Lowering the Playing Field:

Discrimination through Contrast Effects

Judd B. Kessler Corinne Low Xiaoyue Shan

April 2022

Abstract

We document a new source of discrimination in hiring that arises through contrast effects. We analyze data from an incentivized resume rating experiment, where employers evaluate a sequence of randomly generated resumes to be matched with real job seekers. Candidates who follow white men are rated as significantly less desirable than those who follow women or minority candidates. Exploring the mechanisms, we find that employers in our data only display a direct bias in favor of white men when resumes are high-quality, and the contrast effect only arises when the prior resume is a low-quality white man. Our results thus suggest that the contrast effect only arises when employers do not show direct favoritism towards white men in ratings; in these cases, they indirectly favor white men through a contrast effect channel instead.

1 Introduction

Human agents are biased in decisions ranging from consumption and investment to friend-making, voting, and hiring (DellaVigna 2009; O'Donoghue and Rabin 2015; Nickerson 1998; Cohen 1981). A large literature documents gender, racial, and age bias in hiring (Bertrand and Mullainathan 2004; Bertrand and Duflo 2017; Riach and Rich 2002; Neumark, Burn, and Button 2016). More recent studies also investigate strategies that could mitigate these biases, such as diversity training, incentives, behavioral nudges, and quotas (Devine et al. 2012; Beaurain and Masclet 2016; Bhavnani 2017; Bezrukova, Jehn, and Spell 2012).

While these strategies may help agents mitigate direct displays of bias, such bias may manifest indirectly through other channels. This paper documents a new channel through which recruiters display bias that favors white men. We find that after evaluating a white man, recruiters give a lower rating to the next candidate who they evaluate. We call this sequential spillover a contrast effect.

We identify and explore this new channel for bias using data from the original incentivized resume rating (IRR) experiment (Kessler, Low, and Sullivan 2019). The IRR experiment invited employers recruiting graduating seniors at the University of Pennsylvania (Penn) to evaluate a sequence of 40 hypothetical resumes, the components of which (e.g., name, GPA, work experiences) were individually randomized for each resume for each employer. The incentive provided for employers was a list of 10 actual Penn students that were predicted to be good matches for an employer based on their reported preferences. Participating employers were therefore incentivized to truthfully and accurately reveal their preferences in order to receive the most desirable matches for their job opening. The randomization scheme in the IRR experiment allows for the identification of demographic biases through the name associated with the resume. Kessler, Low, and Sullivan (2019) found that, overall, employers' ratings of the desirability of white men were directionally

higher than—but not statistically significantly different from—the desirability ratings of minority or female candidates.[1]

This paper further interrogates the employer ratings from that paper and leverages the fact that—since resume characteristics were randomized for each of the 40 resumes shown to each employer—the data also allows for clean identification of the impact of the *prior* resume's characteristics on the rating of the current candidate. In new analysis, we find that resumes placed after white men are rated statistically significantly worse than those that follow women or minorities. The contrast effect is large. Resumes following white men are rated 4% lower than statistically identical resumes following female or minority candidates (or lower by 7% of a standard deviation in ratings of all resumes). Employers prefer candidates with higher GPAs, and the negative impact of following a white male candidate is equivalent to having a GPA that is 0.1 points lower (e.g., going from a 4.0 to a 3.9). The contrast effect does not significantly differ by the demographics of the current resume, suggesting that it is not a conscious effort to favor white men but, rather, that ratings are uniformly lower when the employer has just evaluated a white man.

To understand this puzzling result, we explore whether a certain subset of the resumes of white men generate this contrast effect. Informed by results from Kessler, Low, and Sullivan (2019), which found that white men received a larger increase in ratings than women or minorities from having a prestigious internship on their resume—a result that itself echoed the findings in Bertrand and Mullainathan (2004)—our main approach is to classify resumes as either high-quality or low-quality, collapsing over the various characteristics that make a resume desirable for employers.

High-quality resumes randomly assigned the names of white men receive statistically significantly higher ratings than those assigned the names of women and minorities. In

[1]That paper also found a statistically significant preference in favor of white men among the subset of employers who were recruiting students with majors in Science, Technology, Engineering, and Math (STEM) fields.

addition, high-quality resumes lead to no contrast effect: resumes that follow high-quality white men are not rated worse than resumes that follow high-quality women or minorities. When the resume is low quality, however, there is no direct preference for white men, and these are the white men who drive the entire contrast effect. Resumes that follow low-quality white men are rated 8% lower than statistically identical resumes following low-quality women and minorities (14% of a standard deviation lower, equivalent to 0.17 GPA points).

The explanation that best fits our pattern of results is that employers have an implicit bias in favor of white men, which they allow to come out when they rate high-quality resumes. When they rate low-quality resumes, however, the direct bias is constrained, but it then manifests through a contrast effect. Employers "lower the playing field" such that the next candidate is rated more harshly, thus making the low-quality white man look better by comparison.

Our paper contributes to the decades-long literature that investigates discrimination in the labor market, especially in the hiring process (Becker 1971; Heckman 1998; Bertrand and Mullainathan 2004; Neumark 2018). Consistent with prior work, we find evidence of discrimination that is conditional on candidate quality. In addition, we show that even when employers might appear to constrain their bias (e.g., by not favoring low-quality white men over low-quality women and minorities), this bias can spill over in unexpected ways, thus raising new equity issues.[2]

[2]This paper also relates to a rich literature on contrast effects. Bhargava and Fisman (2014) find that in speed dating, prior partner attractiveness lowers male evaluators' likelihood to date the current target. Hartzmark and Shue (2018) find that investors' perception of today's earnings news is negatively affected by yesterday's earnings surprises. Radbruch and Schiprowski (2020) show that interview panels' evaluation of candidates for a study grant program decreases with prior candidates' quality.

2 Experimental Design and Data

2.1 Design

Our data comes from the original IRR experiment, described in Kessler, Low, and Sullivan (2019). The experiment was run at the University of Pennsylvania during the 2016–2017 academic year in collaboration with the Penn Career Services office.

The program invited employers to evaluate hypothetical resumes, identified their preferences for candidates, and then recommended to them actual graduating seniors at Penn who were looking for jobs. Participating employers each evaluated 40 hypothetical resumes with randomly assigned candidate characteristics (e.g., name, GPA, major), including curated components from real Penn resumes (e.g., real work experiences and leadership experiences).[3] In addition, resumes were assigned a name that was indicative of race and gender to allow for the exploration of discrimination. Employers rated each hypothetical candidate on two dimensions: their interest in hiring the candidate and the likelihood that the candidate would accept the job if offered one. Appendix Table A.1 and Appendix Figure A.1, both reproduced from Kessler, Low, and Sullivan (2019), show the variation introduced into the resumes in building the survey platform and an example of a hypothetical resume that also shows the question wording. The incentive for the employers was getting to be matched with 10 real graduating seniors at Penn who were looking for jobs and had uploaded their resumes. These matches were based on each employer's preferences for resume characteristics.

In this paper, we analyze data from the original IRR experiment. We focus on hiring interest, which was measured with the following question:

[3]To improve preference elicitation, at the beginning of the survey employers were asked whether they were looking for candidates with "Business (Wharton), Social Sciences, and Humanities" majors or "Science, Engineering, Computer Science, and Math" majors. The candidates they evaluated were then limited to those with related majors and work experiences.

"How interested would you be in hiring [**Name**]?"

Responses were on a 10-point Likert scale, where 1 was "Not interested" and 10 was "Very interested".

2.2 Data

Our dataset includes 72 employers' ratings of 2,880 hypothetical resumes. The employers come from a wide range of industries, including consulting, finance, technology, retail, education, and the non-profit sector. Participating firms also vary in size: about 30% have less than 50 employees, 20% have 50 to 999 employees, and the remaining 30% have 1,000 employees or more. Most of the employers in our sample (70%) are looking for candidates with business, social sciences, or humanities backgrounds; the rest are interested in candidates with a STEM background. In survey data collected after the resume rating exercise, 90% of employers say they consider seeking racial or gender diversity as a factor in their rating of candidates.[4]

To identify contrast effects, we analyze the ratings of 2,808 resumes (i.e., 39 resumes per employer). We exclude the first resume that each employer rates, since there is no prior resume to influence ratings. For these 2,808 resumes, the dependent variable—the rating of hiring interest on the 1–10 scale—has an average value of 4.7 and a standard deviation of 2.6. The main variable of interest in the contrast effect analysis is whether a prior resume had the name of a white man: 32.85% of resumes were assigned the name of a white man, and 67.15% were assigned a name that was indicative of a white woman (32.85%), non-white woman (17.15%), or non-white man (17.15%).[5]

[4]For more details on employers and their survey responses, see Kessler, Low, and Sullivan (2019)

[5]More details on these variables, and the races of the non-White candidates can be found in Appendix Table A.1.

3 Results

3.1 Specification

We use the following regression specification to estimate the contrast effect of following a white man:

$$R_{ij} = \beta N_{i,j-1}^{wm} + \boldsymbol{\gamma_1} N_{ij} + \boldsymbol{\gamma_2} Q_{ij} + \alpha_i + \varepsilon_{ij}. \quad (1)$$

The dependent variable R_{ij} is the rating of hiring interest on the 10-point Likert scale given by employer i about resume j (where $j \in \{2, 3, ..., 40\}$ denotes the order, out of 40, in which the resume was shown). $N_{i,j-1}^{wm}$ is the key variable of interest. It is equal to 1 when the name on the prior resume (resume $j-1$) shown to employer i was indicative of a white man and is 0 otherwise. The regression also controls for the race and gender, N_{ij}, and quality characteristics, Q_{ij}, of the current resume. N_{ij} are dummies for whether the resume has the name of a white woman, a non-white woman, or a non-white man (i.e., white men are the excluded group). In the baseline specification, the quality characteristics include GPA, whether the first work experience is a top or prestigious internship, whether the candidate also has a second internship, whether the candidate has a non-internship "work-for-money" job, and whether the resume has technical skills listed.[6] The regressions always control for employer fixed effects, α_i. Additional specifications also include fixed effects for the college major of the resume, fixed effects for the leadership experiences (i.e., extracurricular activities) on the resume, and fixed effects for the order in which the resume was shown. In additional specifications, we also control for measures of the prior resume's quality, $Q_{i,j-1}$.

Two features of the IRR survey tool ensure the causal identification of the contrast effect β: (1) orthogonal relationships between resume components and (2) the randomized order of resumes. Because all resume components (including demographic and quality indicators)

[6] For more details, see Appendix Table A.1.

were independently and randomly drawn, the measured effect of demographics could not be driven by the possible correlations between demographics and other characteristics. Because the resume contents were randomly populated for each of the 40 resumes, a resume's demographic or quality characteristics are orthogonal to the next resume's characteristics.

3.2 Identifying the Contrast Effect

Table 1 shows the contrast effect using different specifications of equation (1). The dependent variable is the rating of hiring interest on a scale of 1 to 10. In column (1), we estimate the coefficient of $N^{wm}_{i,j-1}$ ("After White Man") controlling for resume quality and demographic indicators and subject fixed effects. In columns (2)–(4), we gradually include major fixed effects, leadership experience fixed effects, and resume order fixed effects. In column (5), we further control for the prior resume's quality indicators ($Q_{i,j-1}$) to rule out potential contrast effects based on resume quality, independent of demographics. All estimations in Table 1 use robust standard errors, but results are very similar when we cluster standard errors at the subject level (see Appendix Table A.2).

The results show a strong and robust negative effect of the prior resume being a white man on the rating of the current resume. Using the fully controlled specification, we find that being placed after white man—rather than after a female or minority candidate—lowers the current rating by 0.19 Likert points, 4% of the average rating of resumes not following white man, or 7% of a standard deviation in ratings of all resumes. Comparing this estimate to the estimate on *GPA* suggests that being after a white man is almost equivalent to having a GPA that is 0.1 points lower.

The table also shows that academic ability and work experience significantly impact employers' ratings, as is also highlighted in Kessler, Low, and Sullivan (2019). For example, employers are much more interested in candidates with higher GPAs, higher quality work experiences, and more work experiences. They are relatively less interested in students who

Table 1: Contrast Effect Regressions

	(1)	(2)	(3)	(4)	(5)
Dependent Variable: Hiring Interest					
After White Man	-0.174**	-0.198**	-0.182**	-0.189**	-0.188**
	(0.083)	(0.082)	(0.084)	(0.085)	(0.085)
GPA	2.077***	2.108***	2.151***	2.166***	2.170***
	(0.126)	(0.126)	(0.129)	(0.131)	(0.131)
Top Internship	0.932***	0.924***	0.909***	0.907***	0.911***
	(0.079)	(0.079)	(0.081)	(0.082)	(0.082)
Second Internship	0.423***	0.450***	0.458***	0.455***	0.451***
	(0.092)	(0.092)	(0.095)	(0.096)	(0.096)
Work for Money	0.120	0.127	0.164*	0.165*	0.162*
	(0.091)	(0.090)	(0.092)	(0.092)	(0.092)
Technical Skills	-0.076	-0.058	-0.065	-0.062	-0.062
	(0.088)	(0.088)	(0.090)	(0.091)	(0.091)
White Woman	-0.105	-0.100	-0.150	-0.154	-0.155
	(0.094)	(0.094)	(0.096)	(0.097)	(0.097)
Non-White Woman	0.002	0.020	0.010	0.021	0.020
	(0.117)	(0.116)	(0.119)	(0.121)	(0.121)
Non-White Man	-0.163	-0.138	-0.163	-0.151	-0.155
	(0.114)	(0.113)	(0.116)	(0.117)	(0.117)
Prior GPA					0.008
					(0.121)
After Top Internship					0.047
					(0.081)
After Second Internship					-0.048
					(0.096)
After Work for Money					-0.025
					(0.094)
After Technical Skills					0.056
					(0.092)
Observations	2,808	2,808	2,808	2,808	2,808
R-squared	0.427	0.446	0.481	0.489	0.489
Subject fixed effects	Yes	Yes	Yes	Yes	Yes
Major fixed effects	No	Yes	Yes	Yes	Yes
Leadership fixed effects	No	No	Yes	Yes	Yes
Order fixed effects	No	No	No	Yes	Yes

Notes: The sample includes 2,808 employer ratings, excluding the first resume reviewed by each employer. Robust standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

spend their summer working for money instead of accumulating human capital through an internship. As shown in column (5), however, these quality indicators do not create contrast effects.[7]

4 Mechanisms Driving the Contrast Effect

Why should having just evaluated a white man systematically lower employers' ratings of the subsequent resume? In this section, we explore potential mechanisms driving this contrast effect. Section 4.1 examines whether the contrast effect varies with prior resume quality (i.e., the quality of resume $j-1$). Section 4.2 examines whether the contrast effect varies with the demographics and quality of the current resume (i.e., the race, gender, and quality of the candidate in resume j). Section 4.3 examines whether the contrast effect varies with the demographics of the employer and the extent to which racial or gender diversity is considered as an important factor in the rating of candidates. In section 4.4, we study whether the contrast effect varies with the employer's industry. Section 4.5 explores the dynamics of the contrast effect over multiple resumes. In Section 4.6 we use the results in this section to suggest a possible explanation for how the contrast effect operates.

4.1 The role of prior resume quality

Evidence from Kessler, Low, and Sullivan (2019) suggests that while there is no statistically significant preference for white men in the ratings data overall, white men are rewarded more than candidates who are not white men for having secured a prestigious internship. This finding relates to results in Bertrand and Mullainathan (2004), which famously found that whites received a higher return to improving resume quality in a resume audit.

Here, we further explore whether resume quality interacts with demographic preferences

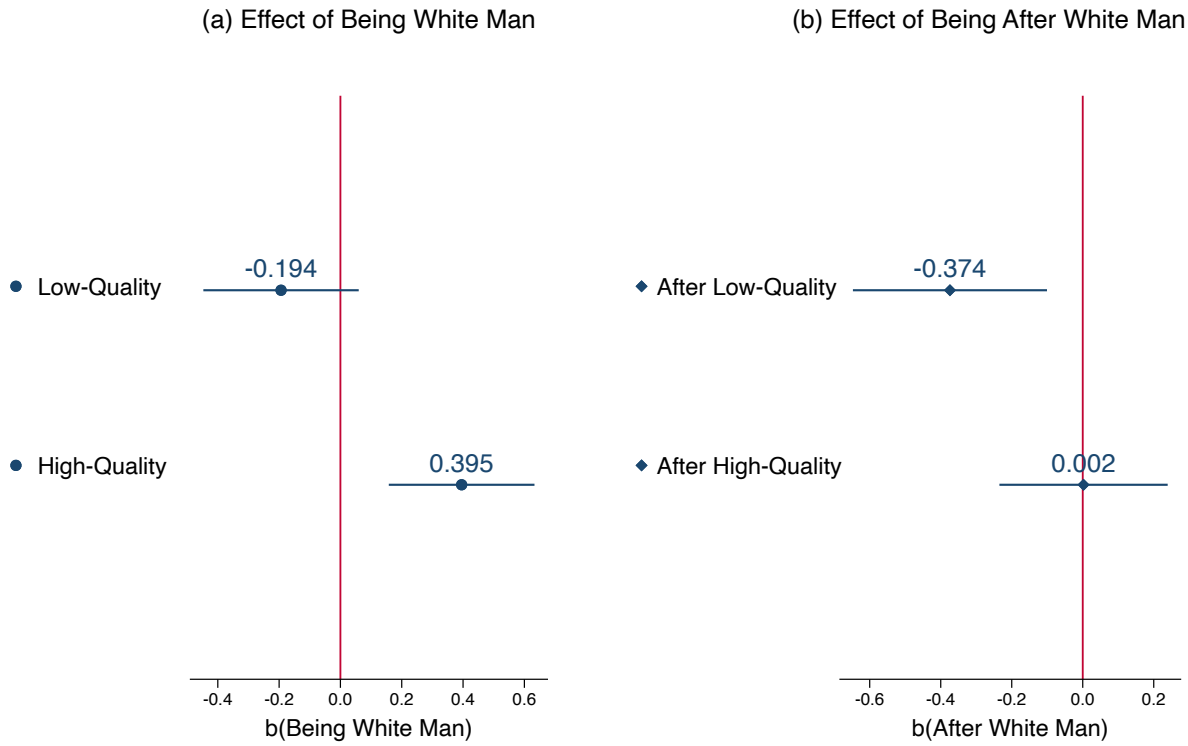
[7] In Appendix Table A.3, we test for quality contrast effects when separately using different quality indicators. Consistently, we find no evidence that previous resume's quality affects current resume's rating.

and the contrast effect. Rather than looking exclusively at the returns to having a prestigious internship, we use a data-driven approach to identify which resume characteristics—excluding race and gender—are associated with a resume receiving high ratings from employers. We then classify resumes as being "high quality" (the top 56% of resumes) or "low quality" (the bottom 44% of resumes) based on this predicted score and explore whether the contrast effect varies with this measure of resume quality.[8]

In Figure 1(a), we test whether employers give higher ratings to white men, both for low-quality resumes and for high-quality resumes. As shown in the bottom figure of panel (a), white men receive higher ratings than women and minorities when their name is randomly assigned to a high-quality resume. For statistically identical resumes, white men are rated about 0.4 Likert points higher than women and minorities. The gap is about 8% of the average rating of high-quality women and minorities (15% of a standard deviation in the ratings of all resumes). These results imply that employers are biased in favor of white men and display this bias in ratings when resume quality is high. This preference is somewhat offset by a directionally negative, although not statistically significant, reduction in ratings for white men when employers evaluate the 44% of resumes that are identified as low-quality. The estimates of being a white man on resumes ratings are statistically significantly different across high-quality and low-quality resumes, highlighting that the way employers treat white men relative to women and minorities depends fundamentally on resume quality.

[8]In particular, we use Lasso to identify the best predictors of resume ratings from the resume quality characteristic variables in our data: GPA dummies (i.e., GPA rounded to the nearest 0.1), dummies for work experiences (i.e., top internship, second internship, work-for-money job), and technical skills. We use adaptive Lasso to select the λ parameter. We then predict resume ratings based on the algorithm-chosen predictors. The distribution of predicted ratings display a few big clusters and a few small ones—which we combine into 7 quality groups. We then estimate the preference for white men in ratings for each quality group. As Appendix Figure A.2 shows, employers give white men directionally lower ratings for all quality groups in the bottom 44% of resumes and give white men directionally higher ratings for all quality groups in the top 56%. We therefore identify 44% as a natural cutoff to divide resumes into low and high quality. The ranking of predicted resume quality is almost identical when we use alternative prediction methods, like OLS, or alternative λ-choosing methods (e.g., cross-validation).

Figure 1: The Effects of Being (After) a White Man by Resume Quality



Notes: Figure (a) shows the effect of being a white man on employer ratings, separately for low-quality and high-quality resumes. Figure (b) shows the impact of being rated after a white man on employer ratings, separately for being after a low-quality white man and after a high-quality white man. All estimations use a sub-sample of the 2,808 observations and control for all the fixed effects and quality indicators in column (4) of Table 1. Figure (a) only uses the white man demographic indicator comparing white men to all women and minorities, and Figure (b) includes the same demographic indicators as in column (4) of Table 1. The differences between point estimates in two sub-figures are statistically significant: p-value$<$0.001 in panel (a) and p-value=0.02 in panel (b). Error bars indicate 95% confidence intervals.

In Figure 1(b), we test whether the impact of being after a white man varies with whether the white man's resume is high-quality or low-quality. We find that the contrast effect is entirely driven by following low-quality white men. Conditional on following a low-quality resume, ratings are 0.37 Likert points (8% of the average rating of resumes following low-quality women or minorities, 14% of a standard deviation in all ratings, equivalent to 0.17 GPA points) lower if the prior candidate was a white man rather than a woman or minority, a magnitude that is almost the same as the partiality displayed toward

high-quality white men in panel (a). Meanwhile, there is no contrast effect when resumes follow high-quality white men.[9] Indeed, the estimate of -0.374 of following a low-quality white man is statistically significantly smaller than the estimate of 0.002 of following a high-quality white man.

Taken together, Figure 1 suggests that the contrast effect relates to whether employers display partiality toward white men when rating the prior resume. It seems that employers only lower ratings of the subsequent resume when they did not show bias in favor of the white man they just evaluated. In the subsequent sections, we therefore pay special attention to the contrast effect from following a low-quality white man.[10]

4.2 The role of the demographics and quality of the current resume

In this section, we examine whether the contrast effect interacts with the demographics and quality of the current resume. Panel A of Table 2 shows the size of the contrast effect when the current resume (i.e., resume j) is a white man and otherwise, or when the current resume is high quality and otherwise. Columns (1) and (2) show that the contrast effect does not statistically significantly differ by whether the current resume is a white man. Column (1) shows the overall contrast effect while column (2) focuses on the contrast effect following a low-quality white man, given the results in Section 4.1. Columns (3) and (4) show that the contrast effect has directionally smaller magnitude when the current resume is high quality, but the difference is not statistically significantly. These results suggest that the documented contrast effect is not driven by a direct favoritism of white men (which is somewhat expected, given that we do not see an overall preference for white men in

[9] Appendix Figure A.3 shows the same results in Figure 1 with bar graphs showing the residualized ratings for white men and for women and minorities separately (i.e., showing the means that generate the differences shown in Figure 1).

[10] As a placebo test, we also examine whether being placed *before* (low-quality) white men affects resume ratings. We find no such effects: the estimated coefficient of "before white man" is 0.038 (p-value=0.66), and the estimated coefficient of "before low-quality white man" is 0.006 (p-value=0.96).

employer ratings). Instead, evaluating a low-quality white man leads to a lowering of the rating of *any* resume that follows, a phenomenon that we call "lowering the playing field." It implicitly gives the prior low-quality white man a boost by lowering the ratings of anyone who follows them.

4.3 The role of employer demographics and recognized importance of diversity

We also test whether the contrast effect is driven by a certain subset of employers. In a survey asked after employers rated all 40 resumes, they reported their own race and gender: 32% reported that they were white men. The survey also asked employers to what extent they considered "seeking to increase gender/racial diversity" as a factor in their rating of candidates. Employers answered the question on a scale of 1 ("do not consider at all") to 10 ("this is among the most important things I consider"). Less than 10% of them chose 1 as the answer, meaning that most employers in the setting consciously wanted to hire more women or minority candidates. We use this as a measure of employers' recognized importance of increasing diversity in their hiring and define the above-medium answers as "high importance".

As Panel B of Table 2 show, the contrast effect does not significantly differ with whether or not the employer evaluating the resume is a white man. Point estimates suggest that white men display a directionally smaller contrast effect. More generally, we find no evidence that the effect varies with employer gender or race. Also, the contrast effect does not significantly differ between employers who consider increasing diversity has low and high importance. The results suggest that the contrast effect mainly work through an unconscious mechanism; even for employers who consciously and actively seek to increase racial and gender diversity in their hiring process, the contrast effect is similarly strong.

Table 2: Heterogeneity in the Effect of Being After a (Low-Quality) White Man

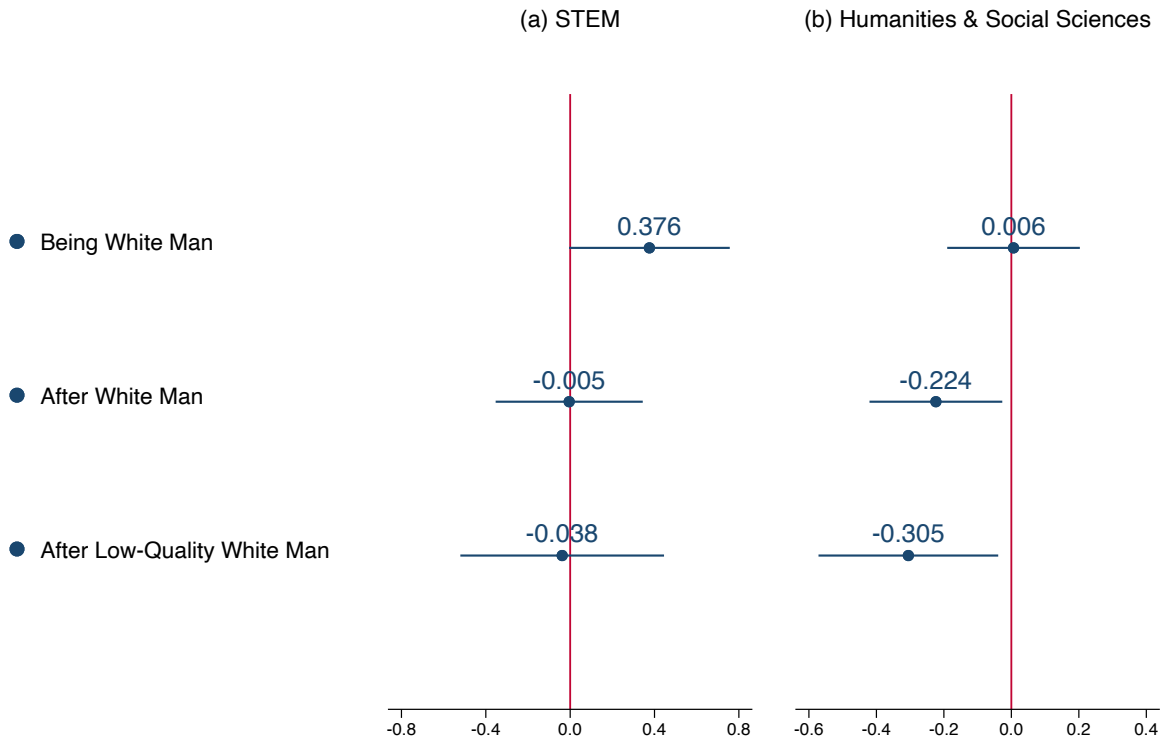
Dependent Variable: Rating of Hiring Interest	(1)	(2)	(3)	(4)
Panel A: By current resume's demographics & quality				
After White Man	-0.171*		-0.250**	
	(0.101)		(0.123)	
After Low-Quality White Man (LQWM)		-0.326**		-0.447***
		(0.140)		(0.161)
After White Man × White Male Resume	-0.056			
	(0.184)			
After LQWM × White Male Resume		0.096		
		(0.243)		
After White Man × High-Quality Resume			0.101	
			(0.170)	
After LQWM × High-Quality Resume				0.259
				(0.226)
High-Quality Resume			0.248	0.248*
			(0.155)	(0.149)
Panel B: By employer's demographics & reported importance of diversity				
After White Man	-0.217**		-0.172	
	(0.102)		(0.112)	
After Low-Quality White Man		-0.354***		-0.330**
		(0.134)		(0.150)
After White Man × White Male Employer	0.087			
	(0.185)			
After LQWM × White Male Employer		0.181		
		(0.256)		
After White Man × High Importance of Diversity			-0.039	
			(0.171)	
After LQWM × High Importance of Diversity				0.087
				(0.231)

Notes: All estimations include the control variables specified in column (4) of Table 1. Robust standard errors are in parentheses. *After LQWM* is the estimated impact on ratings of following a low-quality white man. Note that most of these regressions do not report the direct effect of a white male resume or of a white male employer because these are already absorbed by the fixed effects included in the regression specification from column (4) of Table 1. The *High-Quality* indicator is not perfectly collinear with resume quality characteristics, so it still shows up in the results. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

4.4 The role of industry

Finally, we test whether the employer's industry matters for the display of contrast effects. As documented in Kessler, Low, and Sullivan (2019), employers looking for candidates with science, technology, engineering, mathematics (STEM) backgrounds display a more salient preference for white men, while employers hiring candidates with humanities and social sciences (including business) backgrounds do not show a direct bias toward white men. Guided by our results in Section 4.1, we further explore whether the contrast effect is driven by employers in humanities and social sciences industries—because they do not display direct biases in their ratings. Indeed, we find results consistent with that.

Figure 2: Preference for White Men and the Contrast Effect by Industry



Notes: The figure shows the estimated effects of being a white man and being after a (low-quality) white man on resume ratings by the employer's industry type. Each point estimate is derived from one regression. The top two regressions use the same specification as in Figure 1(a), and the bottom four regressions use the same specification as in Figure 1(b). Error bars indicate 95% confidence intervals.

Figure 2 shows the estimated preference for white men and the contrast effect generated by (low-quality) white men by employers' industry type. We find that STEM employers on average rate white men higher than minority or female candidates by 0.38 Liker points. These employers display no significant contrast effect in their evaluation of resumes: resumes placed after white men and resumes placed after other candidates receive similar ratings. By contrast, employers in humanities and social sciences industries do not give white men explicitly higher ratings; only for these employers, we find strong and statistically significant contrast effects.

4.5 Dynamics of the contrast effect

All of our previous analysis has focused on the effect of *immediately* following a white man. To better understand the contrast effect, we examine its dynamics. In particular, we explore whether the effect of following a white man also has longer-lasting impacts on subsequent resume ratings.

In Table 3, we explore whether the demographics of resume $j-2$ (i.e., the resume before the prior resume) still has an impact on the current resume rating, once we account for the impact of resume $j-1$ (i.e., the prior resume). Column (1) estimates the impact of following a white man in resume $j-1$ (i.e., WM_{j-1}) and of following a white man in resume $j-2$ (i.e., WM_{j-2}). While the former is statistically significant, the latter is directionally negative and not significant. Column (2) separately identifies the impact of following one or more white men in the three possible cases in which at least one of the prior two resumes is a white man. It shows that the case when both of the past two resumes are white men has the largest negative impact on resume ratings; the coefficient on $(WM_{j-2}, WM_{j-1}) =$ (Yes, Yes) suggests that resumes are rated 0.35 Likert points (14% of a standard deviation) lower when the last two resumes were white men than when the last two resumes were women or minorities (the excluded group in the regression). However, this coefficient is not

statistically significantly different from the coefficient on (WM_{j-2}, WM_{j-1}) = (No, Yes), which estimates that ratings are 0.13 Likert points (5% of a standard deviation) lower when the prior resume is a white man and the resume before that was not. In addition, the coefficient on (WM_{j-2}, WM_{j-1}) = (Yes, No) is effectively 0, highlighting that if the prior resume is not a white man, the fact that resume $j-2$ was a white man has no independent impact.

Columns (3) and (4) replicate this analysis focusing on low-quality white men—estimating the effect of following a low-quality white man compared to following anyone else. The pattern of results looks very similar to columns (1) and (2) but with more-negative coefficient estimates, since the contrast effect is driven by following low-quality white men. Taken together, the results suggest that the impact of following a low quality white man is short lived.[11]

4.6 How the contrast effect may operate

The results above document a contrast effect that manifests after an employer has evaluated the resume of a low-quality white man. Any resume that follows a low-quality white man is rated significantly more harshly—on the order of 8% worse—than if the resume had followed a low-quality woman or minority. The contrast effect does not depend on the race or gender of the subsequent resume—it is just as large for white men as for women and minorities. Consequently, we say that the contrast effect "lowers the playing field," imposing an equal penalty on anyone that follows the low-quality white man. In addition, the contrast effect has limited persistence; it influences the evaluation of the next resume

[11] We limit our analysis of the dynamics of the contrast effect to the prior two resumes for reasons of statistical power. Across all of our data, we only observe 58 pairs of low quality white men back-to-back in our data. For similar power reasons, we do not extensively analyze the effects of exposure to white male resumes in positions $j-3$ and earlier. That said, in Appendix Figure A.4, we show the estimated independent effect of following a (low-quality) white man in resume $j-5$ through resume $j-1$. Similar to the results in Table 3, we find no evidence that having been exposed to (low quality) white men earlier than resume $j-1$ affects the rating independently.

Table 3: The Duration of the Effect of Being After a (Low-Quality) White Man

Dependent Variable: Hiring Interest	(1)	(2)	(3)	(4)
WM_{j-2} = Yes	-0.073			
	(0.088)			
WM_{j-1} = Yes	-0.198**			
	(0.086)			
WM_{j-2}, WM_{j-1} = No, Yes		-0.130		
		(0.103)		
WM_{j-2}, WM_{j-1} = Yes, No		-0.005		
		(0.106)		
WM_{j-2}, WM_{j-1} = Yes, Yes		-0.352**		
		(0.144)		
$LQWM_{j-2}$ = Yes			-0.030	
			(0.122)	
$LQWM_{j-1}$ = Yes			-0.307***	
			(0.117)	
$LQWM_{j-2}, LQWM_{j-1}$ = No, Yes				-0.292**
				(0.126)
$LQWM_{j-2}, LQWM_{j-1}$ = Yes, No				-0.016
				(0.130)
$LQWM_{j-2}, LQWM_{j-1}$ = Yes, Yes				-0.404
				(0.304)
Observations	2,736	2,736	2,736	2,736
R-squared	0.491	0.491	0.491	0.491

Notes: All regressions use the specification as column (4) of Table 1. Because we examine the impact of two resumes prior, we limit the sample to the resumes 3–40 that the employers rate. $(LQ)WM_{j-2}$ = Yes means that the resume before the prior resume is a (low-quality) white man. $(LQ)WM_{j-1}$ = Yes means that the prior resume is a (low-quality) white man. Robust standard errors are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

but does not have a longer lasting impact. However, there is some evidence that it can compound, as the contrast effect is directionally bigger after the employer has just evaluated two low-quality white men back-to-back.

When considering the causes of the contrast effect, we find it important to consider the direct favoritism employers display, and the favoritism they do not display. Employers directly favor high-quality white men; conditional on a resume being high quality, it gets higher ratings when it is assigned the name of a white man than when it is assigned the

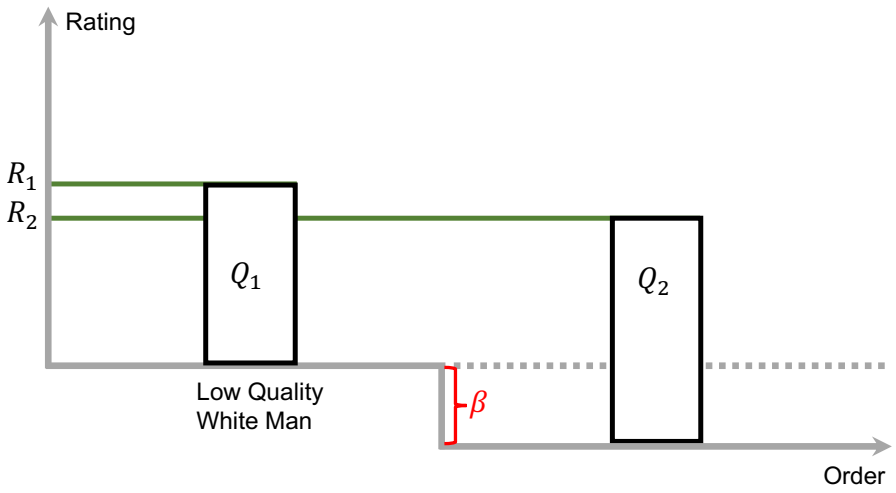
name of a woman or minority. However, employers *do not* display favoritism towards low-quality white men. It is particularly striking that the contrast effect only arises after the white men that the employer does not directly favor.

Pulling these insights together, the explanation that best fits our pattern of results is that employers have an implicit bias in favor of white men, which comes out when they rate high-quality resumes. When they rate low-quality resumes, however, this implicit bias is somehow constrained—perhaps because the candidate is obviously unremarkable or because the employer is disappointed that the candidate is not as impressive as they expect. However, the constrained implicit bias still manifests to indirectly favor the low-quality white men through a contrast effect. Employers "lower the playing field" by rating the next candidate more harshly, thus making the low-quality white man look somewhat better by comparison.

As noted in the prior paragraph, our evidence suggests that the direct favoritism and the indirect favoritism through a contrast effect are both forms of implicit bias. First, employers report valuing diversity in their recruiting. Second, if they wanted to display direct explicit bias, they might instead do that by uniformly rating the resumes of white men more highly.[12] Third, if they wanted to explicitly favor a specific low-quality white man through a contrast effect, it would need to be longer lasting (e.g., rating everyone else lower to make the low-quality white man look better by comparison); the short-term nature of the contrast effect implies that it is driven by a psychological urge rather than a rational response.

[12]It is also possible that employers may intuit that there is no benefit to displaying explicit bias in the incentivized resume rating paradigm, since we do not use employers' demographic preferences when identifying which real candidates we match to them, we instead only use the preferences we identify about other resume characteristics.

Figure 3: An Example of Lowering the Playing Field



Notes: In this example, the first resume is a low-quality white man, whose quality is lower than the next resume: $Q_1 < Q_2$. However, due to the contrast effect (β), the second resume receives a lower rating than the low-quality white man: $R_2 < R_1$.

Despite being short-lived, however, such a contrast effect could still have big impacts. In a setting with two candidates, as shown in Figure 3, a contrast effect (of size β) could lead the employer to favor a low-quality white man over a somewhat more impressive candidate who follows the low-quality white man but cannot overcome the contrast effect.

5 Conclusion

This paper documents a new, subtle channel through which employers display bias in resume rating. We leverage data that randomizes the components of a series of 40 resumes—including a randomized name indicative of race and gender—shown to employers who have an incentive to rate the desirability of the 40 candidates. We observe that employers display a preference for resumes randomly assigned the names of white men when the underlying resume components suggest the candidate is of high quality. We observe no such preference when the underlying resume components suggest the candidate is low quality. However, we find that while employers fail to show partiality toward these low-quality white men, after

evaluating a low-quality white man, they lower their rating of the subsequent candidate. This effect creates a large gap in the ratings received by candidates following low-quality white men and those following women or minorities. Our findings suggest the power of implicit bias; even when no direct bias is shown towards low-quality white men, the bias manifests in unexpected ways, benefiting the low-quality white men indirectly through the contrast effect.

Our data—with 40 ratings per employer of randomly generated resumes—provides an ideal environment to cleanly identify and explore a contrast effect. Future research should explore the presence of such contrast effects in other settings where such bias may operate indirectly. Better understanding the psychological causes of such biases, and how biases may manifest in unexpected ways, can be an important step to help mitigate them.

References

Beaurain, Guillaume, and David Masclet. 2016. "Does affirmative action reduce gender discrimination and enhance efficiency? New experimental evidence." *European Economic Review* 90:350–362.

Becker, Gary. 1971. *The Economics of Discrimination.* Technical report. University of Chicago Press.

Bertrand, Marianne, and Esther Duflo. 2017. "Field experiments on discrimination." *Handbook of Economic Field Experiments* 1:309–393.

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

Bezrukova, Katerina, Karen A Jehn, and Chester S Spell. 2012. "Reviewing diversity training: Where we have been and where we should go." *Academy of Management Learning & Education* 11 (2): 207–227.

Bhargava, Saurabh, and Ray Fisman. 2014. "Contrast effects in sequential decisions: Evidence from speed dating." *Review of Economics and Statistics* 96 (3): 444–457.

Bhavnani, Rikhil R. 2017. "Do the effects of temporary ethnic group quotas persist? Evidence from India." *American Economic Journal: Applied Economics* 9 (3): 105–23.

Cohen, L Jonathan. 1981. "Can human irrationality be experimentally demonstrated?" *Behavioral and Brain Sciences* 4 (3): 317–331.

DellaVigna, Stefano. 2009. "Psychology and economics: Evidence from the field." *Journal of Economic literature* 47 (2): 315–72.

Devine, Patricia G, Patrick S Forscher, Anthony J Austin, and William TL Cox. 2012. "Long-term reduction in implicit race bias: A prejudice habit-breaking intervention." *Journal of Experimental Social Psychology* 48 (6): 1267–1278.

Hartzmark, Samuel M, and Kelly Shue. 2018. "A tough act to follow: Contrast effects in financial markets." *The Journal of Finance* 73 (4): 1567–1613.

Heckman, James J. 1998. "Detecting discrimination." *Journal of Economic Perspectives* 12 (2): 101–116.

Kessler, Judd B, Corinne Low, and Colin D Sullivan. 2019. "Incentivized resume rating: Eliciting employer preferences without deception." *American Economic Review* 109 (11): 3713–44.

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

Neumark, David, Ian Burn, and Patrick Button. 2016. "Experimental age discrimination evidence and the Heckman critique." *American Economic Review* 106 (5): 303–08.

Nickerson, Raymond S. 1998. "Confirmation bias: A ubiquitous phenomenon in many guises." *Review of General Psychology* 2 (2): 175–220.

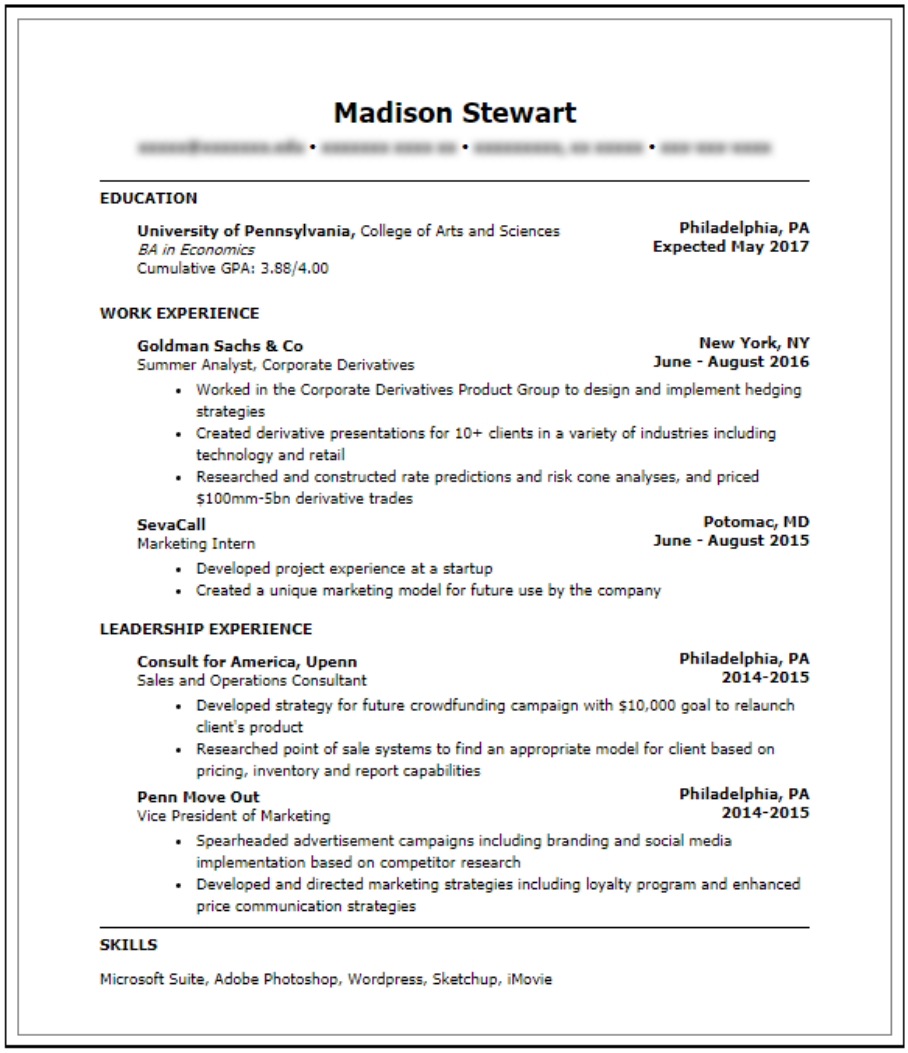
O'Donoghue, Ted, and Matthew Rabin. 2015. "Present bias: Lessons learned and to be learned." *American Economic Review* 105 (5): 273–79.

Radbruch, Jonas, and Amelie Schiprowski. 2020. *Interview Sequences and the Formation of Subjective Assessments.* Technical report.

Riach, Peter A, and Judith Rich. 2002. "Field experiments of discrimination in the market place." *The Economic Journal* 112 (483): F480–F518.

A Online Appendix

Figure A.1: A Sample Resume Generated in the Survey Tool



Wharton
University of Pennsylvania

Madison Stewart

[redacted]

EDUCATION

University of Pennsylvania, College of Arts and Sciences — **Philadelphia, PA**
BA in Economics — **Expected May 2017**
Cumulative GPA: 3.88/4.00

WORK EXPERIENCE

Goldman Sachs & Co — **New York, NY**
Summer Analyst, Corporate Derivatives — **June - August 2016**

- Worked in the Corporate Derivatives Product Group to design and implement hedging strategies
- Created derivative presentations for 10+ clients in a variety of industries including technology and retail
- Researched and constructed rate predictions and risk cone analyses, and priced $100mm-5bn derivative trades

SevaCall — **Potomac, MD**
Marketing Intern — **June - August 2015**

- Developed project experience at a startup
- Created a unique marketing model for future use by the company

LEADERSHIP EXPERIENCE

Consult for America, Upenn — **Philadelphia, PA**
Sales and Operations Consultant — **2014-2015**

- Developed strategy for future crowdfunding campaign with $10,000 goal to relaunch client's product
- Researched point of sale systems to find an appropriate model for client based on pricing, inventory and report capabilities

Penn Move Out — **Philadelphia, PA**
Vice President of Marketing — **2014-2015**

- Spearheaded advertisement campaigns including branding and social media implementation based on competitor research
- Developed and directed marketing strategies including loyalty program and enhanced price communication strategies

SKILLS

Microsoft Suite, Adobe Photoshop, Wordpress, Sketchup, iMovie

Table A.1: Randomization of Resume Components

Resume Component	Description	*Analysis Variable*
Personal Information		
First & last name	Drawn from list of 50 possible names given selected race and gender Race drawn randomly from U.S. distribution (65.7% White, 16.8% Hispanic, 12.6% Black, 4.9% Asian) Gender drawn randomly (50% male, 50% female)	*Female, White* (32.85%) *Male, Non-White* (17.15%) *Female, Non-White* (17.15%) *Not a White Male* (67.15%)
Education Information		
GPA	Drawn $Unif[2.90, 4.00]$ to second decimal place	*GPA*
Major	Drawn from a list of majors at Penn	*Major*
Degree type	BA, BS fixed to randomly drawn major	*Wharton* (40%)
School within university	Fixed to randomly drawn major	*School of Engineering and Applied Science* (70%)
Graduation date	Fixed to upcoming spring (i.e., May 2017)	
Work Experience		
First job	Drawn from curated list of top internships and regular internships	*Top Internship* (20/40)
Title and employer	Fixed to randomly drawn job	
Location	Fixed to randomly drawn job	
Description	Bullet points fixed to randomly drawn job	
Dates	Summer after candidate's junior year (i.e., 2016)	
Second job	Left blank or drawn from curated list of regular internships and work-for-money jobs	*Second Internship* (13/40) *Work for Money* (13/40)
Title and employer	Fixed to randomly drawn job	
Location	Fixed to randomly drawn job	
Description	Bullet points fixed to randomly drawn job	
Dates	Summer after candidate's sophomore year (i.e., 2015)	
Leadership Experience		
First & second leadership	Drawn from curated list	
Title and activity	Fixed to randomly drawn leadership	
Location	Fixed to Philadelphia, PA	
Description	Bullet points fixed to randomly drawn leadership	
Dates	Start and end years randomized within college career, with more recent experience coming first	
Skills		
Skills list	Drawn from curated list, with two skills drawn from {Ruby, Python, PHP, Perl} and two skills drawn from {SAS, R, Stata, Matlab} shuffled and added to skills list with probability 25%.	*Technical Skills* (25%)

Notes: Resume components are listed in the order that they appear on hypothetical resumes. Italicized variables in the right column are variables that were randomized to test how employers responded to these characteristics. Degree, first job, second job, and skills were drawn from different lists for Humanities & Social Sciences resumes and STEM resumes (except for work-for-money jobs). Name, GPA, work-for-money jobs, and leadership experience were drawn from the same lists for both resume types. Weights of characteristics are shown as fractions when they are fixed across subjects (e.g., each subject saw exactly 20/40 resumes with a *Top Internship*) and percentages when they represent a draw from a probability distribution (e.g., each resume a subject saw had a 32.85% chance of being assigned a white female name). Additional details can be found in Kessler, Low, and Sullivan (2019).

Table A.2: Contrast Effect Regressions (Clustered S.E.)

	(1)	(2)	(3)	(4)	(5)
Dependent Variable: Hiring Interest					
After White Man	-0.174**	-0.198**	-0.182**	-0.189**	-0.188**
	(0.078)	(0.076)	(0.078)	(0.079)	(0.079)
Observations	2,808	2,808	2,808	2,808	2,808
R-squared	0.427	0.446	0.481	0.489	0.489
Subject fixed effects	Yes	Yes	Yes	Yes	Yes
Major fixed effects	No	Yes	Yes	Yes	Yes
Leadership fixed effects	No	No	Yes	Yes	Yes
Order fixed effects	No	No	No	Yes	Yes
Previous Resume Quality	No	No	No	No	Yes

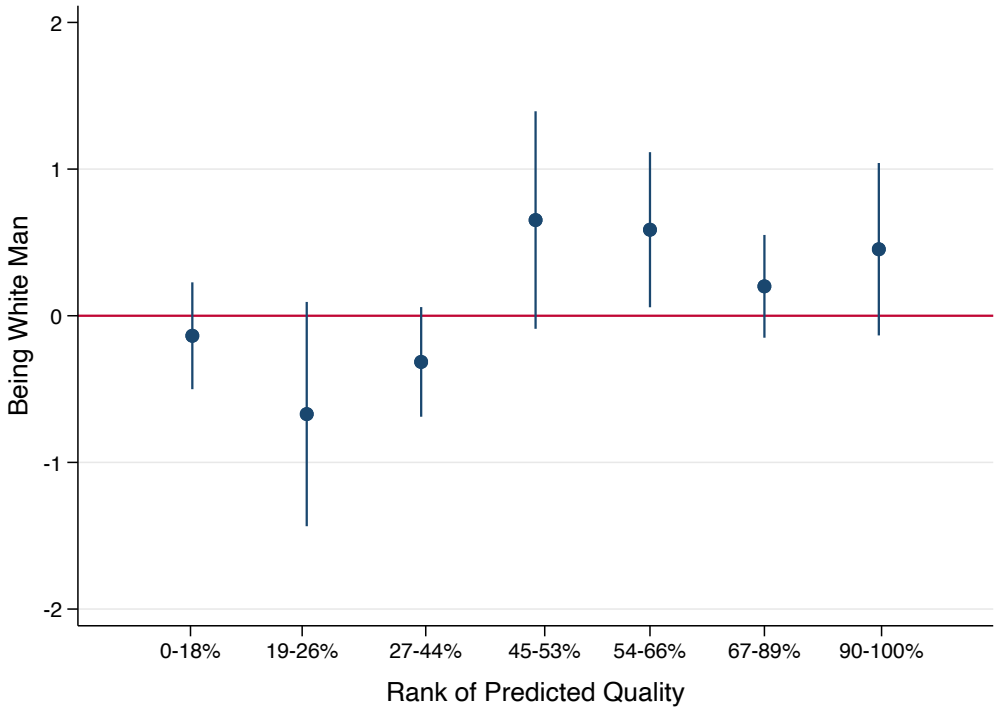
Notes: The table uses the same sample and specifications as in Table 1. Standard errors in parentheses are clustered at the subject level. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

Table A.3: No Evidence on Quality Contrast Effects

Dependent Variable: Hiring Interest	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Prior GPA	0.008						0.009
	(0.120)						(0.121)
After Top Internship		0.050					0.048
		(0.080)					(0.081)
After Second Internship			-0.040				-0.050
			(0.084)				(0.096)
After Work for Money				-0.001			-0.026
				(0.083)			(0.095)
After Technical Skills					0.057		0.055
					(0.092)		(0.092)
After Low-Quality						-0.015	
						(0.083)	
Observations	2,808	2,808	2,808	2,808	2,808	2,808	2,808
R-squared	0.488	0.488	0.488	0.488	0.488	0.488	0.488

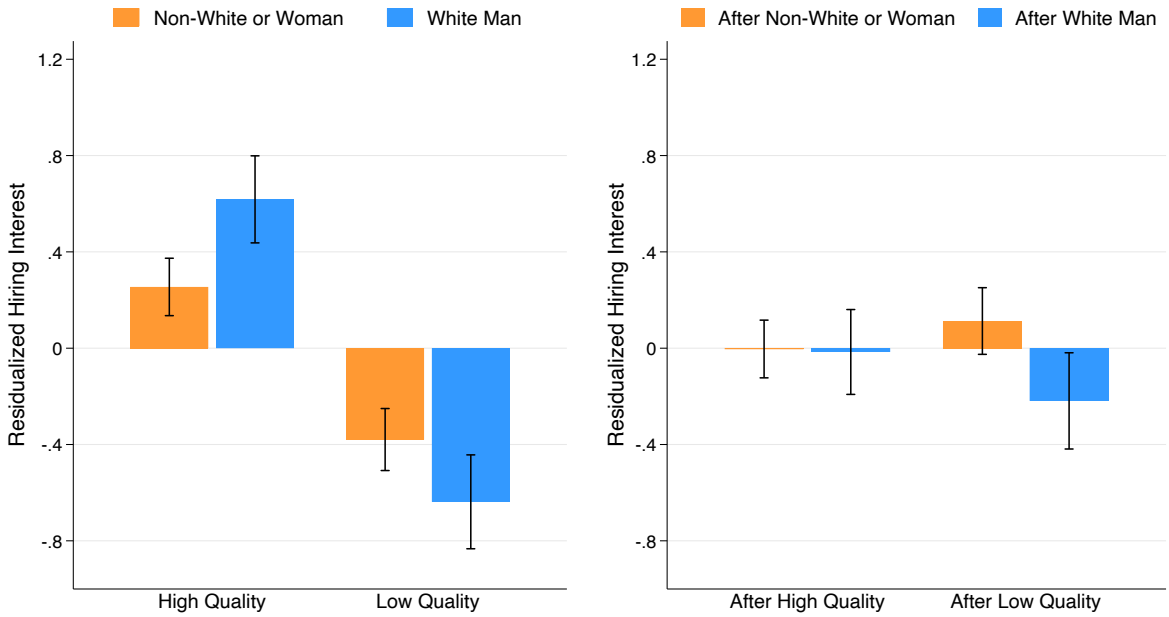
Notes: All regressions control for all the resume characteristics and fixed effects specified in column (4) of Table 1—without the "After White Male" variable. See Section 4.1 for the definition of "low quality". When separately looking at the sample of white male resumes and other resumes, We find no significant quality contrast effects either. Robust standard errors are in parentheses.

Figure A.2: Preference for White Men in Ratings by Predicted Resume Quality



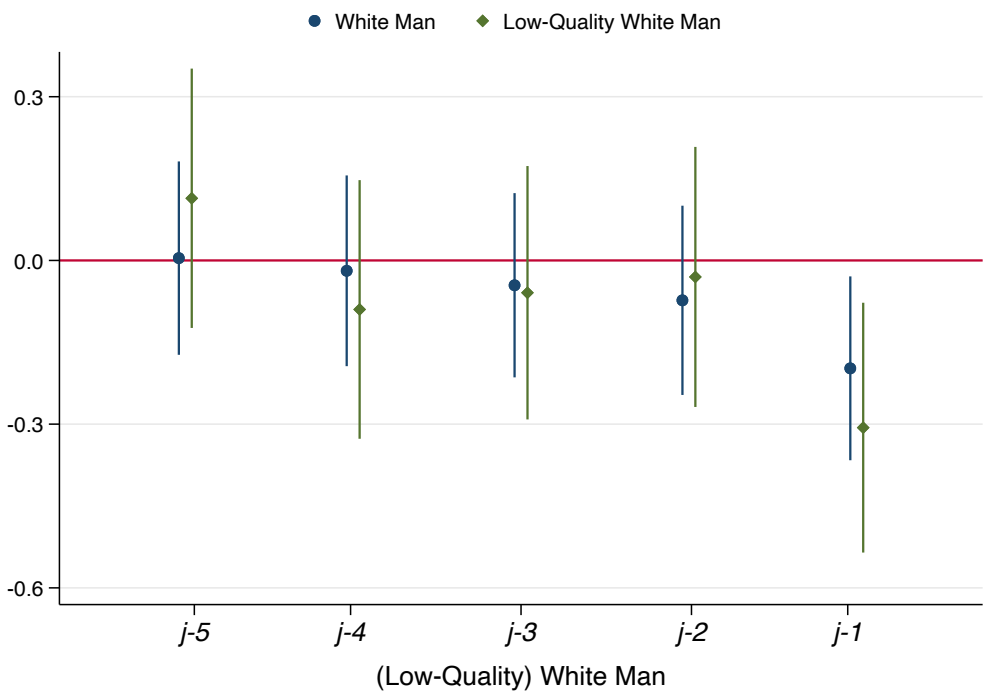
Notes: We use Lasso to predict resume ratings using GPA dummies (i.e., GPA rounded to the nearest 0.1), dummies for work experiences (i.e., top internship, second internship, work-for-money job), and technical skills. The distribution of predicted ratings is discrete and can be naturally grouped into 7 clusters, shown on the x-axis. For each cluster, we estimate the preference for white men in ratings, controlling for GPA, work experience dummies, technical skills, major fixed effects, resume order fixed effects, and subject fixed effects. (We do not include leadership experience fixed effects because there is not enough variations within clusters.) Error bars indicate 95% confidence intervals.

Figure A.3: Rating by Current/Previous Resume being a White Man and by Resume Quality



Notes: The left panel shows how employers' hiring interest varies with resumes demographics (being a white man or not being a white man) and the quality (low or high). The right panel shows how hiring interest varies with the previous resume's demographics and quality type. We derive the residualized hiring interest by controlling for the fixed effects for subject, major, leadership experiences, and resume order. Error bars indicate 95% confidence intervals.

Figure A.4: Effects of Previous (Low-Quality) White Men



Notes: Figure shows the effects on ratings of a (low-quality) white man being placed: immediately before a resume $j-1$, two resumes before $j-2$, three resumes before $j-3$, four resumes before $j-4$, and five resumes before $j-5$. Each point estimate represents one regression. When estimating the effect of a resume belonging to a (low-quality) white man in a given period (e.g., $j-5$), we control for whether later resumes before the current resume (e.g., $j-4$ to $j-1$) are also (low-quality) white men. The pattern we find is very similar if we do not control for whether more-recent resumes are (low-quality) white men. We also control for all resume characteristics and fixed effects in column (4) of Table 1. Error bars indicate 95% confidence intervals.