Lowering the Playing Field: Discrimination through Sequential Spillover Effects*

Judd B. Kessler Corinne Low Xiaoyue Shan

August 2022

Abstract

We document a new source of discrimination that arises through sequential spillover effects. Employers in an incentivized resume rating experiment evaluate a sequence of hypothetical candidates with randomly assigned characteristics. Candidates are rated worse when following white men than when following women or minorities. Exploring the mechanisms, we find that spillover effects are inversely related to explicit bias. When reviewing high-quality resumes or recruiting in STEM industries, employers directly favor white men and display no spillover effects. For low-quality resumes or non-STEM industries, we find no direct bias but strong spillover effects. Our results highlight the power of implicit bias.

*The Wharton School, University of Pennsylvania, 3733 Spruce St, Vance Hall, Philadelphia, PA 19104 (email: judd.kessler@wharton.upenn.edu, corlow@wharton.upenn.edu, xiaoyues@wharton.upenn.edu).
We thank the participants of the seminar on the Economics of Discrimination and Disparities, the ESA 2022 World Meeting, and SITE Experimental Economics for helpful comments and suggestions.

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 male candidate, recruiters give a lower rating to the next candidate that they evaluate. We call this a sequential spillover 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. The subset of employers recruiting students with majors in science, technology, engineering, and math (STEM) fields, however, did display statistically significant bias in favor of white men.

This paper further interrogates the employer ratings from that paper and leverages the fact that—since all 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. We show that resumes placed after white men are rated statistically significantly worse than those that follow women or minorities. The spillover 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). 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 spillover effect does not significantly differ by the demographics or quality 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. Moreover, the effect does not significantly vary with the employer's demographics or the employer's explicitly stated preference for hiring diverse candidates.

To understand this puzzling result, we explore whether a certain subset of the resumes of white men generate this spillover 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 other words, recruiters display direct bias in favor of white men when evaluating high-quality candidates. Following these high quality resumes, we see no spillover effect: candidates that follow high-quality white men are not rated worse than candidates that follow high-quality women or minorities. When the resume is low quality, however, there is no direct preference for white men, but a large spillover effect impacting the following candidate. 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).

We find a similar pattern that the spillover effect inversely relates to explicit bias in ratings when we split the data by whether the employer is recruiting in STEM industries. STEM employers—who on average display a direct bias in favor of white men—do not display a spillover effect. On the other hand, employers recruiting humanities, social science, and business majors—who on average do not display a direct bias in favor of white men—do display a strong spillover effect in their ratings.

Taken together, the explanation that best fits our pattern of results is that employers have an implicit bias in favor of white men, which is operative in some circumstances (e.g., in STEM fields) but constrained in others (e.g., when the candidate is clearly unremarkable). But when employers do not act on this implicit bias in their rating of the candidate—when it is constrained—they instead "lower the playing field" such that the next candidate is rated more harshly, thus making the white man look better by contrast. In other words, the sequential spillover effect serves as an alternative channel to display implicit bias toward white men.

This paper provides new evidence on how implicit bias may operate. By showing that implicit bias generates direct favoritism only in certain circumstances—which can be explained by employers needing some "justification" for favoring a white man—this study connects to literature on the role of self-signaling (Bem 1972; Bénabou and Tirole 2011; Grossman and Van der Weele 2017) and new work showing evidence of more bias when there is more ambiguity about quality (Chan 2022). More broadly, related work shows that acting on "undesirable" preferences is more common when there is some perceived justification or "excuse" (Exley 2016; Exley and Kessler 2019). We show that when direct favoritism is constrained by the absence of such a justification, implicit bias may spill over and manifest as a penalty on the following candidate.

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). Our finding that the display of bias can be indirect and context-specific contributes to a growing literature on implicit discrimination and stereotypes (Bertrand, Chugh, and Mullainathan 2005; Tetlock and Mitchell 2009; Carlana 2019; Hangartner, Kopp, and Siegenthaler 2021; Barron et al. 2022; Cunningham and De Quidt 2022). By uncovering a dynamic feature of discrimination, this paper also relates to a few recent studies exploring these dynamics; see Bohren, Imas, and Rosenberg (2019), Miller and Schmutte (2021), and Benson and Louis-Pierre (2022). In exploring sequential spillover effects in hiring, our paper also closely relates to work, including Abel (2017), Phillips (2019), and Radbruch and Schiprowski (2020), on how the composition of an applicant pool (in terms of quality or immigration status) can affect job seekers' application outcomes. Relative to this work, we document a new source of bias that arises through spillover effects even when direct favoritism is constrained, raising new equity issues.[1]

[1]Our paper also relates to the literature on contrast effects, which document sequential spillover effects in other settings (Pepitone and DiNubile 1976; Simonson and Tversky 1992). For example, Bhargava and Fisman (2014) find that in speed dating, prior partner attractiveness lowers male evaluators' desire 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. Unlike with contrast effects, however, the sequential spillover in our setting does not result from the contrast between resumes. Instead, it arises simply because the prior resume was a white man.

2 Experimental Design, Data, and Specification

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).[2] 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 (excluding demographics).

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

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

[2] 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.

Responses were on a 10-point Likert scale, where 1 was “Not interested” and 10 was “Very interested”. By separately capturing candidate “gettability” in the likelihood of acceptance question, we ensure that the hiring interest question is not “top-coded,” and thus can be interpreted as a measure of perceived candidate quality.

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.[3]

To identify spillover 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 spillover 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%).[4]

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

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

2.3 Empirical specification

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

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

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). $N^{wm}_{i,j-1}$ 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 most recent work experience is a 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.[5] 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 spillover effect β: (1) orthogonal relationships between resume components and (2) the randomized order of resumes. Because all resume components (including demographic and quality indicators) were independently and randomly drawn, the measured effect of demographics could not be driven by the possible correlations between demographics and other char-

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

acteristics. 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 Results

Table 1 shows the spillover effect using different specifications of equation (1). The dependent variable is the rating of hiring interest on a scale of 1 to 10. Our key dependent variable of interest is whether the preceding resume was a white male, $N^{wm}_{i,j-1}$, (*After White Man*). In column (1), we estimate the coefficient of *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 spillover 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 a white man—rather than after a female or minority candidate—lowers the current rating by 0.19 Likert points. This represents 4% of the average rating of resumes not following white men, or 7% of a standard deviation in the ratings of all resumes. Comparing this estimate to the estimate on *GPA* suggests that being after a white man is equivalent to having a GPA that is approximately 0.1 points lower.

As highlighted in Kessler, Low, and Sullivan (2019), the table also shows that academic ability and work experience naturally impact employers' ratings. For example, employers evaluate much more favorably candidates with higher GPAs, higher quality work experiences, and more work experiences. However, these quality indicators do not

Table 1: Spillover 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. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

create spillover effects, as shown in column (5). In Appendix Table A.3, we test for quality spillover effects when separately using different quality indicators. We find no evidence that previous resume quality affects the rating of the current resume.

4 Explaining the Sequential Spillover Effect

In this part of the paper, we explore potential mechanisms driving the sequential spillover effect generated by white male candidates. We first examine whether the effect varies with the demographics and quality of the current resume and whether the effect varies with the employer's demographics and explicitly stated preference for hiring diverse candidates (Section 4.1). Next, we show evidence that the spillover effect is negatively related to the direct display of bias shown toward white men: the effect arises from low-quality white male candidates and is concentrated in humanities, social sciences, and business industries, where a direct preference for white men is not detected in resume ratings (Section 4.2). We then explore the dynamics of the spillover effect over multiple resumes (Section 4.3). Finally, we provide an explanation for how the spillover effect may operate (Section 4.4).

4.1 Spillover is not driven by characteristics of current resume or employer

The role of current resume's demographics and quality Table 2, Panel A, column (1) shows the size of the spillover effect when the current resume (i.e., resume j) is a white man and when it is not. We find that the effect of being after a white man does not statistically significantly differ by whether the current resume is a white man.[6] These results suggest that the documented spillover effect does not appear driven by conscious favoritism of white men, which would imply white men would be exempted from the spillover effect. Instead, evaluating a 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 white man a boost by lowering the ratings of anyone who follows him.

Columns (2) and (3) test whether the effect varies with the current resume's quality. Rather than looking at specific quality indicators such as GPA and work experiences,

[6]More generally, the spillover effect does not statistically significantly vary with the gender or race of the current resume.

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, using Lasso to identify which characteristics should be used to create both a standardized quality measure and a binary split into high- and low-quality resumes.[7] Using either the continuous or binary quality measure, we find that the effect on *After White Man* does not vary by the quality of the current resume, again indicating it is a general "lowering of the field".

The role of employer demographics and stated diversity preference We also test whether the spillover effect is driven by a certain subset of employers. In the post-survey, employers 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 diversity" and "seeking to increase racial diversity" as factors in their rating of candidates (each on a scale of 1 to 10). Over 90% of employers indicated that they considered these factors favorably in their hiring. We use responses to these survey questions to create a continuous measure of employers' explicitly stated preference for diversity and to construct a binary classification classifying employers as placing "high importance" on diversity if they provided above-median responses on the diversity questions.

Table 2, Panel B, column (1) shows that the spillover effect does not differ with whether the employer is a white man.[8] Point estimates suggest that white men display a directionally smaller spillover effect, but the difference is not statistically significant.

[7] 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 and use the predicted ratings as the quality measure. The ranking of predicted resume quality is very similar when we use alternative prediction methods, like OLS or alternative λ-choosing methods (e.g., cross-validation). In some of our analysis we use a standardized measure of quality. We also create a binary quality measure, dividing resumes into "low quality" (the bottom 44% of resumes) and "high quality" (the top 56% of resumes). See further discussion about this cutoff in footnote 9.

[8] We find no significant differences in the effect by employer's race or gender.

Table 2: Heterogeneity in the Effect of Being After a White Man

Dependent Variable: Rating of Hiring Interest	(1)	(2)	(3)
Panel A: By Current Resume Demographics and Quality			
After White Man	-0.171*	-0.189**	-0.250**
	(0.101)	(0.085)	(0.123)
After White Male × White Male Resume	-0.056		
	(0.184)		
After White Male × Std. Resume Quality		0.011	
		(0.082)	
After White Male × High-Quality Resume			0.101
			(0.170)
Panel B: By Employer Demographics and Diversity Preference			
After White Man	-0.217**	-0.189**	-0.172
	(0.102)	(0.085)	(0.112)
After White Male × White Male Employer	0.087		
	(0.185)		
After White Male × Std. Importance of Diversity		0.032	
		(0.084)	
After White Male × High Importance of Diversity			-0.039
			(0.171)

Notes: All estimations include the control variables specified in column (4) of Table 1. The table does not report the effects of a white male resume, or any features of the employer, because these are already absorbed by the fixed effects included in the regression specification. *Std.* indicates that we are using a standardized version of the variable (i.e., mean is set to 0 and standard deviation is set to 1). White not reported, the resume quality variables are also included in the regressions in columns (2)–(3) of Panel A since the predicted quality is not perfectly collinear with resume quality characteristics. Robust standard errors are in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Columns (4) and (5) test for potential interaction effects between being placed after a white man and the employer's explicitly stated diversity preference. We find that the spillover effect also does not significantly differ between employers who place more or less importance on diversity in hiring. If anything, it is directionally larger for employers actively seeking to increase racial and gender diversity in their hiring.

These findings indicate that the sequential spillover effect is not heterogenous by either demographic traits that may be linked to explicit preferences nor by employer's stated preferences for diversity. This further suggests the bias may be unintentional, or implicit.

4.2 Spillover is larger where there is less direct bias

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 white candidates received a higher return to improving resume quality in a resume audit. Here, we further explore whether resume quality interacts with the preference for white men and the spillover effect generated by white men.

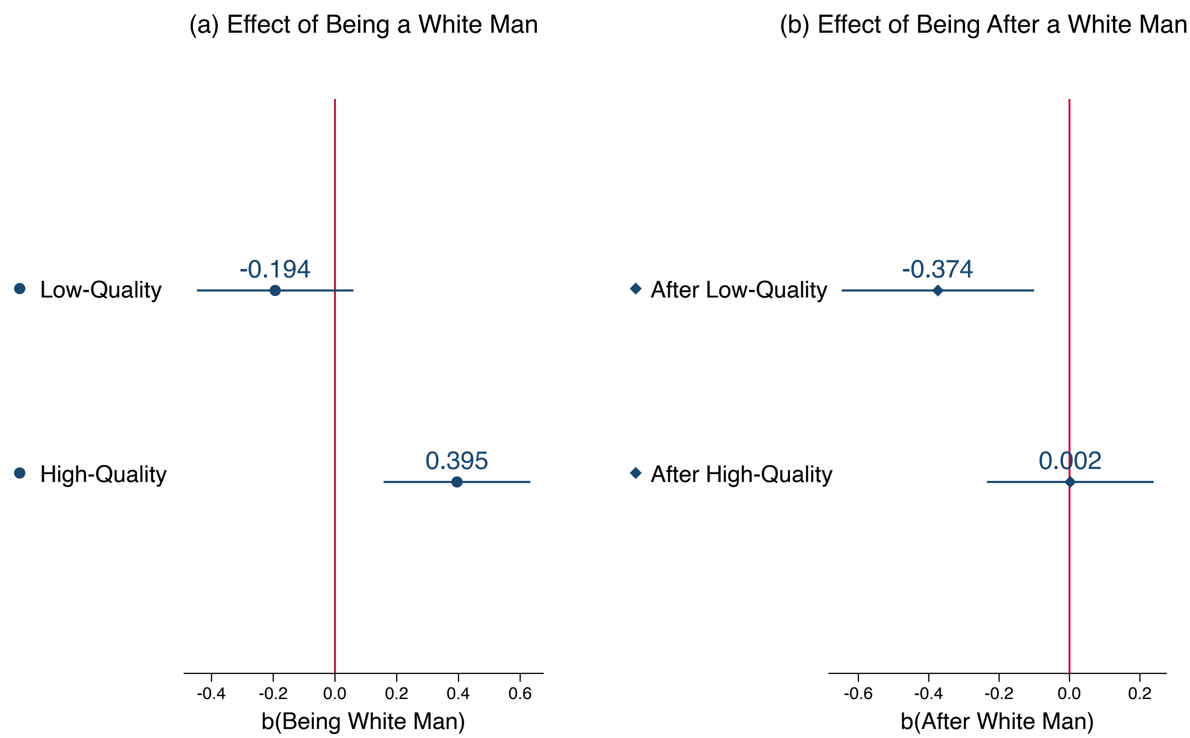
We first employ the Lasso predicted resume quality described in footnote 7 and estimate the preference for white men in ratings, both for low-quality resumes and for high-quality resumes.[9] As Figure 1(a) shows, white men receive higher ratings than women and minorities for high-quality resumes. For statistically identical high-quality 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 show that employers are *biased* in favor of white men when resume quality is high, since the demographic assignment is random and orthogonal to all other resume traits.

On the other hand, there is a directional (i.e., 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 resume ratings are statistically significantly different across high-quality and low-quality resumes, highlighting that resume

[9] The distribution of predicted ratings display a few big clusters and a few small ones—which we combine into 7 quality groups. For each quality group, we estimate the gap in ratings between white male candidates and other candidates. 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%. This pattern further corroborates our choice of 44% as the natural cutoff to define resumes as either "low quality" or "high quality."

quality is an important determinant of whether employers display bias.

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 following a white man's resume varies with whether that resume was high or low quality. We find that the spillover 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) lower if the prior candidate was a white man rather than a woman or minority, equivalent to 0.17 GPA points. The magnitude of this spillover effect is almost the same as the partiality displayed toward high-quality white men in panel (a). Meanwhile, there is no spillover effect when resumes follow high-quality white men (the estimate of -0.374 of following a low-quality white man is statistically significantly different than the estimate of 0.002 of following a high-quality white man).[10] To confirm these results are not driven by spurious serial correlation, we run a placebo test of the effect of preceding a low-quality white man, and find no effect.[11]

Taken together, Figure 1 suggests that the spillover effect relates to whether employers display favoritism toward white men when rating the prior resume. Employers only lower ratings of the current resume when they did not show bias in favor of the white man they just evaluated.[12]

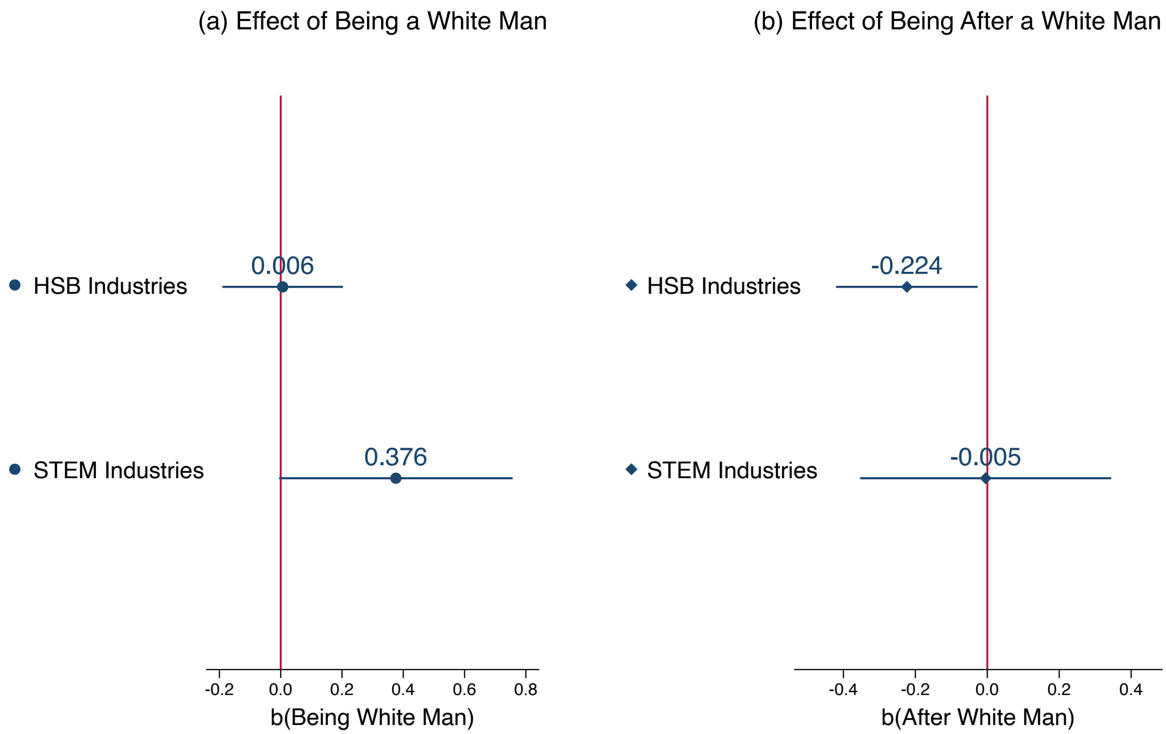
The role of industry We further test whether spillover effects vary by the employer's industry. As documented in Kessler, Low, and Sullivan (2019), employers looking for candidates with science, technology, engineering, mathematics (STEM) backgrounds display a direct preference for white men, while employers hiring candidates with humanities, social science, and business backgrounds (HSB, for short) do not show such a direct bias. Guided by the above results, we explore whether the spillover effect differs across employers recruiting for STEM and HSB and, in particular, whether the spillover effect is concentrated among the HSB employers who do not display a direct bias toward white men.

[10] Using an alternative specification where we replace *After White Man* in column (4) of Table 1 with two variables, *After Low-Quality White Man* and *After High-Quality White Man*, we find similar results: the estimated coefficient of the former is -0.312 (p-value=0.008) and the latter is -0.09 (p-value=0.392).

[11] Appendix Table A.5 shows that the estimated coefficient of *Before (Low-Quality) White Man* is close to zero and has large standard errors, regardless of whether we also control for *After (Low-Quality) White Man*. The null effect of preceding a white man is expected, because employers can only move forward but not backward in their evaluation of candidates, but it helps to confirm that our estimated spillover effects are not caused by spurious temporal correlations in evaluations or mean reversion.

[12] Since we have documented that the spillover effect is almost entirely driven by low-quality white men, we rerun the heterogeneity results shown in Table 2 but replace *After White Man* with *After Low-Quality White Man*. Results are shown in Appendix Table A.4. The sequential spillover effect from following a low-quality white man does not significantly vary with the current resume's demographics or quality or the employer's demographics or explicitly stated preference for diversity.

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



Notes: The figure shows the estimated effects of being a white man and being after a white man on resume ratings by the employer's industry type. "STEM" refers to science, technology, engineering, mathematics, and "HSB" refers to humanities, social science, and business. Each point estimate is from one regression. *b(Being White Man)* is estimated using the same specification as in Figure 1(a), restricted to the industry being analyzed, and *b(After White Man)* is from the same specifications as in Figure 1(b), restricted to the industry being analyzed rather than quality of the preceding resume. Error bars indicate 95% confidence intervals.

Figure 2 shows the estimated preference for white men and the sequential spillover effect in the STEM and HSB industries. STEM employers on average rate white men higher than minority or female candidates by 0.38 Likert points. These employers display no spillover effect in their evaluation of subsequent resumes. On the other hand, employers looking to hire in HSB do not show a direct bias in favor of white men; instead, they display a strong and statistically significant spillover effect, rating resumes following white man 0.22 Likert points lower.[13]

[13]The point estimates have larger standard errors because our sample includes fewer STEM employers (30%) and more HSB employers (70%).

This pattern is consistent with the pattern by resume quality. Where there is direct bias shown to white men, there is no spillover effect on a candidate following one; where there is no direct bias, on the other hand, we observe a negative spillover effect.

4.3 Dynamics of the spillover effect

All of our previous analyses have focused on the effect of *immediately* following a white man. To better understand the sequential spillover 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 $(\mathrm{WM}_{j-2}, \mathrm{WM}_{j-1})$ 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 $(\mathrm{Other}_{j-2}, \mathrm{WM}_{j-1})$, 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 $(\mathrm{WM}_{j-2}, \mathrm{Other}_{j-1})$ 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 spillover 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.[14]

4.4 How the sequential spillover effect may operate

The results above document a spillover effect that manifests after an employer has evaluated the resume of a white man. Any resume that follows a white man is rated significantly more harshly than if the resume had followed a woman or minority. The 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 spillover effect "lowers the playing field," imposing an equal penalty on anyone that follows the white man.

This bias is driven entirely by low-quality resumes. We find that there is no sequential spillover effect when following a high-quality white male resume. In addition, the spillover effect has limited persistence; it influences the evaluation of the next resume but does not have a longer lasting impact. However, there is some evidence that it can compound, as the spillover effect is directionally bigger after the employer has just evaluated two low-quality white men back-to-back.

When considering the causes of the spillover effect, we find it important to consider

[14] We limit our analysis of the dynamics of the spillover 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.3, 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

	(1)	(2)	(3)	(4)
	Dependent Variable: Hiring Interest			
WM_{j-2}	-0.073			
	(0.088)			
WM_{j-1}	-0.198**			
	(0.086)			
$\text{Other}_{j-2}, \text{WM}_{j-1}$		-0.130		
		(0.103)		
$\text{WM}_{j-2}, \text{Other}_{j-1}$		-0.005		
		(0.106)		
$\text{WM}_{j-2}, \text{WM}_{j-1}$		-0.352**		
		(0.144)		
LQWM_{j-2}			-0.030	
			(0.122)	
LQWM_{j-1}			-0.307***	
			(0.117)	
$\text{Other}_{j-2}, \text{LQWM}_{j-1}$				-0.292**
				(0.126)
$\text{LQWM}_{j-2}, \text{Other}_{j-1}$				-0.016
				(0.130)
$\text{LQWM}_{j-2}, \text{LQWM}_{j-1}$				-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. WM_{j-2} means that the resume before the prior resume is a white man. LQWM_{j-2} means that the resume before the prior resume is a low-quality white man. WM_{j-1} means that the prior resume is a white man. LQWM_{j-1} means that the prior resume is a low-quality white man. Other_{j-2} and Other_{j-1} are indicators that the relevant resume (i.e., $j-2$ or $j-1$) is not a (low-quality) white man. Robust standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

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 spillover effect only arises after the white men that the employer does not directly favor. We see the same pattern when splitting our data by industry; employers recruiting in STEM industries display a direct bias in favor of white men and display no spillover effect, while those recruiting in humanities, social sciences, and business industries do not display a direct bias in favor of white men but do display strong spillover effects.

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 or when they recruit candidates in STEM fields. When employers rate low-quality resumes or look for candidates with non-STEM backgrounds, however, this implicit bias is somehow constrained (perhaps because the candidate is obviously unremarkable or the industry norm against discrimination is more salient, and thus showing the bias explicitly would create a negative self-signal for the employer). This suggests that employers may need an "excuse" for their implicit bias to manifest as direct discrimination—a high-quality resume or the greater prevalence of white men in STEM fields may provide an adequate unconscious justification for rating white men more highly.

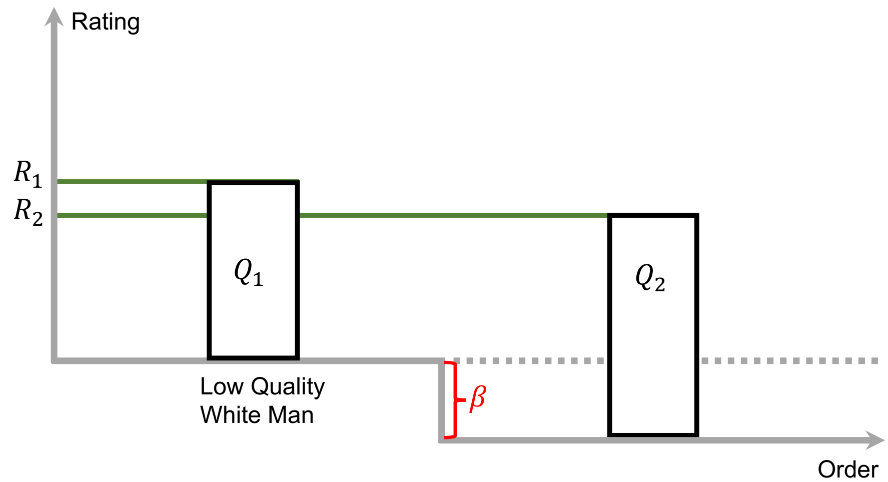
However, even when direct bias is constrained, employers may still indirectly favor white men through a spillover effect imposed on subsequent candidates. Employers "lower the playing field" by rating the next candidate more harshly, thus making the (low-quality) white man look somewhat better by comparison.

As noted in the prior paragraph, our evidence suggests that the direct favoritism and the indirect favoritism through a spillover 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.[15] Third, if they wanted to explicitly favor a specific low-quality white man through the sequential spillover 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 spillover effect implies that it is driven by a psychological urge

[15]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.

rather than a rational response.

Figure 3: Impact of the Spillover Effect—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 spillover 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 spillover effect could still have big impacts. In a setting with two candidates, as shown in Figure 3, a spillover 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 spillover effect.

5 Conclusion

This paper documents a new 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 find that resumes following white male candidates are rated significantly lower than resumes following female or minority candidates.

Digging into the mechanisms for this effect, we observe that employers display a direct preference for resumes randomly assigned the names of white men only when the candidate is of high quality or when recruiting in STEM fields. When the white male candidate is low quality or from a non-STEM field, we observe no such preference. However, while employers fail to show partiality toward low-quality white men and to men in non-STEM fields, after evaluating one, they lower their rating of the subsequent candidate.

Our findings suggest the power of implicit bias. Even when no direct favoritism is shown towards white men, the bias manifests in unexpected ways, benefiting the white men indirectly through the spillover effect. This has important policy implications for designing systems for mitigating implicit bias.

Our data—with 40 ratings per employer of randomly generated resumes—provides an ideal environment to cleanly identify and explore a sequential spillover effect. Future research should explore the presence of such spillover effects in other settings. 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

Abel, Martin. 2017. "Labor market discrimination and sorting: evidence from South Africa." *World Bank Policy Research Working Paper No. 8180.*

Barron, Kai, Ruth Ditlmann, Stefan Gehrig, and Sebastian Schweighofer-Kodritsch. 2022. "Explicit and implicit belief-based gender discrimination: A hiring experiment." *CESifo Working Paper.*

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.

Bem, Daryl J. 1972. "Self-perception theory." In *Advances in experimental social psychology,* 6:1–62. Elsevier.

Bénabou, Roland, and Jean Tirole. 2011. "Identity, morals, and taboos: Beliefs as assets." *The Quarterly Journal of Economics* 126 (2): 805–855.

Benson, Alan, and Lepage Louis-Pierre. 2022. "Learning to Discriminate on the Job." *Working Paper.*

Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. 2005. "Implicit discrimination." *American Economic Review* 95 (2): 94–98.

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.

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

Carlana, Michela. 2019. “Implicit stereotypes: Evidence from teachers’ gender bias.” *The Quarterly Journal of Economics* 134 (3): 1163–1224.

Chan, Alex. 2022. “Discrimination and Quality Signals: A Field Experiment with Healthcare Shoppers.” *Unpublished manuscript.*

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

Cunningham, Tom, and Jonathan De Quidt. 2022. “Implicit Preferences.” *Working Paper.*

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.

Exley, Christine L. 2016. "Excusing selfishness in charitable giving: The role of risk." *The Review of Economic Studies* 83 (2): 587–628.

Exley, Christine L, and Judd B Kessler. 2019. "Motivated errors." *NBER Working Paper No. 26595.*

Grossman, Zachary, and Joël J Van der Weele. 2017. "Self-image and willful ignorance in social decisions." *Journal of the European Economic Association* 15 (1): 173–217.

Hangartner, Dominik, Daniel Kopp, and Michael Siegenthaler. 2021. "Monitoring hiring discrimination through online recruitment platforms." *Nature* 589 (7843): 572–576.

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.

Miller, Conrad, and Ian M Schmutte. 2021. "The Dynamics of Referral Hiring and Racial Inequality: Evidence from Brazil." *NBER Working Paper No. 29246.*

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.

Pepitone, Albert, and Mark DiNubile. 1976. "Contrast effects in judgments of crime severity and the punishment of criminal violators." *Journal of Personality and Social Psychology* 33 (4): 448.

Phillips, David C. 2019. "Do comparisons of fictional applicants measure discrimination when search externalities are present? Evidence from existing experiments." *The Economic Journal* 129 (621): 2240–2264.

Radbruch, Jonas, and Amelie Schiprowski. 2020. "Interview Sequences and the Formation of Subjective Assessments." *ECONtribute Discussion Paper.*

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

Simonson, Itamar, and Amos Tversky. 1992. "Choice in context: Tradeoff contrast and extremeness aversion." *Journal of Marketing Research* 29 (3): 281–295.

Tetlock, Philip E, and Gregory Mitchell. 2009. "Implicit bias and accountability systems: What must organizations do to prevent discrimination?" *Research in Organizational Behavior* 29:3–38.

A Online Appendix

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

Wharton
University *of* Pennsylvania

Madison Stewart

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	*Female, White* (32.85%)
		Male, Non-White (17.15%)
	Race drawn randomly from U.S. distribution (65.7% White, 16.8% Hispanic, 12.6% Black, 4.9% Asian)	*Female, Non-White* (17.15%)
		Not a White Male (67.15%)
	Gender drawn randomly (50% male, 50% female)	
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: Spillover 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 Spillover Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: Hiring Interest						
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

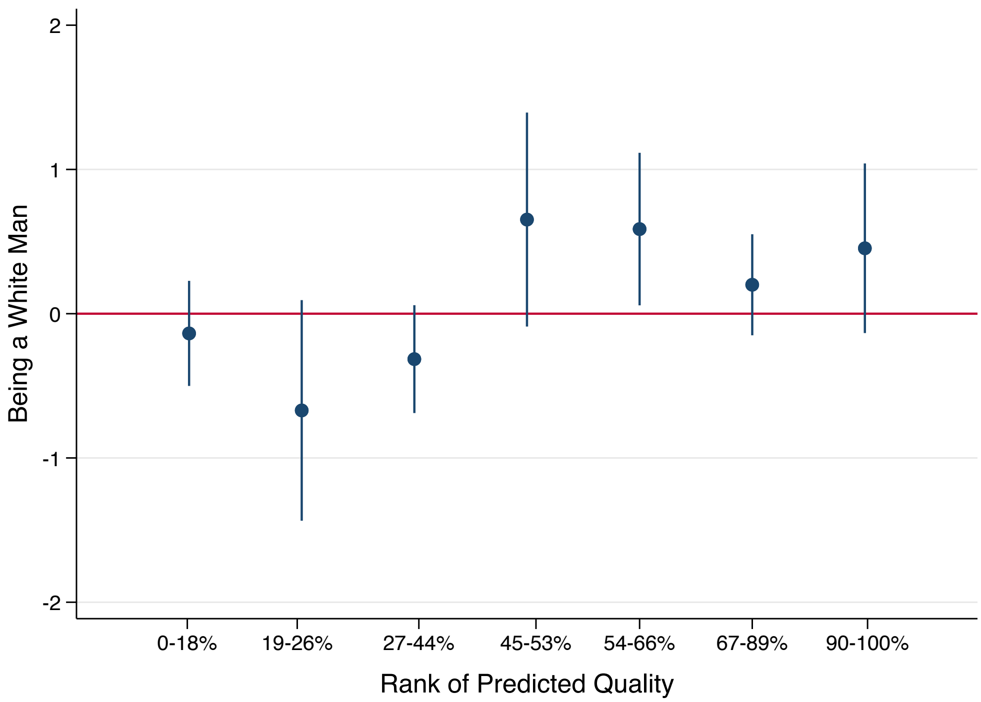
Notes: All regressions control for all the resume characteristics and fixed effects specified in column (4) of Table 1—without the *After White Man* variable. See Section 4.2 for the definition of "low quality". When separately looking at the sample of white male resumes and other resumes, We find no significant quality spillover effects either. Robust standard errors are in parentheses.

Table A.4: Heterogeneity in the Effect of Being After a Low-Quality White Man

Dependent Variable: Rating of Hiring Interest	(1)	(2)	(3)
Panel A: By Current Resume Demographics and Quality			
After Low-Quality White Man	-0.326**	-0.275**	-0.447***
	(0.140)	(0.116)	(0.161)
After Low-Quality White Man × White Male Resume	0.096		
	(0.243)		
After Low-Quality White Man × Std. Resume Quality		0.151	
		(0.120)	
After Low-Quality White Man × High-Quality Resume			0.259
			(0.226)
Panel B: By Employer Demographics and Diversity Preference			
After Low-Quality White Man	-0.354***	-0.291**	-0.330**
	(0.134)	(0.115)	(0.150)
After Low-Quality White Man × White Male Employer	0.181		
	(0.256)		
After Low-Quality White Man × Std. Importance of Diversity		0.072	
		(0.109)	
After Low-Quality White Man × High Importance of Diversity			0.087
			(0.231)

Notes: The table replicates Table 2 by only replacing the *After White Man* variable with *After Low-Quality White Man*. Robust standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.

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



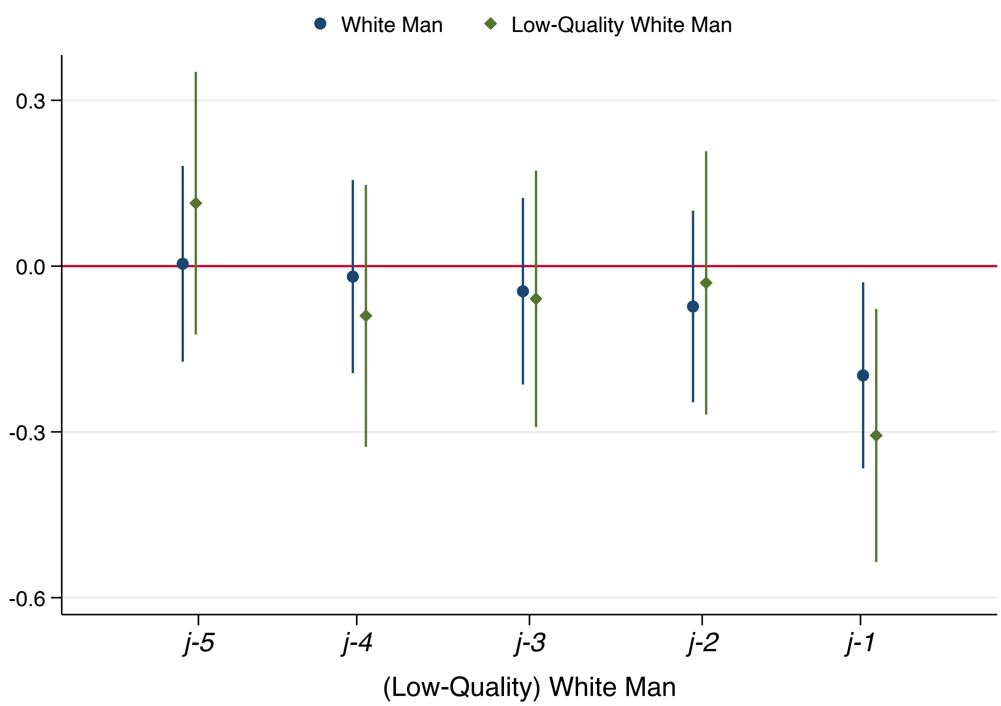
Notes: We use Lasso to predict resume ratings using GPA dummies (i.e., GPA rounded to the nearest 0.1), dummies for work experiences (i.e., top internship, second internship, work-for-money job), and technical skills. The distribution of predicted ratings is discrete and can be naturally grouped into 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.

Table A.5: Placebo Tests

	(1)	(2)	(3)	(4)
	Dependent Variable: Hiring Interest			
After White Man	-0.172**			
	(0.086)			
Before White Man	0.038		0.041	
	(0.086)		(0.086)	
After Low-Quality White Man		-0.264**		
		(0.116)		
Before Low-Quality White Man		0.006		0.015
		(0.118)		(0.118)
Observations	2,736	2,736	2,808	2,808
R-squared	0.494	0.494	0.488	0.488

Notes: All regressions use the specification in column (4) of Table 1 but replace the original *After White Male* variable with different variables shown in the table. The number of observations in columns (1)–(2) is smaller because the sample also excludes the last resume for each employer ($2736 = 2808 - 72$). Robust standard errors are in parentheses.

Figure A.3: 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.