicial law clerk | 'This employee will be responsible for assisting judges in court, conducting research, and preparing legal documents.' |
| Gem and diamond worker | 'This employee will be responsible for fabricating, finishing, and evaluating the quality of gems and diamonds used in jewelry or industrial tools.' |
| Exercise physiologist | 'This employee will be responsible for assessing, planning, and implementing fitness programs that include exercise or physical activities such as those designed to improve cardiorespiratory function, body composition, muscular strength, muscular endurance, or flexibility.' |
| Computer hardware engineer | 'This employee will be responsible for researching, designing, developing, and testing computer or computer-related equipment for commercial, industrial, and scientific use. He or she must supervise the manufacturing and installation of computer or computer-related equipment and components.' |
| Management analyst | 'This employee will be responsible for conducting organizational studies and evaluations. His or her duties include designing systems and procedures, conducting work simplification and measurement studies, and preparing operations and procedures manuals to assist management in operating more efficiently and effectively.' |

Note: Job functions and descriptions were provided by O*NET, as described in Section 2.2.

*Table A–2. Comparison between Participant Characteristics and the Characteristics of HR Professionals in ACS*

| Participant characteristics | (1) Mean among participants in the experiment | (2) Mean among HR professionals in ACS |
|---|---|---|
| Female gender | 0.474 | 0.670 |
| Age | 39.632 | 45.363 |
| Race-ethnicity: only white | 0.818 | 0.811 |
| Highest educational degree: secondary education or lower | 0.171 | 0.140 |
| Highest educational degree: tertiary education | 0.829 | 0.860 |

Notes: We combined ACS data from 2010–2019 and selected all respondents with occ2010 occupation codes 0130 (Human Resources Managers), 0620 (Human Resources, Training, and Labour Relations Specialists), and 5360 (Human Resources Assistants, Except Payroll and Timekeeping).

*Table A–3. Relationships between Candidates' Ethnic Backgrounds and Employers' Perceptions of Candidates (Sample of Participants with Low Scores on MC-SDS)*

| Reference: White Anglo-Saxon Americans | German Americans | | | Asian Americans | | | African Americans | | | Hispanic Americans | | | Arab Americans | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N = 1,844 | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names | All first names | Anglo-sounding first names | Ethnic-sounding first names |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) |
| | Regression coefficients | | | Regression coefficients | | | Regression coefficients | | | Regression coefficients | | | Regression coefficients | | |
| **A. HIRING INTENTIONS** | | | | | | | | | | | | | | | |
| Invitation probability | 0.186* (0.106) | 0.201 (0.135) | 0.177 (0.144) | 0.155** (0.063) | 0.153* (0.082) | 0.16 (0.098) | −0.004 (0.06) | −0.117 (0.085) | 0.065 (0.078) | 0.038 (0.062) | 0.082 (0.087) | −0.015 (0.088) | 0.124* (0.064) | 0.246*** (0.094) | 0.007 (0.082) |
| **B. SIGNALS** | | | | | | | | | | | | | | | |
| Perceived intellectual abilities | 0.087 (0.128) | 0.019 (0.185) | 0.125 (0.170) | 0.200*** (0.068) | 0.090 (0.086) | 0.368*** (0.115) | −0.004 (0.073) | 0.082 (0.125) | −0.056 (0.089) | −0.012 (0.066) | 0.086 (0.090) | −0.127 (0.100) | 0.049 (0.081) | −0.049 (0.108) | 0.142 (0.119) |
| Perceived social abilities | 0.151 (0.101) | −0.074 (0.157) | 0.268** (0.124) | −0.023 (0.084) | −0.043 (0.105) | 0.008 (0.139) | −0.141* (0.082) | 0.047 (0.103) | −0.254** (0.110) | 0.018 (0.065) | 0.039 (0.081) | −0.004 (0.102) | −0.020 (0.085) | 0.002 (0.108) | −0.040 (0.129) |
| Perceived physical abilities | −0.082 (0.126) | 0.098 (0.171) | −0.178 (0.165) | 0.019 (0.076) | 0.000 (0.099) | 0.048 (0.120) | −0.033 (0.084) | 0.121 (0.137) | −0.126 (0.103) | −0.088 (0.066) | −0.095 (0.090) | −0.080 (0.097) | 0.059 (0.080) | 0.084 (0.103) | 0.034 (0.121) |
| Perceived technological knowledge and skills | −0.040 (0.126) | −0.198 (0.197) | 0.043 (0.159) | 0.066 (0.066) | 0.028 (0.087) | 0.123 (0.099) | −0.066 (0.076) | −0.036 (0.111) | −0.084 (0.103) | −0.032 (0.067) | 0.006 (0.089) | −0.076 (0.099) | 0.026 (0.072) | 0.081 (0.085) | −0.025 (0.110) |
| Perceived language skills | 0.050 (0.119) | −0.075 (0.154) | 0.115 (0.160) | 0.011 (0.084) | −0.025 (0.108) | 0.065 (0.133) | 0.066 (0.083) | 0.248* (0.137) | −0.044 (0.102) | −0.022 (0.068) | 0.010 (0.096) | −0.057 (0.099) | −0.077 (0.082) | −0.079 (0.102) | −0.075 (0.123) |
| Perceived leadership skills | 0.003 (0.117) | −0.159 (0.185) | 0.091 (0.146) | 0.092 (0.075) | 0.080 (0.101) | 0.109 (0.114) | −0.142* (0.079) | −0.111 (0.116) | −0.161 (0.105) | 0.008 (0.066) | −0.012 (0.090) | 0.032 (0.097) | −0.070 (0.077) | −0.130 (0.088) | −0.012 (0.123) |
| Perceived assertiveness | 0.010 (0.138) | −0.141 (0.215) | 0.091 (0.175) | −0.001 (0.073) | 0.043 (0.088) | −0.068 (0.132) | −0.019 (0.080) | −0.032 (0.114) | −0.011 (0.106) | −0.050 (0.066) | −0.017 (0.089) | −0.088 (0.099) | −0.111 (0.077) | −0.239** (0.091) | 0.012 (0.119) |
| Perceived ambition | 0.128 (0.115) | 0.066 (0.173) | 0.163 (0.146) | 0.201*** (0.075) | 0.176** (0.088) | 0.239* (0.135) | −0.121* (0.073) | −0.100 (0.118) | −0.134 (0.090) | 0.097 (0.069) | 0.126 (0.092) | 0.064 (0.105) | 0.047 (0.076) | −0.089 (0.103) | 0.176* (0.104) |
| Perceived motivation | 0.171 (0.113) | 0.105 (0.207) | 0.208 (0.129) | 0.163** (0.076) | 0.095 (0.099) | 0.264** (0.117) | −0.039 (0.072) | 0.054 (0.086) | −0.096 (0.100) | 0.077 (0.068) | 0.034 (0.088) | 0.128 (0.108) | 0.091 (0.079) | 0.015 (0.107) | 0.162 (0.110) |
| Perceived autonomy | 0.105 (0.108) | 0.135 (0.133) | 0.089 (0.148) | 0.067 (0.074) | −0.011 (0.089) | 0.188 (0.133) | −0.083 (0.084) | −0.051 (0.119) | −0.103 (0.113) | −0.024 (0.067) | 0.041 (0.084) | −0.101 (0.108) | −0.103 (0.077) | −0.114 (0.098) | −0.093 (0.118) |
| Perceived efficiency | 0.144 (0.113) | 0.191 (0.188) | 0.122 (0.142) | 0.135* (0.070) | 0.053 (0.088) | 0.259** (0.114) | −0.024 (0.075) | −0.081 (0.107) | 0.009 (0.102) | 0.040 (0.069) | 0.074 (0.095) | −0.002 (0.099) | −0.023 (0.078) | −0.073 (0.103) | 0.023 (0.113) |
| Perceived organizational skills | 0.176 (0.115) | 0.157 (0.146) | 0.188 (0.157) | 0.170** (0.076) | 0.118 (0.089) | 0.248* (0.139) | −0.072 (0.076) | −0.137 (0.098) | −0.035 (0.105) | 0.053 (0.067) | 0.062 (0.087) | 0.041 (0.104) | −0.048 (0.080) | −0.116 (0.103) | 0.015 (0.116) |

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Perceived reliability | 0.084 | 0.113 | 0.069 | 0.140* | 0.096 | 0.207 | −0.109 | −0.106 | −0.111 | 0.032 | 0.076 | −0.019 | −0.014 | −0.010 | −0.018 |
| | (0.120) | (0.213) | (0.145) | (0.077) | (0.097) | (0.129) | (0.077) | (0.103) | (0.105) | (0.069) | (0.089) | (0.106) | (0.074) | (0.093) | (0.110) |
| Perceived trainability | 0.173* | 0.126 | 0.198 | 0.118 | 0.033 | **0.245**** | −0.148* | 0.019 | **−0.249**** | −0.047 | −0.019 | −0.078 | 0.023 | 0.010 | 0.035 |
| | (0.096) | (0.135) | (0.126) | (0.080) | (0.104) | (0.124) | (0.077) | (0.108) | (0.102) | (0.072) | (0.093) | (0.115) | (0.079) | (0.102) | (0.118) |
| Perceived flexibility | **0.245**** | 0.147 | **0.299**** | 0.133* | 0.069 | **0.229**** | −0.056 | −0.104 | −0.029 | 0.034 | 0.051 | 0.014 | −0.024 | −0.088 | 0.035 |
| | (0.106) | (0.145) | (0.140) | (0.078) | (0.105) | (0.109) | (0.084) | (0.111) | (0.115) | (0.071) | (0.096) | (0.103) | (0.077) | (0.095) | (0.116) |
| Perceived accuracy | 0.133 | 0.156 | 0.122 | 0.106 | 0.079 | 0.146 | −0.025 | 0.023 | −0.055 | 0.029 | 0.038 | 0.019 | −0.025 | −0.137 | 0.081 |
| | (0.115) | (0.206) | (0.138) | (0.075) | (0.099) | (0.121) | (0.074) | (0.107) | (0.099) | (0.063) | (0.086) | (0.094) | (0.074) | (0.094) | (0.110) |
| Perceived openness to new people and experiences | 0.201* | 0.115 | 0.246 | **0.226***** | 0.196* | **0.271**** | 0.010 | 0.122 | −0.058 | 0.006 | 0.007 | 0.006 | −0.041 | −0.040 | −0.041 |
| | (0.119) | (0.155) | (0.161) | (0.085) | (0.113) | (0.129) | (0.073) | (0.112) | (0.093) | (0.067) | (0.086) | (0.106) | (0.080) | (0.106) | (0.118) |
| Perceived respect toward authority | 0.179* | 0.134 | 0.201 | 0.151* | 0.179* | 0.107 | −0.039 | 0.046 | −0.089 | 0.063 | 0.048 | 0.081 | −0.062 | 0.013 | −0.133 |
| | (0.104) | (0.153) | (0.137) | (0.087) | (0.108) | (0.147) | (0.080) | (0.120) | (0.105) | (0.072) | (0.096) | (0.109) | (0.089) | (0.118) | (0.132) |
| Perceived sense of responsibility | 0.108 | 0.033 | 0.148 | 0.138* | 0.118 | 0.169 | −0.119* | −0.106 | −0.127 | 0.069 | 0.113 | 0.018 | −0.005 | −0.075 | 0.062 |
| | (0.104) | (0.115) | (0.146) | (0.073) | (0.089) | (0.126) | (0.069) | (0.098) | (0.090) | (0.063) | (0.085) | (0.095) | (0.078) | (0.098) | (0.119) |
| Employer's attitude toward collaboration | 0.038 | −0.099 | 0.112 | 0.090 | 0.074 | 0.115 | 0.052 | 0.101 | 0.022 | 0.066 | 0.134 | −0.013 | 0.040 | −0.107 | 0.179 |
| | (0.111) | (0.150) | (0.147) | (0.072) | (0.095) | (0.116) | (0.070) | (0.100) | (0.092) | (0.067) | (0.090) | (0.099) | (0.076) | (0.094) | (0.113) |
| Employees' predicted attitude toward collaboration | 0.070 | −0.201 | 0.213 | −0.045 | −0.089 | 0.022 | −0.008 | 0.067 | −0.054 | −0.034 | 0.026 | −0.104 | −0.073 | −0.082 | −0.065 |
| | (0.114) | (0.149) | (0.147) | (0.075) | (0.097) | (0.118) | (0.070) | (0.099) | (0.092) | (0.066) | (0.086) | (0.099) | (0.080) | (0.093) | (0.128) |
| customers' predicted attitude toward collaboration | 0.071 | −0.102 | 0.163 | 0.045 | −0.007 | 0.122 | −0.074 | 0.024 | −0.133 | −0.032 | 0.012 | −0.084 | −0.131 | −0.172* | −0.092 |
| | (0.097) | (0.150) | (0.122) | (0.073) | (0.098) | (0.107) | (0.074) | (0.096) | (0.102) | (0.069) | (0.096) | (0.102) | (0.080) | (0.103) | (0.118) |

Note: We present coefficient estimates for the 23 distinct linear regression models outlined in Sections 4.1 and 4.2 in which we regress the standardized version of (i) the invitation probability (Panel A) and (ii) the stereotypes and attitudes toward race-ethnicity minority groups (Panel B) based on each candidate's ethnic background. Each row of this table, thus, contains the results from one regression analysis. Standard errors are presented in parentheses and corrected for the clustering of observations at the participant level. *, ** and *** indicate significance at the 10%, 5% and 1% levels. The coefficients related to $p$-values below 5% are in bold.