From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

Doctoral Dissertations, 2011-2026 Sociology

2021

Organizational Diversity Programs and Individual Outcomes

Xu, Janet

Princeton, NJ : Princeton University

http://arks.princeton.edu/ark:/88435/dsp01dr26z1540

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Organizational Diversity Programs and Individual Outcomes

Jiayi Janet Xu

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Sociology
Adviser: Matthew J. Salganik

November 2021

© Copyright by Jiayi Janet Xu, 2021.

All Rights Reserved

Abstract

In recent decades, opportunity-enhancing affirmative action programs – such as selective diversity scholarships and pipeline programs – have proliferated in universities and a wide range of industries. Emphasizing individual merit and conceptualizing "contributions to diversity" in a broad, race-neutral manner, these initiatives are appealing policy tools for organizations facing demands to diversify while navigating a legal environment that constrains what they can do to achieve those goals. Previous social scientific research has focused on the organizational-level outcomes of diversity initiatives, providing few insights on how programs affect their intended beneficiaries' individual-level labor market outcomes. This dissertation provides new empirical evidence on the individual-level effects of opportunity-enhancing diversity initiatives and contributes to our understanding of how organizations shape individual-level outcomes and social inequality.

The first empirical chapter uses a national audit study to investigate the reputational effects of diversity merit scholarships on recent college graduates' entry into the labor market, and how these signals vary by the racial cues associated with applicants' names. The second empirical chapter draws on an original survey experiment of hiring managers to further elucidate the determinants and mechanisms of affirmative action stigma. The third empirical chapter leverages archival administrative data from a large diversity postdoctoral pipeline program to compare the career trajectories of postdoctoral fellows and the alternates and finalists who were not selected for a fellowship.

Acknowledgements

This dissertation would not exist without the help and support of many, many people.

Matt Salganik has provided encouragement and nuggets of wisdom throughout graduate school. By pushing me to ask bigger questions, pursue more creative research strategies, and write more accessibly, Matt has made me a better scholar. Even more importantly, Matt has shown me how to be a good member of the scientific community. I have always been inspired by Matt's vision of what science should be – collaborative rather than competitive, committed to advancing collective knowledge over individual gains – and I will strive to uphold and pass on those ideals.

The other members of my committee, Adam Goldstein, Nicole Shelton, and Janet Vertesi, have also given me valuable feedback and steadfast support throughout the years. Our conversations frequently prompted me to look at my research from a new angle, and it is better because of it. I am also grateful to Fred Wherry for his insight and for serving as an external reader.

Rola Adebogun, Mel Hornyak, and Yael Marans provided excellent research assistance for the field experiment in the first chapter. The learning curve for conducting an audit study felt steep at times, and I am grateful to the other audit study researchers who made themselves available for conversations. I am lucky to have received multiple rounds of feedback from David Pedulla and S. Michael Gaddis, in particular.

Mark Lawson generously allowed a graduate student outside his institution to study the program he directs. I am grateful for the support I've received from him, Kim Adkinson and the rest of the PPFP staff.

This research was made possible by grants from the National Science Foundation, the Washington Center for Equitable Growth, and from multiple internal sources at Princeton, including the Fellowship of SPIA Scholars, the Center for the

Study of Social Organization, Princeton Research in Experimental Social Science, and the Data-Driven Social Science Initiative.

I am fortunate to have found many pockets of intellectual community at Princeton. I will forever be grateful to Stacey Sinclair, Nicole Shelton, Betsy Levy Paluck, and their lab members for welcoming a sociologist into their midst. Many faculty members in the Sociology department and the Office of Population Research have shaped my academic trajectory, but I would like to thank Brandon Stewart for teaching me how to think about statistics and Bob Wuthnow for teaching me how to think about culture, in particular.

I am indebted to all the administrative, library, custodial, computing, and statistical support staff in Wallace Hall who made Princeton such a comfortable and convenient place to learn and work. In particular, Amanda Rowe and Lynne Johnson have helped me stay on top of requirements throughout graduate school, Steve Mars and Eric Altman have provided essential technical support, and Boriana Pratt and Dawn Koffman helped me stay sane whenever I needed help debugging code.

My experience at Princeton has been enriched by the other graduate students I was lucky to befriend and learn from. There are too many to mention each individually, but some warrant special acknowledgement. Jeremy Cohen, Daniela Urbina-Julio and Simone Zhang have been a steady source of support, feedback, and friendship. They have had the dubious honor of hearing and reading my ideas at their earliest stages, and their thoughtful reactions have improved my work. Above all, our writing and accountability meetings gave me something to look forward to every week. Thank you for getting me through grad school. I also have to thank Gracie Himmelstein, Lauren Feldman, and Chris Felton for making me laugh and for many thought-provoking random conversations throughout the years.

My family and friends outside of grad school have been my biggest cheerleaders, and I feel so lucky to have support from such an incredible group of people. I

would especially like to thank my parents, Xu Feng and Xu Baohong, who cultivated my curiosity, modeled kindness and generosity, and always supported my decisions, even when they didn't understand them.

Finally, to Eugene, whose love, patience, and humor have sustained me. From proofreading the first grant application to letting me take over your personal computer to run Visual Basic code, your support runs through every page of this dissertation. Thank you for making the lows feel higher and the highs feel stratospheric.

Contents

Abstract iii

Acknowledgements iv

1 Introduction **1**

2 Race and the Reputational Effects of Diversity Scholarships in the Labor Market **5**

2.1 Signaling and Affirmative Action Stigma 9

2.2 Hypotheses 12

2.3 Methods 15

2.3.1 Materials 15

2.3.2 Sampling and Randomization 19

2.4 Results 21

2.5 Discussion 25

3 The Social Psychological Roots of Diversity and Affirmative Action Stigma **29**

3.1 Background 30

3.1.1 Competence, Warmth, Affirmative Action Stigma 30

3.1.2 Achievement, Selectivity, and Signal Clarity 32

3.1.3 The Racialized Effects of Awards 34

3.2 Methods 36

3.2.1 Experimental Design and Randomization 36
3.2.2 Survey and Variable Construction 40
3.3 Results 42
3.3.1 Recommendation for Interview 42
3.3.2 Revisiting the Affirmative Action Stigma of Incompetence and Low Warmth 51
3.3.3 Perceived Selectivity and Signal Uncertainty 53
3.3.4 Alternative Explanation: Are Diversity Scholarships Altering Sociodemographic Cues? 55
3.4 Discussion 60
3.4.1 Limitations and Questions for Future Research 62
3.4.2 Policy Implications 64

4 Pipeline Solutions for Promoting Diversity: Lessons from Academia 66

4.1 Background 67
4.1.1 Inequality in the Academy 67
4.1.2 The Structural Role of Postdocs 69
4.2 Case Study: UCPPFP 72
4.2.1 Overview and History 72
4.2.2 Eligibility and Selection Process 76
4.3 Methods and Data 80
4.3.1 Constructing the Dataset 80
4.3.2 Overview of the 2014-2015 Cohort 83
4.4 Results 89
4.4.1 Fellows and Finalists' Current Positions 89
4.4.2 Retaining Fellows in the UC System 93
4.4.3 Alternative Paths 95
4.5 Discussion 97

4.5.1 Limitations 99
4.5.2 Future research and extensions 101

5 Conclusion **103**

A Supplemental Materials for Chapter 2 **105**
A.1 University Reputations and Scholarship Interpretations 105
A.2 Racial and Class Cues of Names 108

B Supplemental Materials for Chapter 3 **111**
B.1 Causal Mediation Analyses of Order Effects 111
B.2 Additional Causal Mediation Analyses of Perceived Race and Sexual Orientation 112

C Supplemental Materials for Chapter 4 **113**
C.1 Disciplines and Interdisciplinarity in the UCPPFP Evaluation Process 113

Chapter 1

Introduction

Contemporary organizations face growing public demands to promote diversity and inclusion while navigating a legal environment that constrains what they can do to diversify. As a result, many institutions have implemented diversity initiatives that take a particular form: selective scholarships, fellowships, or pipeline programs that emphasize individual merit and conceptualize "contributions to diversity" in a broad, race-neutral manner. Far from the original stereotype of affirmative action as a form of unearned "special favors", these initiatives are depicted as a form of earned honors and a recruiting tool for attracting highly sought-after candidates. Such initiatives can now be found across sectors as varied as education, law, arts and entertainment, finance, and science and technology.

The effectiveness of diversity programs is of great interest to social scientists, advocates, and decision-makers who seek to broaden access to labor market opportunity, and these newer merit-oriented initiatives are no exception. Do they increase organizational diversity? Do they have a positive impact on the career outcomes of people from historically excluded groups? To date, studies have mostly focused on the former question of organizational-level diversity, rather than the latter question of individual-level outcomes. Yet these are two distinct issues, and gains at the orga-

nizational level do not necessarily reflect improvements for the people that diversity initiatives target.

In this dissertation, I draw on experimental methods and observational administrative data to examine two types of university diversity initiatives: undergraduate merit scholarships and postdoctoral fellowship programs. Together, these two cases capture universities' dual status as both distributors of educational credentials and employers in a specialized labor market. Tasked with different goals and target populations, the two strategies carry different labor market implications. For example, merit scholarships used by admissions officials to shape the composition of the student body may function as an additional informational signal for recipients when they apply for entry-level jobs, but they do not construct new pathways into the labor market. Postdoctoral pipeline programs, on the other hand, are more akin to internal labor markets found in a variety of other industries and organizational settings. Critically, however, neither case formally limits eligibility by race, ethnicity, or other demographic traits. Thus, I am able to examine how the returns of diversity initiatives differ for members of various racial and ethnic groups, illustrating their implications for inequality.

In chapter 2, I investigate how diversity merit scholarships affect recent college graduates' entry into the labor market. I focus specifically on diversity scholarships' reputational consequences and what they signal to prospective employers. In an audit study, I sent fictitious yet realistic applications to job openings to compare the rate at which employers call back applicants with diversity scholarships, applicants with other types of merit scholarships, and applicants with no scholarships at all. I also investigate how the scholarship signals differ when paired with applicant names that are widely perceived as Black rather than White. Results suggest that the effects of the scholarships diverge for putatively Black and white applicants. While putatively White applicants with diversity scholarships and non-diversity merit

scholarships receive significantly more callbacks than White applicants without scholarships, putatively Black applicants do not receive similar benefits from either type of scholarship. Black diversity scholarship winners are treated more similarly to Black applicants without scholarships than Black applicants with non-diversity merit scholarships. However, contrary to older theories of affirmative action stigma, I do not find evidence that Black diversity scholarship applicants are penalized relative to Black applicants without scholarships.

To better understand the mechanisms underlying these outcomes, I use the same study materials to conduct a survey experiment on a sample of hiring professionals. Chapter 3 evaluates existing social psychological theories of affirmative action stigma and examines the role of perceived selectivity in explaining the reputational effects of merit scholarships. Contrary to older theories based on a conception of affirmative action as preferential treatment, I find that newer, merit-oriented diversity awards do not shape perceptions of job applicants' warmth and competence. However, hiring decision-makers perceive diversity merit scholarships as less selective than scholarships that do not explicitly reference diversity. Moreover, perceptions of selectivity causally mediate the overall reputational gains of a having merit scholarship for Black applicants, but not for White applicants.

Experimental approaches are ideal for examining immediate reputational effects, but they cannot speak to the longer-term impacts of more structural interventions. Thus, in Chapter 4, I draw on administrative data from one of the oldest and largest diversity postdoctoral fellowship programs in the country to examine how pipeline programs shape professionals' career trajectories. I leverage a natural experimental set-up: in this competitive system, fellowship recipients and shortlisted finalists are similarly qualified, and there is a high degree of randomness to receiving a fellowship in a given year. For one fellowship cohort, I merge administrative data on award decisions with public information manually compiled from the Internet to

compare the career trajectories of fellowship recipients against those of finalists who did not receive a fellowship. I find that fellows are more likely to hold tenure-track faculty positions than finalists, and this trend is mostly driven by scholars in the life sciences. I also identify the alternative paths that finalists took in lieu of the selective fellowship, finding that most finalists took other postdoctoral research positions or non-tenure-track academic positions. Scholars who eventually left the academy did not leave immediately after obtaining their Ph.D.s. These results suggest that the strength of the diversity postdoctoral program is not in retaining talented women and minority scholars in the academy immediately after the Ph.D., but in converting postdoctoral opportunities into more permanent faculty positions.

Chapter 2

Race and the Reputational Effects of Diversity Scholarships in the Labor Market

Although diversity has become a widely touted organizational goal, racial minorities, especially Black Americans, remain underrepresented in many U.S. universities and workplaces. In response, organizations have heightened efforts to recruit individuals that will contribute to diversity. Among these efforts, opportunity-enhancing programs such as scholarships enjoy particular prominence because they are consistent with the narrative of individual effort (Bobo and Kluegel 1993). By portraying scholarships as competitive prizes based on past achievement, organizations can provide more attractive offers to certain desired recruits, meet diversity goals, and publicly signal a commitment to diversity without violating meritocratic norms. Indeed, diversity initiatives are often portrayed as good business practices rather than strictly efforts to remedy inequality for historically disadvantaged groups (Berrey 2015; Dobbin 2009).

Whether diversity programs such as specialized scholarships or pipeline programs are actually effective at improving racial representation in organizations has been of great interest to social scientists (Dobbin 2009; Holzer and Neumark 2000; Kalev et al. 2006; Reskin 1998). Yet a key feature of diversity scholarships and pipeline programs is that they advance organizational diversity goals by singling out particular individuals. The impact of these initiatives on individual recipients' future labor market outcomes are less well-understood. For example, universities trying to increase the number of racial minorities in their student body may recruit high-achieving students from underrepresented groups by offering them special merit scholarships, thereby differentiating them from other students, including their demographically similar peers. These awards may help persuade students to enroll, but how are they received when students transition into the labor market?

When organizations bestow diversity awards and resources, they publicly identify recipients as targets of diversity and affirmative action initiatives, and a body of management scholarship has shown that diversity initiatives may have varied unintended consequences for racial minorities and women at work (see Leslie et al. 2014 for review). However, much of the prior research focuses on subjective self-assessment at work (e.g., Brown et al. 2000; Major et al. 1994; Taylor 1994) rather than more material outcomes like hiring and promotion. Studies on perceptions of affirmative action hires have found that White women and racial minorities are stigmatized as less competent when they are framed as beneficiaries of preferential hiring programs (Heilman et al. 1992, 1996, 1997), but these studies rely on outdated cues and signals that are meaningfully different from the presentation and substance of contemporary diversity management practices (Gilbert and Stead 1999; Scarborough et al. 2019). Moreover, the bulk of this research was conducted two or three decades ago. As Americans' stated support for affirmative action has risen in the past decade (Norman 2019) and public demands for equity and representation have intensified,

beneficiaries of contemporary diversity initiatives may be less susceptible to the stigma of affirmative action.

Past work on affirmative action stigma is further complicated by a legal landscape that has been increasingly hostile to its practice (Kellough et al. 1997; Harper and Reskin 2005), especially in higher education. For example, since the mid-1990s, the percentage of colleges and universities publicly claiming their consideration of race in undergraduate admissions has declined dramatically (Hirschman and Berrey 2017). Court rulings such as 2003's Gratz v. Bollinger prompted the institutionalization of new individualized systems of evaluation that conceptualize "contributions to diversity" in a broad manner (Kirkland and Hansen 2011), encompassing not just race and ethnicity but other sociodemographic and even personal dimensions (Edelman et al. 2001). Diversity management practices, therefore, are just another tool that managers can use to achieve broader organizational goals. Take, for instance, university admissions officials' use of merit scholarships – including diversity scholarships – to shape the composition of an entering class in terms of not just race and ethnicity, but also geography, socioeconomic background, and standardized test scores (Kantrowitz 2011). As such, older studies that only examine White women and racial minorities as beneficiaries of diversity practices may overstate organizations' focus on targeting these groups specifically and fail to capture the full scope of diversity initiatives' impact on other populations.

In this chapter, I investigate the reputational effects of diversity scholarships for recent college graduates entering the labor market. Using an audit study, I compare the rate at which employers call back applicants with diversity scholarships, applicants with other types of merit scholarships, and applicants with no scholarships at all. I also examine how the scholarship signals differ when applicant names have a Black racial cue versus a White racial cue. In essence, do employers discriminate

against beneficiaries of diversity policies, and if so, are the effects unique to putatively Black applicants?

I find that the effects of the scholarships diverge for Black and White applicants[1]. While putatively White applicants with diversity scholarships and non-diversity merit scholarships receive more callbacks than White applicants without scholarships, putatively Black applicants do not receive similar benefits from either type of scholarship. Black diversity scholarship winners are treated more similarly to Black applicants without scholarships than Black applicants with non-diversity merit scholarships. However, contrary to older theories of affirmative action stigma, I do not find strong evidence that Black diversity scholarship applicants are penalized relative to Black applicants without scholarships.

This research has implications for several areas of sociological research. First, it provides an important update to knowledge on organizations and diversity policy effectiveness. By focusing on the individual-level impact of scholarship programs on future labor market outcomes, I show that Black individuals who are recruited for contributions to organizational-level diversity may not always personally reap the rewards of such recognition. Additionally, because the study design isolates the reputational effects of diversity scholarships from the impact of any material intervention, this research demonstrates the limits of purely symbolic diversity measures. Furthermore, this research presents a less-considered counterfactual for thinking about who diversity initiatives impact. Frequently, research conceptualizes diversity and affirmative action practices as policies that enable individuals from historically excluded groups to gain entry into organizations that they would not have had opportunity to access otherwise. While this was the original ethos of affirmative

[1]I refer to applicants as "Black" and "White" throughout this chapter for brevity, but acknowledge that these applicants are not really "Black or "White". Race, as a construct, cannot be fully experimentally manipulated (Sen and Wasow 2016). What I actually experimentally manipulated was the employers' exposure to a particular racial cue, as signaled by the applicants' names on the resumes.

action when first instituted, this narrow focus overlooks the class of contemporary outreach initiatives are in fact "in-reach" efforts (Rivera 2012) oriented towards recruiting high-achieving individuals who often already have access to conventional opportunities.

This research also contributes substantively to recent research on diversity in cultural sociology and to social stratification research on the differential effects of credentials and achievements accrued in educational organizations. Cultural sociologists have consistently shown how organizational gatekeepers and managers use "diversity" in a broad, ambiguous, and colorblind manner that obscures historical power asymmetries and elides concrete efforts at addressing racial inequality (Berrey 2015; Edelman et al. 2001; Embrick 2011). This type of diversity ideology suggests that everyone, regardless of race, contributes to it (Kirkland and Hansen 2011). The present study extrapolates this premise by probing how contributions to "diversity" are perceived differently for applicants who present different racial cues, and what consequences these signals have in the labor market.

2.1 Signaling and Affirmative Action Stigma

According to classical signaling theory, credentials like scholarships provide a meaningful signal to employers, as they distinguish recipients from the rest of the student body (Spence 1973). Indeed, in her ethnographic research on elite recruiting, Rivera (2012) describes how one recruiter told applicants at a diversity job fair to list "full scholarship" on their resumes if they attended a less prestigious school for financial reasons. Moreover, because scholarships are usually viewed as signs of individual effort (Bobo and Kluegel 1993), they may be particularly helpful for members of groups bearing negative stereotypes, such as Black Americans, by providing information that counters such stereotypes (Pager and Karafin 2009).

However, diversity scholarships could send different signals than other types of merit scholarships due to their association with affirmative action. Many managers tend to see affirmative action as an unacceptable violation of merit (Foley and Williamson 2019) and people who appear to be hired under affirmative action policies are stigmatized (see Leslie et al. 2014 for review). White women and racial minorities who are identified as affirmative action hires are generally seen as less competent than White women and minorities who are unaffiliated with affirmative action initiatives (Heilman et al. 1992, 1996, 1997). A recent meta-analysis suggests that minority affirmative action beneficiaries are also seen as less warm and more threatening because affirmative action policies can make their group appear more competitive (Leslie et al. 2014). Indeed, this type of threat is consistent with the prior documentation that a "mere mention" of preferential treatment for Black Americans led to negative generalizations about them as a group (Sniderman and Piazza 1995).

Yet the literature on affirmative action stigma also contains several caveats. For one, the modern rhetoric of diversity management is a sharp contrast to the explicit references to "preferential hiring" found in the experimental stimuli of the canonical studies on affirmative action stigma. Indeed, as Scarborough and colleagues have pointed out, "few workplaces use the language of affirmative action to describe their workplace policies", and "although there are many different types of workplace diversity policies used today, the vast majority of them do not fit the prototypical and controversial preferential hiring programs that people may think of when they are asked about affirmative action" (Scarborough et al. 2019, 195). The framing and rhetorical portrayal of modern diversity policies have consequences for how targets are perceived. For example, when selection is framed as a result of demographics and merit, the negative consequences of affirmative action are lessened (Major et al. 1994). Racial minorities and women who are portrayed as being hired under "diversity

management" are seen more favorably than those who are framed as being hired under "affirmative action" (Gilbert and Stead 1999).

More recent accounts have even posited that evaluators view contributions to diversity as one component of merit after candidates have already been screened for traditional indicators of achievement (Posselt 2016). In this view, an individual's proven contribution to organizational diversity is an instrumental and sought-after quality in its own right, especially when organizations rely on racial minorities to perform the "racial tasks" of smoothing interactions between White and minority peers (Wingfield and Alston 2014) or feel pressured to increase their "diversity numbers" (Rivera 2012).

Finally, attitudes towards affirmative action policies – especially when they are framed as preferential treatment— may be sensitive to the social identity to the policy targets. For example, public opinion polls show that affirmative action on the basis of race tends to receive slightly less public support than affirmative action on the basis of other social dimensions (Norman 2019). Social psychologists have theorized that affirmative action policies targeting Black Americans are viewed more negatively than affirmative action for other groups, such as White women (Eberhardt and Fiske 1994), and indeed, there is evidence that White people are less likely to support race-targeted diversity workplace policies than gender-targeted policies (Scarborough et al. 2019). Thus, the presence and degree of affirmative action stigma that individuals experience may be contingent on other social cues, such as information about their race.

Given that diversity management differs from older affirmative action practices rhetorically and structurally, the importance of social cues about recipients' demographic identities is especially noteworthy. As diversity practices have become increasingly decoupled with legally protected social categories, the presentation of identities among diversity scholarship recipients may, themselves, be more diverse.

"Diversity" may refer to a wide array of dimensions of difference, which could refer to race and ethnicity, but could also include other types of social statuses not protected by anti-discrimination law (e.g., socioeconomic class, geographic origin.), or even personality traits (Edelman et al. 2001). Some individualistic evaluation systems may even consider actions that dominant group members' have taken to demonstrate commitment to diversity as contributions to diversity (Kirkland and Hansen 2011).

2.2 Hypotheses

Taken together, existing theories of signaling and affirmative action stigma suggest four potential outcomes for how employers will evaluate diversity merit scholarship honorees relative to other types of merit scholarship winners and applicants who do not have scholarship awards. (summarized in Figure 2.1). First, if diversity scholarships provide the same signals of distinction and achievement as other types of merit scholarships, we would expect employers to call back demographically similar applicants with diversity and non-diversity scholarships at comparable rates. Under this hypothesis, diversity scholarships do not carry a detectable stigma associated with affirmative action, and candidates are not treated differently for being associated with organizational diversity efforts.

Alternatively, employers may dole out a diversity penalty, calling back diversity scholarship winners at lower rates than demographically similar applicants who have won other types of merit scholarships. Under this hypothesis, employers may prefer applicants with diversity scholarships to applicants who don't have any scholarship awards at all or treat them similarly. That is, employers may discount diversity awards relative to other awards, but the diversity scholarship does not delegitimate the applicants' other achievements.

On the other hand, a more extreme version of a diversity penalty results in a diversity backlash, wherein employers do penalize diversity scholarship holders even relative to applicants without any scholarships. Under this hypothesis, employers don't just discount applicants' achievements relative to other applicants who hold non-diversity related merit scholarships, but actually treat them as less preferable than applicants who don't have any accolades. Like the theory of affirmative action stigma documented in the 1990s, this hypothesis predicts that an individual's association with diversity or affirmative action will call their overall ability into question. Whereas the diversity penalty hypothesis predicts that employers may devalue the meaning of the diversity scholarship award itself, the diversity backlash hypothesis predicts that employers will devalue the person with the diversity scholarship.

Finally, there is the possibility that diversity scholarship winners may actually experience a diversity bonus, because diversity-related merit scholarships recognize contributions that non-diversity merit scholarships don't. Under this hypothesis, employers may prefer applicants with diversity scholarships over demographically similar applicants with non-diversity scholarships and applicants without scholarships.

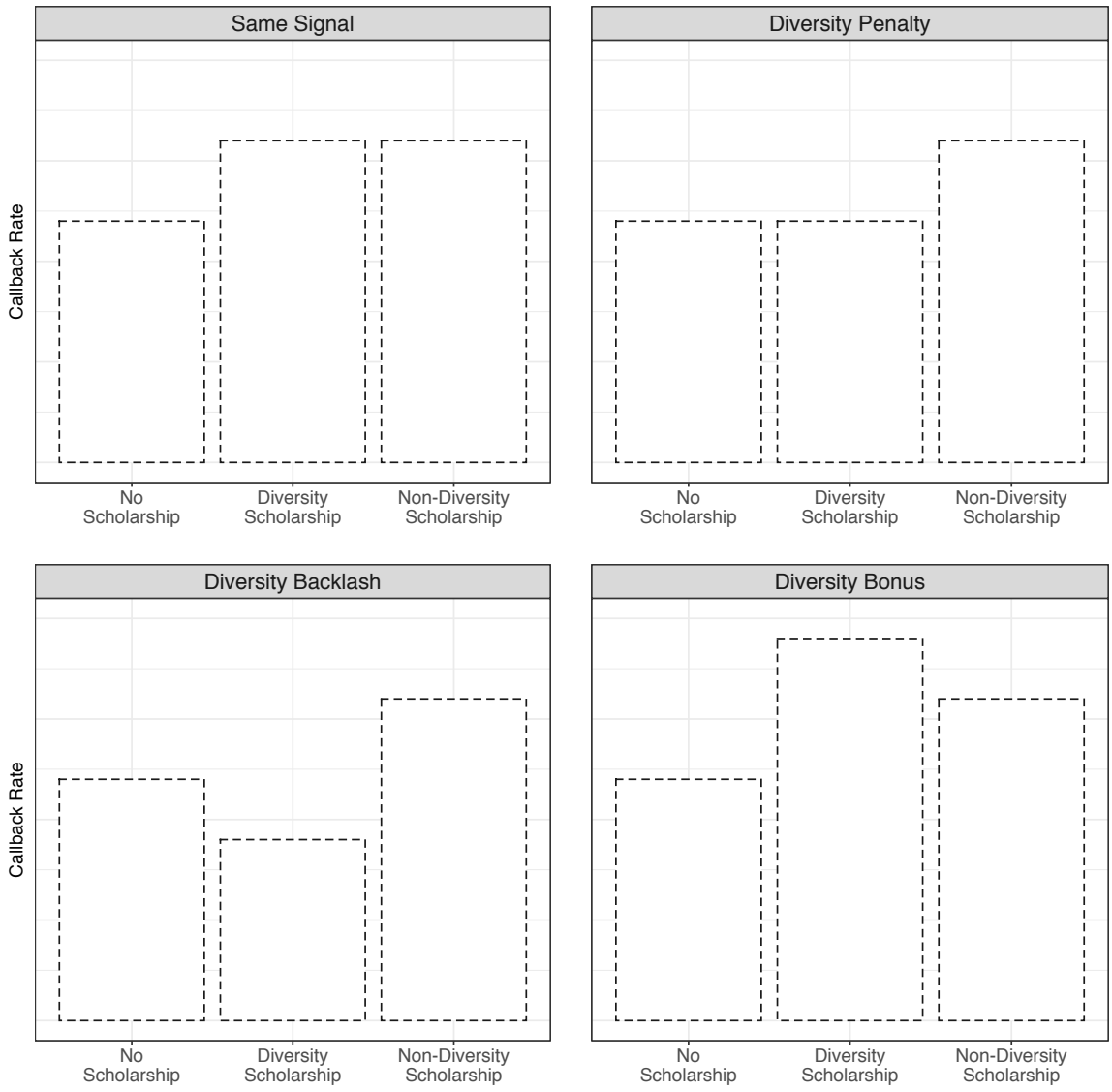


Figure 2.1: Potential outcomes for how employers will evaluate diversity merit scholarship honorees relative to other types of merit scholarship winners and applicants who do not have scholarship awards.

The signal of various scholarships may also be interpreted differently depending on exposure to the racial cues provided by other aspects of the application, such as racialized names. For example, predominantly White educational institutions often evince a demand for Black students, but in practice, White admissions counselors "weed out" correspondence from Black students who they perceive as "too concerned" with race and racism Thornhill (2019). On the other hand, White or male organizational leaders are not typically penalized for diversity-valuing behavior Hekman et al. (2016). Thus, diversity scholarships may be more likely to generate a penalty or backlash for applicants with widely perceived Black names than applicants with stereotypically White names.

2.3 Methods

To examine how employers evaluate applicants with diversity merit scholarships relative to applicants with other types of merit scholarships and applicants without scholarships, and how this may differ by applicants' racial cue, I conducted a multi-city field experiment. Between July and October 2020, I submitted 3,456 fictitious job applications to 1,728 real job openings. The primary outcome of interest is whether a job applicant receives a callback – a positive response from the employer that advances them to the next stage of the hiring process. I tracked all positive employer responses received by each application via email, voicemail, and text messages. The experiment was pre-registered on Open Science Framework.

2.3.1 Materials

The job applications featured experimentally manipulated racial cues (signaled with names that are often perceived as Black or names that are often perceived as White) and scholarship statuses (diversity merit scholarship, non-diversity merit scholarship,

or no scholarship) for a total of six conditions. All applicants were portrayed as recent college graduates and had names with male cues. I focus on men to avoid signaling gender-based affirmative action in addition to race-based affirmative action, especially because women are not underrepresented in undergraduate populations. I selected names of real diversity and non-diversity merit scholarships from three mid-sized, selective (but not elite) private universities (schools ranked 15 to 45 by U.S. News and World Report). I then created two standardized versions of stylized descriptions to be used across all three schools. These descriptions are composites of language found in the real scholarship descriptions (see Table 2.1). On resumes, diversity and non-diversity scholarships are presented in the education section, which is at the top of the document. Resumes that do not feature a scholarship simply omit those lines in the education section.

Table 2.1: Scholarship Award Desciptions used in Experimental Conditions

Experimental Condition	**Version A**	**Version B**
Diversity Scholarship	Recognizes students who enhance the overall quality and diversity of the student body through academic achievements, leadership, community service, and diverse life experiences	Awarded to select students on the basis of academic accomplishments, leadership, service to their communities, diverse life experiences, and contributions to diversity
Non-Diversity Scholarship	Recognizes students who enhance the overall quality of the student body through academic achievements, community service, and leadership	Awarded to select students on the basis of academic accomplishments, service to their communities, and leadership
No Scholarship	–	–

The descriptions of the two scholarships were designed with three criteria in mind. First, they must model the language of real scholarships as closely as possible. Notably, this language makes no explicit reference to race and instead uses the abstract rhetoric of diversity management. Second, both types of scholarships needed to convey that the prizes are awarded on the basis of academic and extracurricular achievement; however, the diversity scholarship needed to additionally signal the consideration of demographic identity and/or personal experiences as a selection criterion. Finally, the scholarships should not be nationally known or overtly prestigious, since hiring managers could have personal attachments with special meaning to well-known awards that they have prior familiarity with. Thus, I use generically named university-specific scholarships that most people are unlikely to be very familiar with, which gives me greater control in crafting comparable but realistic award descriptions while reducing the likelihood that they would carry personalized meanings for evaluators [2].

I pre-tested the names and descriptions of the scholarships and names of universities on an online convenience sample of U.S.-based adults with hiring experience. I chose three universities that were rated nearly identically for perceived selectivity and alumni reputation. When asked to speculate on the criteria associated with the diversity and non-diversity scholarships at these three schools, respondents attributed both scholarships to good academic performance, extracurricular involvement, and leadership. However, respondents were more likely to think that underrepresentation

[2]To further confirm the external and ecological validity of using a diversity merit scholarship and non-diversity merit scholarship as counterfactuals, I researched merit-based awards on scholarship search engines and reviewed the financial aid websites of universities that offer merit aid (current industry estimates do not specify the prevalence of merit-based scholarships). Almost all schools offering diversity scholarships also offer scholarships that do not mention diversity. Selection processes are opaque, and I could not find any institutional data on the demographic composition of merit scholarship recipients. To confirm that diversity scholarship recipients and other merit scholarship recipients resemble each other in profile, and that both sets of recipients are racially heterogenous, I searched the scholarship names on LinkedIn to identify previous recipients. My searches showed that putatively Black and putatively White students both receive both kinds of scholarships.

or minority status was considered as a criterion for the diversity scholarship. Results from these pre-tests can be found in the Supplemental Materials.

To signal race, I manipulated the applicant names. I chose names from a list published by Gaddis (2017), which measured the perceived race of common first names from New York state birth records and common U.S. surnames. In another pre-test with a different online convenience sample of U.S.-based adults with hiring experience, I presented participants with a subset of names from Gaddis (2017) and asked them to associate each name with race and socioeconomic class. I selected two names that were perceived as Black and middle-class most of the time, and two names that were perceived as White and middle-class most of the time. Results from these pre-tests are also presented in the Supplemental Materials.

Other parts of the resume (e.g., major, GPA, work experience) are based on example resumes found on LinkedIn and resume banks, including sample resumes from university career centers. All applicants were the equivalent of business management majors at their respective universities. To maintain consistency with the idea that these applicants may be particularly high-achieving students, all the resumes listed 3.7 GPAs, which is higher than average at all three institutions. The resumes also listed two internships: one as a "summer analyst" at a boutique consulting or financial management firm (industry varied across cities but were always matched within cities), and one as a business operations/product management intern at a small technology start-up. Contact information (area code of phone numbers and residential addresses) and the names and locations of previous employers were customized to match the local labor market so that employers would not penalize applicants for being non-local. Every piece of information on the resumes, including applicant names, scholarship names, and the template itself, were randomized and counter balanced.

Finally, because many companies now use automated Applicant Tracking Systems (ATS) software to screen resumes as a first step, I used a resume-optimization service that ran both resume templates through generic ATS software. This step ensured that my resumes had a higher-than-average chance of bypassing automated screening. More importantly, this step also allowed me to test that the two templates in each pair were parsed and scored similarly and improves my confidence that the two resumes in a matched pair would be treated similarly by ATS. While I cannot observe if an employer rejected an application through ATS software or human judgement, this step provides some evidence that it is unlikely for any one resume in a matched pair to get rejected by ATS while the other resume makes its way to a human manager.

2.3.2 Sampling and Randomization

I created a multi-city sample of job openings by web-scraping a large job board aggregator for full-time, entry-level positions. Using public LinkedIn profiles of recent college graduates who are similar to the fictitious applicants (e.g., business majors from selective universities) as well as publicly available post-graduation data published by the three university career centers, I compiled a list a common post-graduate jobs and used these jobs as the occupational search terms in my web-scraping script. The search terms were business analyst, financial analyst, operations analyst, program analyst, product manager, project manager, consultant, business development, account executive, marketing analyst, and executive/administrative assistant[3]. Job listings related to these search terms tend to require or strongly prefer a college degree but no specialized training[4]. To further ensure that these positions were realistic oppor-

[3]I did not specify exact match for the search terms, so each term could return job listings for titles that were similar but not identical to the key terms (e.g., "business operations analyst" when the search term is "business analyst").

[4]Administrative assistant positions tended to state a wider range of qualifications. Some required college degrees while others were open to applicants who did not graduate high school or hold a GED. To maintain a realistic match between the applicant profiles and the job opening, I applied to administrative assistant positions that required or indicated explicit preference for a college degree.

tunities that these applicants would actually pursue, I filtered out jobs that explicitly required fluency in a foreign language, specialized licenses, active security clearances, or niche industry-specific experiences that they did not possess.

I pulled job listings for positions within 25 miles of eight major cities: New York, Boston, Chicago, Detroit, Dallas, Washington D.C., Los Angeles, and San Francisco. I chose the two metropolitan areas with the largest GDP in each of the four geographical regions (Northeast, Midwest, South, and West). None of the universities listed on the resumes were located in one of these metropolitan areas. The distribution of job applications by metropolitan areas can be found in Table 2.1.

Table 2.2: Distribution of Job Applications across Metropolitan Areas

Metropolitan Areas	**% Applications**
Northeast	
New York City	18.1
Boston	11.7
Midwest	
Chicago	12.2
Detroit	6.9
South	
Washington D.C.	16.4
Dallas	9.5
West	
Los Angeles	16.2
San Francisco	9.0

N = 3,456 Applications

The field experiment used a matched-pair design, which consists of sending two pieces of correspondence to each job opening. When it is more common for pairs to receive the same outcome (e.g. no contact from the employer) than to receive different outcomes (e.g. only one gets a callback), matched designs tend to be more statistically efficient (Vuolo et al. 2016). Though the risk of detection and the risk of spillover effects both increase when more than one job application is sent to each

job opening, these risks decrease as applicant pools for each job grows larger, which tends to be the case in looser labor markets like the recessionary market during this study period.

Pairs of correspondence were always matched on racial cue (as signaled by names) but differ by scholarship status. To avoid detection, the two pieces of correspondence in a pair always featured different universities (and their associated majors and extracurricular activities), slightly different employer histories, extracurricular activities, and scholarship descriptions. They are also presented on different templates. Though the two versions feature the same list of skills at the bottom, the skills are listed in different order. All application features are randomized and counterbalanced, so they do not confound the main treatments of interest. Each pair of applications was randomly assigned to a job opening. Applications were submitted one day apart, with the order also randomized. To minimize the chance that an employer would respond positively to a second application only because the first applicant declined an interview, I waited 5-7 days before declining callbacks. Among all job openings that called back both applicants, 99% called back the applicants within 7 days of each other.

2.4 Results

The overall positive response rate was around 11.3%. However, as Figure 2.2 shows, the callback rates differed by experimental condition. Among job applications with Black racial cues, applicants without scholarships listed on their resumes and applicants with diversity scholarships listed on their resumes had similar callback rates (8.7% and 8.2%, respectively; t = -0.32, p = 0.75). Applicants with non-diversity merit scholarships listed on their resumes had a slightly higher callback rate of 10.4%, but this difference is not statistically significantly larger than the callback rates for

diversity scholarship recipients and applicants without scholarships (t = -1.32, p = 0.19 and t = -1.00, p = 0.32, respectively).

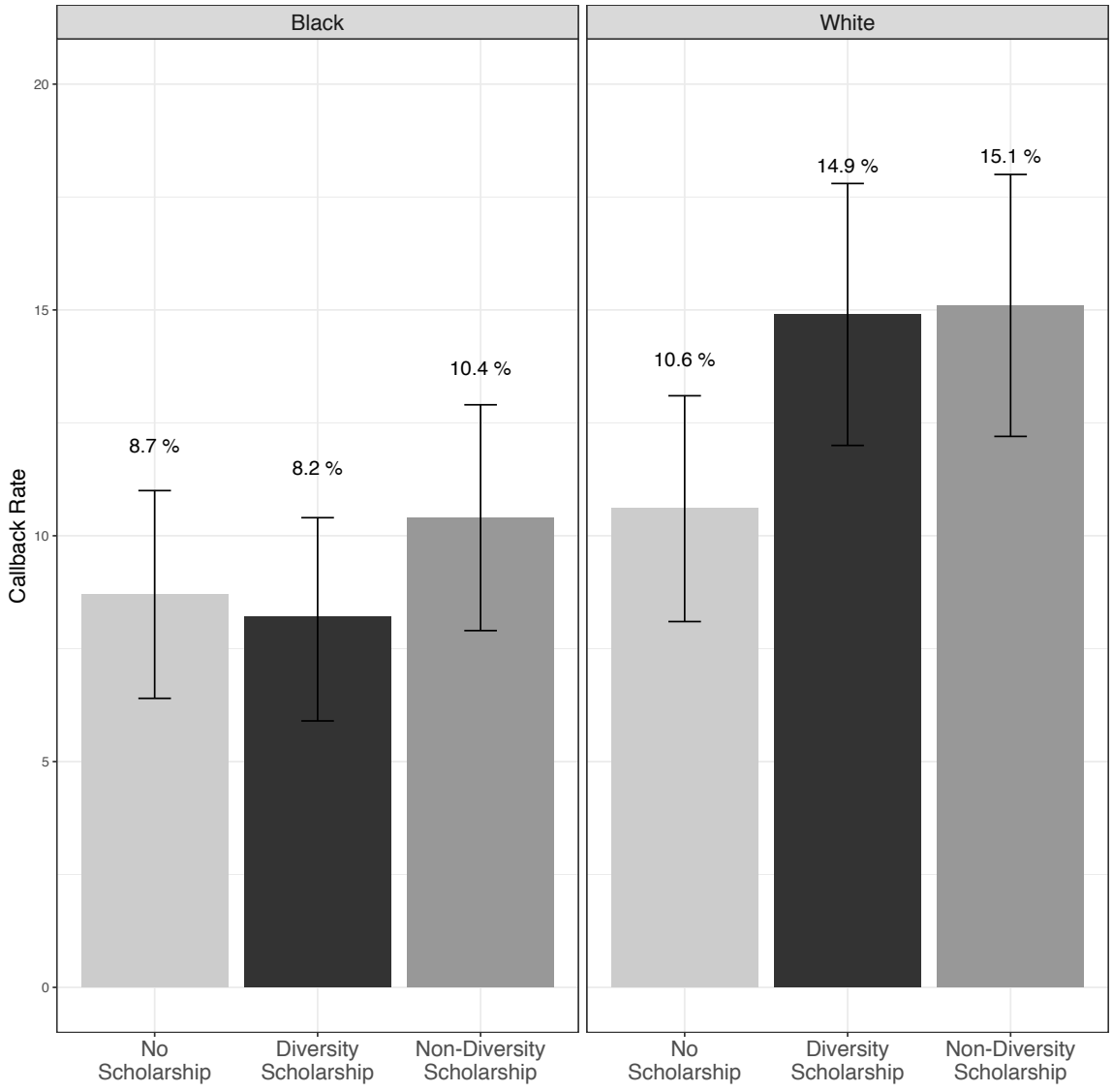


Figure 2.2: Callback rates for Black and White applicants without scholarships, with diversity scholarships, and with non-diversity scholarships.

On the other hand, among job applications with White racial cues, both applicants with diversity and non-diversity scholarships listed on their resumes received significantly more callbacks than applicants without scholarships. While 14.9% of applicants with diversity merit scholarships received positive responses, only 10.6% of applicants without scholarships did ($t = 2.21$, $p = 0.03$). About 15.1% of applicants with non-diversity merit scholarships received positive responses, which is significantly higher than applicants without scholarships ($t = 2.29$, $p = 0.02$), but comparable to applicants who had diversity scholarships listed on their resumes ($t =$ -0.08, $p = 0.27$).

Applicants with White racial cues consistently received more callbacks from employers than applicants with Black racial dues did. However, these Black-White differences are greater among applicants with scholarships listed on their resumes than among applicants without scholarships. The callback probability for putatively Black applicants is about 1.9 percentage points lower than that of putatively White applicants for those without scholarships ($t = 1.10$, $p = 0.27$), but it is 6.8 percentage points lower and 4.7 percentage points lower for those with diversity and non-diversity scholarships, respectively ($t = 3.61$, $p < 0.01$ for diversity scholarships; and $t = 2.39$, $p = 0.02$ for non-diversity scholarships). Scholarship awards appear to exacerbate racial discrimination because they increase callback probability for putatively White applicants more than they do for putatively Black applicants. Further, because putatively Black candidates receive virtually no benefit from having a diversity scholarship on their resume, but putatively White applicants gain as much from having a diversity scholarship listed as they do a non-diversity scholarship, the difference in callback rates between applicants with Black and White cues is starkest in the diversity scholarship condition.

Logistic regression models that account for the matched pair design by clustering standard errors at the job opening level and including city and month fixed

Table 2.3: Logistic Regressions Predicting Callback

	Black Applicants	White Applicants	All Applicants
	(1)	(2)	(3)
Scholarship Status (ref. No Scholarship)			
Diversity Scholarship	−0.090	0.401*	−0.080
	(0.214)	(0.181)	(0.183)
Non-Diversity Scholarship	0.186	0.409*	0.195
	(0.203)	(0.180)	(0.173)
Racial Cue (ref. Black Cue)			
White Cue			0.213
			(0.202)
Scholarship x Racial Cue (ref. No Scholarship x Black Cue)			
Diversity x White			0.483*
			(0.240)
Non-Diversity x White			0.224
			(0.226)
Constant	−2.113***	−1.928***	−2.100***
	(0.362)	(0.316)	(0.311)
Order, City and Month Fixed Effects	Yes	Yes	Yes
Observations	1,728	1,728	3,456
Log Likelihood	−507.491	−660.485	−1,181.463
Akaike Inf. Crit.	1,042.981	1,348.970	2,396.926
Note:			*p<0.05; **p<0.01; ***p<0.001

effects recover substantively similar estimates for the effect of scholarships on callbacks (Table 2.3). In Model 1, I regress callback outcomes on scholarship status among applicant profiles with Black racial cues and find no statistically significant effects for either type of merit scholarships. In Model 2, I regress callback outcomes on scholarship status among applicant profiles with White racial cues and find that both scholarships significantly increase the likelihood of receiving a callback.

Finally, in Model 3, I regress callback outcomes on scholarship status, racial cue, and the interaction of the two among all applicants. The interaction term for diversity scholarship and White racial cue is positive and statistically significant, suggesting that the reputational gains of diversity scholarships is greater for putatively White applicants than it is for putatively Black candidates. However, the interaction term for non-diversity scholarship and White racial cue is not statistically significant, suggesting that the reputational gains of the non-diversity scholarship does not vary significantly by applicants' racial cues. In sum, both types of awards significantly increase the chances of getting an employer callback for applicants racialized as White, but not for applicants racialized as Black, and the diversity scholarship award provides a uniquely positive boost for putatively White applicants.

2.5 Discussion

The results from the audit study suggest that diversity scholarships may pose a penalty for job applicants whose names provide a Black racial cue, but not for applicants who have a White racial cue. While the results do not support the idea that diversity scholarships might stigmatize applicants so much that they would receive worse treatment than peers without scholarships, they indicate that employers devalue the accomplishments of putatively Black applicants when their accolades are associated with diversity.

Moreover, the results also indicate that even non-diversity merit scholarships fail to generate rewards for putatively Black applicants at the same rate as they do for putatively White applicants. This is consistent with results from prior audit studies that have found also found employers to be more responsive to improvements in resume quality for applicants with "White" names than "Black" names (e.g., Bertrand and Mullainathan 2004; Kessler et al. 2019; Nunley et al. 2015). Thus, the results have implications not just for theories of affirmative action stigma, but for more general theories of cumulative advantage and "success breeds success" dynamics.

There are several potential reasons and social psychological mechanisms for these patterns. First, consistent with theories arguing that affirmative action decreases perceived competence, employers might devalue Black applicants' awards, and especially awards associated with diversity, because they perceive these awards to be less competitive than awards given to White applicants, which makes them a less informative signal. This, in turn, could be because employers already think that members of the same racial category only compete against each other for recognition, such that there is less competition when the size of the group is smaller. It could also be because biased employers perceive the quality of minority applicant pools to be lower. A second line of explanation – which is not mutually exclusive from the first – is that employers might think of an applicant less favorably after the "mere mention" of affirmative action. Consistent with theories arguing that affirmative action lowers perceived warmth, employers might view Black applicants as more competitive, and even threatening, after associating them with affirmative action and diversity efforts.

There are also several scope conditions and limitations to this study that future research can build on. First, it is difficult to predict how the effects observed in this study may extrapolate to population-level estimates because there is such a dearth of institutional data on the volume and composition of diversity scholarship recipients. Like many other types of organizational diversity initiatives, merit scholar-

ships are implemented in a patchwork fashion, and not necessarily with standardized documentation. More research on the prevalence of these practices and data on the composition of their policy targets would advance our knowledge of their effectiveness. Relatedly, future research should also be attentive to when diversity scholarship or fellowship recipients are more likely to disclose their scholarship status on application materials, and how evaluation outcomes might differ when the awards are more familiar to evaluators. For example, in recent ethnographic research, Portocarrero (2020) argues that intra-organizational practices play a key role in constructing stigma around being a national scholarship recipient. In her research site in Peru, scholarship recipients at an elite university often hid their scholarship status with other members of the organization. However, they were less ashamed of disclosing their status to people outside of the organization. For example, one scholarship recipient would list her scholarship on her CV when she applied to jobs outside of her university (similar to what the applicants in this study are doing) but would hide the scholarship when applying to jobs on-campus (Portocarrero 2020). The reputational consequences of being affiliated with a diversity program are likely to be different for internal mobility within organizations than for movement across organizations.

Another scope condition is that the racial cues were limited to Black and White. Moreover, the applicant names chosen for this study were designed to generate the strongest and most unambiguous racial cues. Research on "resume whitening" shows that when Black and Asian job applicants try to conceal their race on their resumes, they experience less discrimination (Kang et al. 2016). Future research is needed to understand if the patterns of diversity stigma and devaluation observed here would also apply to White women and Hispanic/Latino(a), Asian, and Indigenous people, as well as applicants who might have more racially ambiguous names.

Finally, this study was fielded during a very particular time period. Though the labor market was more stable during July - October 2020 than it was earlier in the

year, the economy was in a severe recession caused by the novel coronavirus pandemic. This not only reduced the pool of available jobs, but many traditional office jobs such as the ones in this sample were changed into remote positions. Though there is no existing evidence that perceptions of affirmative action and diversity beneficiaries are contingent on certain labor market conditions, it is possible that the effects found in this study might be attenuated if there is less supply-side competition. Indeed, prior research has suggested that discrimination around "weak" resume signals are more consequential when local unemployment is higher (Weisshaar 2018). Finally, racial inequality may have been particularly salient during this study period because of increased mainstream attention to the Black Lives Matter movement. As support for the social movement waned over the summer of 2020 among non-Black Americans (Thomas and Horowitz 2020), backlash towards diversity-related initiatives may have risen. Future research is needed to tease out the relationship between the Black Lives Matter movement and genuine support for affirmative action policies.

Limitations aside, the findings presented in this chapter has important implications for the design and implementation of diversity initiatives. Scholars across a range of disciplines have pointed out that a race-neutral and instrumental rationale for diversity – which has been institutionalized as the legal justification for achieving diversity in higher education – has moved the cultural and legal framework away from the original aim of affirmative action. While affirmative action was originally oriented towards righting historic wrongs by expanding opportunities for historically oppressed and excluded groups, diversity management is touted as a framework that benefits everyone in the organization. Yet the results of this study suggest that the spoils of such organizational efforts may not be shared evenly. In fact, they may be least effective for the people who stand to benefit the most.

Chapter 3

The Social Psychological Roots of Diversity and Affirmative Action Stigma

The previous chapter presented field experimental evidence that two different kinds of merit scholarships – those that mention diversity specifically and those that don't – have reputational effects in labor market and these effects differ by race. Having a non-diversity merit scholarship increases the probability of getting a callback for both Black and White applicants, but the effect was larger and statistically significant for White applicants[1] For White applicants, having a diversity scholarship also significantly increased the chances of getting a callback. By contrast, diversity scholarships were devalued relative to other types of merit scholarships for Black applicants. White applicants also received more callbacks across the board.

However, while field experiments are the "gold standard" method for detecting the *existence* of discrimination, they are inadequate for investigating the *how*

[1]Again, I refer to applicants as "Black" and "White" throughout this chapter for brevity, but acknowledge that these applicants are not really"Black or "White". Race, as a construct, cannot be fully experimentally manipulated (Sen and Wasow 2016).

and *why* aspects of discrimination (Gaddis 2019). In this chapter, I use a survey experiment to replicate the finding from the field experiment and to examine potential causal mechanisms that may explain why diversity and non-diversity scholarships are received differently by hiring decision-makers. I focus on the role of perceived competence and warmth, the role of perceived selectivity and signal clarity, and how these factors might be racialized.

3.1 Background

3.1.1 Competence, Warmth, Affirmative Action Stigma

Social psychological theories of affirmative action stigma would predict employers devaluing diversity scholarship recipients because diversity-related labels alter perceptions of candidates' competence and warmth. Competence and warmth are the primary dimensions of the stereotype content model (Fiske et al. 2002), an influential theoretical framework of interpersonal perception. This model posits that status is the key antecedent of competence-based stereotypes, while lack of competition is the key antecedent of warmth-based stereotypes (Cuddy et al. 2008; Fiske et al. 2002).

Early landmark studies on the unintended negatives consequences of affirmative action plans focused on the stigma of incompetence (Heilman et al. 1992, 1997). However, in a more recent meta-analytic review, Leslie et al. (2014) proposed that evaluations of affirmative action targets are also driven by perceptions of low warmth.

In particular, Leslie and colleagues argue that diversity and affirmative action programs simultaneously convey two things about their targets: they have low status, and they have received help in their ability to compete for resources. First, diversity initiatives raise the idea that the targeted group must have poor outcomes because they *require* extra aid (Maio and Esses 1998), and this perception of poor

outcomes forms the basis for low status. Second, the presence of diversity initiatives suggest that targeted groups have *received* extra help, enabling them to compete for resources more effectively and making them a bigger threat.

Thus, if diversity scholarships create the illusion that their recipients needed and received extra help at a prior institution, employers may perceive recipients as less competent and be less likely to interview them than their peers. But, as discussed in the previous chapter, the stimuli in these prior experiments used explicit, unrealistic language that portrayed affirmative action as "preferential treatment" and made no mention of merit. A more realistic, meritocratic framing of diversity initiatives might lessen the impression that their targets needed extra help, which would reduce the stigma of incompetence. (However, a more meritocratic framing would not necessarily make the target appear warmer, as the scholarship still enhances the competitiveness of the target.)

This theoretical framework can be extended to other forms of merit scholarships as well. Non-diversity scholarships might not convey the same message of needing extra help, thereby not damaging targets' status and perceived competence. Indeed, people tend to see general, "unrestricted" awards as higher status than diversity-specific awards (Germano and Cheryan 2021). Under this logic, people who receive non-diversity scholarships would not only be seen as more competent than their peers without any scholarship awards, but they would also be perceived as more competent than their peers who receive diversity-related awards.

In sum, this line of thinking suggests that merit scholarships (whether they are diversity-related or not) affect employers' appraisals of job applicants by making the candidates appear higher (or lower) status and more competitive, which in turn influences their perceptions of the applicants' warmth and competence. However, not only are most of the studies in this line of research based on an unrealistic conceptualization of affirmative action as "preferential treatment", they also limited focus

to evaluations of people's *past* performance. It is less clear how a more meritocratic framing of participation in diversity efforts might shape perceptions of competence and warmth, or what role competence and warmth might play in determining evaluations of people's *future* performance in a different organizational context.

3.1.2 Achievement, Selectivity, and Signal Clarity

In a hiring setting, the signal fit and signal clarity of various types of merit scholarships may be as important as perceptions of warmth and competence. Signaling theory is applicable in any setting in which two parties need to reduce information asymmetry (Connelly et al. 2011; Spence 2002), but it has been particularly influential in studies of labor markets. In the original formulation of signaling theory, Spence (1973) proposed that education has value for job candidates because educational credentials can serve a signaling function (in contrast to merely being human capital accumulation). Employers lack information about the quality of job candidates, but candidates who have educational credentials can use them to signal their higher quality. Importantly, signals only have informational value when they are seen as reliable indicators of relevant unobservable attributes. The stronger this correlation, the better the signal fit (Connelly et al. 2011). For example, college degrees have signaling value in the original Spence model because of the presumption that "lower quality candidates would not be able to withstand the rigors of higher education" (Connelly et al. 2011).

However, recent research in the sociology of labor markets suggests that this simple signaling model under-theorizes the heterogeneity in signal clarity, or "the degree to which signals are unambiguous, have a known and verifiable origin, and can be evaluated quickly with minimum errors by signal recipients" (Mai 2021, 6). An unclear signal can decrease a job applicants' prospects because it is cognitively arduous to interpret, which goes against decision-makers' tendency to rely on heuristics

in the early stages of the screening process. Moreover, the ambiguous meaning of unclear signals threatens their actual informational value.

What unobservable qualities are merit scholarships signaling, and are they signaling them clearly? I argue that the signaling value of merit scholarships, like that of other awards and prizes, derives from their selectivity. Merit scholarships, whether they focus on a specific type of criteria or not, signal that their recipients are especially "high-ability" or "high-quality". These people are not only good; they are exceptional. Moreover, prior research has found that selectivity has signaling value in its own right – graduates of selective, elite universities are more likely to get a callback for an interview than job applicants who graduated from less selective institutions (Gaddis 2015).

The clarity of the merit scholarship signal depends on the extent to which decision-makers perceive them to be unambiguously selective. Since these awards are not likely to be known beyond the boundaries of the institutions that awarded them, their selectivity is not likely to be known or verifiable to most employers, which might make them rather unclear signals. Merit scholarships that do not explicitly focus on diversity might still be perceived as unambiguously difficult achievements because the pool of potential awardees appears to be large and unrestricted. On the other hand, scholarships that explicitly refer to diversity may vary more in perceived selectivity because the eligibility seems restricted to a smaller pool of applicants, or based on a more specialized set of criteria. Additionally, as will be elaborated in the next section, the perceived selectivity of the award could be influenced by the attributes of its recipient.

Thus, the signaling perspective suggests that the overall reputational effects of merit scholarships on hiring could operate through the perceived selectivity of the scholarship. If decision-makers question the informational value of the awards, then their overall effect will be influenced by how selective they appear.

3.1.3 The Racialized Effects of Awards

Affirmative action stigma and signaling theory both offer broad predictions about how employers might evaluate diversity merit scholarships (and other types of merit scholarships), but they are agnostic about how these effects may differ by race or other types of sociodemographic characteristics. Yet results from the field experiment showed that the reputational effects of these awards are indeed racialized: having a diversity scholarship increases callback rates for White applicants but not for Black applicants, and non-diversity scholarship raise White applicants' interview prospects more than they do for Black applicants'. What are the limitations of these theories, and how might they be expanded to explain racialized effects?

The affirmative action stigma model is unable to explain racialized effects because it only focuses on beneficiaries who are clearly perceived as women or racial minorities. By conceptualizing affirmative action as a form of preferential treatment for minorities, rather than more diffuse form of diversity management, this framework has a narrow interpretation of who might be associated with organizational diversity initiatives. This theory developed by comparing target groups who were all already lower-status in gender and racial hierarchies – for example, women who were branded as affirmative action hires and women who did not receive these labels. Thus, affirmative action's negative effect on perceived competence can be interpreted as a result of emphasizing lower-status groups' need for extra help – in other words, drawing further attention to their marginalized status. It is unclear if participation in diversity initiatives will convey the same message of "needing extra help" under more diffuse conceptualizations of diversity, or when the minority status of the target is more ambiguous (i.e. not clearly female or non-White). Diversity initiatives can contribute to the further stigmatization of already stigmatized groups, but members of high-status groups who participate in them might be buffered from these effects by their pre-existing high status.

The signaling model is more flexible for accounting for racialized effects. The model already considers that the usefulness of the information signal depends on how well the the signal correlates with the sought-after quality of the signaler in the eyes of the signal recipient (Connelly et al. 2011). This makes clear that biased evaluations due to pre-existing status biases can affect signal quality. For example, if hiring decision-makers already have biased status beliefs about Black applicants, they might think that merit scholarships are highly indicative of being exceptional when White applicants have them, but are less convinced of this association for Black applicants.

Finally, the racialized appraisal of diversity-valuing behavior may be an additional reason why diversity scholarships have stronger positive reputational effects for White applicants than Black applicants. Women and racial minorities who engage in diversity-valuing behaviors tend to be negatively stereotyped and receive lower performance ratings (Hekman et al. 2016), Black employees tend to be seen as more self-interested than White employees are when they promote diversity initiatives (Gardner and Ryan 2020), and even though racial minorities are more likely to be assigned to perform social justice-related tasks at work (e.g. equal-opportunity or corporate social responsibility efforts), the justice-related initiatives they lead are perceived as less effective or necessary (Saguy et al. 2020). Moreover, because racial minorities are often expected to do unpaid "racialized equity labor" (Lerma et al. 2020) to address racism in organizations and to smooth interactions with White peers (Wingfield and Alston 2014), this kind of diversity-related work may be taken for granted rather than recognized as extra labor. This body of research suggest that when putatively White applicants receive diversity-related scholarships, these awards may be construed as well-deserved acknowledgement of positive efforts. However, when putatively Black applicants receive the same awards, those accolades may be seen as signs of self-interest or unnecessary recognition.

3.2 Methods

3.2.1 Experimental Design and Randomization

In order to test these potential mechanisms in the same context as the prior field experiment, the survey experiment replicated the design of the field experiment as closely as possible. Although it is impossible to replicate the external context of the field in a survey setting, I am able to rule out other confounding factors by keeping the experimental materials and randomization scheme across the two studies the same.

Like the field experiment, the survey experiment also used a matched pairs design. The two main independent variables of interest (scholarship status and racial cue) are operationalized, manipulated, and randomized in the same way. This means each respondent viewed and rated two resumes, which always differed in the type of scholarship it presented (diversity scholarship, non-diversity scholarship, or no scholarship at all) but shared the same racial cues (Black or White). Because there are three scholarship conditions and each respondent only reviews two resumes, each respondent is assigned three possible pairs: i) resume with diversity scholarship and resume with non-diversity scholarship; ii) resume with diversity scholarship and resume with no scholarship; or iii) resume with non-diversity scholarship and resume with no scholarship. The order that the two resumes were shown was randomized. As with the field experiment, this design produces within-subject estimates for scholarship conditions and between-subject estimates for racial cue.

Estimates of racial discrimination tend to be smaller in survey experiments than field experiments due to social desirability bias (Pager and Quillian 2005; see Pedulla 2016 for an analogous finding with respect to gender). However, social desirability bias is less likely to be a problem for estimating the effect of various types of merit-based awards. Whereas anti-discrimination laws and social norms discourage employers from even admitting that they consider job applicants' race during the

screening process, merit is considered a legitimate and fair principle for guiding hiring decisions (Scully 2000). In fact, the logic of meritocracy would justify treating people differently if some of them have awards that others don't. Moreover, since questions about diversity and affirmative action often appear on polls and in the news, they are more likely to be regarded as public opinion issues in which both "pro" and "con" perspectives are normatively and legally acceptable. Although most employers face pressure to legally comply with affirmative action efforts, rejecting a job candidate because they seem to have benefited from prior diversity initiatives is unlikely to be seen as a violation of anti-discrimination policy. Thus, within-race scholarship effects may still be observable in a survey experimental setting, even if the effects of the racial cue itself will be muted by social desirability bias.

While the two main variables of interest were operationalized and randomized identically between the survey and field experiments, the counterbalancing of "nuisance" factors (factors that are not of theoretical interest, like resume template, university name, and home address) in the survey experiment diverged from the field experiment in two minor ways. First, while the template and work histories in the field experiment were fully randomized (e.g. work history A could be presented on template A or B, and work history B could be presented on template A or B), I blocked them in the survey experiment so that they would always co-vary (e.g. work history A would always appear on template A, and work history B would always appear on template B). This reduced the number of nuisance factors by one, which was helpful because the sample for the survey experiment is smaller.

Second, while resumes in the field experiment matched addresses, area codes, and names of previous employers to the local labor market, I fixed all resumes in the survey experiment to one city: Washington D.C. Since the respondents in the survey experiment are not restricted to a certain set of locations, it would be near impossible to match every resume's place-specific information to its viewer's geographic location.

I decided to pick one city for all the resumes rather than randomizing the resumes across the eight cities featured in the field experiment to reduce noise and the number of nuisance factors to block on. I choose Washington D.C. as the "representative" city because 1) in the audit study, Washington D.C. callback rates were about the same as the sample-wide callback rates across all eight cities (about 11%), 2) the Washington D.C. callback rate falls in the middle among all city-specific callback rates, and 3) Washington D.C. was one of the biggest labor markets in the field study sample (only NYC was greater; D.C. and LA were about the same).

I worked with Qualtrics, a survey research company, to field the survey on a sample of U.S.-based adults with experience making hiring decisions. Since there are no national sampling frames of individuals who make hiring decisions or national estimates on the demographic characteristics of such a population, probability-based sampling and quota sampling were not possible. To best approximate the population of interest with convenience sampling, Qualtrics recruited participants from a large opt-in panel of business professionals whose identities were verified by the panel provider. Two screening questions determined eligibility: 1) in the past two years, respondents had to have made decisions about whether or not to hire job applicants as part of their job, and 2) respondents' current occupation had to be business owner, business executive, mid-level manager, human resources manager, human resources associate or assistant, or corporate recruiter[2]. Between June and July 2021, 751 respondents completed the survey. Descriptive statistics of the sample are presented in Table 3.1.

[2]The second screening question asked people to choose a category that best described their current occupation, not their current employment. Therefore, people who are not employed at the time of the survey but have recently made hiring decisions as part of their usual work were eligible for the study. About 4.8% of the sample were people who were not working at the time of the survey. Only 2.4% were actively looking for work (the rest were temporarily on leave from a current job, retired, or not looking for work).

Table 3.1: Descriptive Statistics of Survey Sample (N = 751)

	Percentage
Race	
White	64.4
Black/African American	14.6
Hispanic/Latino(a)	11.6
Asian/Asian American	5.2
Gender	
Man	51.8
Woman	47.8
Age	
Mean (Range)	37.6 (18-73)
Education	
Less than Bachelor's Degree	37.3
Bachelor's Degree	37.4
Master's Degree or Higher	25.3
Employment Status	
Paid Employee	76.4
Self-Employed	18.8
Currently Between Jobs	4.8
Firm Size	
Fewer than 15 employees	18.4
Between 16-99 employees	19.2
Between 100-499 employees	25.4
Between 500 and 1,500 employees	21.2
Greater than 1,500 employees	15.8
Industry	
Agriculture, Mining, Construction, or Manufacturing	20.7
Education and Health Services	14.7
Finance, Real Estate, or Information	17.8
Leisure, Hospitality, Transportation, or Utilities	6.4
Professional or Other Services and Public Administration	30.5
Wholesale or Retail Trade	10.0
Region	
Northeast	17.4
Midwest	17.0
South	45.1
West	20.4

Note: "Other" category for race and gender are omitted.

Compared the broader U.S. population, the respondents in the sample are more likely to have an advanced degree, more likely to be self-employed[3], and slightly more likely to self-identify as a man. Black people are slightly overrepresented in the sample, while White and Hispanic people are slightly underrepresented. Respondents are also more likely to work in professional, business, and scientific services, public administration, or other services, more likely to work in firms based in the South, and more likely to work in large firms with more than 500 employees.

3.2.2 Survey and Variable Construction

Participants were told that the study involves reviewing resumes of recent college graduates who conducted job searches for full-time entry-level positions in the past year, and they will review two resumes randomly selected from a larger pile. Participants who provided informed consent and passed the two screening questions and a bot check were then presented with the first resume. On the same page as the resume, they were asked on a seven-point scale, "How likely are you to recommend that your organization interview [Name] for an entry-level position?". Following other survey experiments of hiring (Pedulla 2016; Quadlin 2018), I recoded this item into a dichotomized variable with the positive endpoint of the scale ("Extremely Likely") equaling "1" and all other categories equaling "0". Dichotomizing the interview likelihood variable makes the item more comparable to a callback in a field-experimental context because in real hiring scenarios, only applicants who receive the highest rating would likely advance to the next stage (Pedulla 2016).

After answering the interview likelihood question, respondents rated the applicants' on competence, warmth, and social skills. Perceptions of competence were measured with a scale averaged from three items. Respondents were asked on a seven-point scale how strongly they agreed or disagreed with the following three

[3]This could be because "business owner" was one of the occupations eligible for the study.

statements: "He is competent", "He is skillful", and "He is capable" (statements taken from Fiske et al. 2002). The endpoints of the scale were "Strongly Disagree" and "Strongly Agree", with "Neither Agree nor Disagree" as the midpoint. The three combined items had a Cronbach's Alpha of 0.87, indicating high internal consistency. Perceptions of warmth were also measured with a scale averaged from three items. Respondents answered the same agree-disagree question, but with the statements "He is friendly", "He is good-natured", and "He is agreeable" (statements taken from Fiske et al. 2002). This scale also had high internal consistency, with a Cronbach's Alpha of 0.89. Finally, to measure perceptions of social skills, respondents were asked to rate on a five-point scale how accurately they thought the following statements represented the applicant: "He would interact well with customers and clients", "He would fit in well at the company", and "He might make coworkers feel uncomfortable" (reverse-coded). The endpoints of the scale were "Extremely Accurate" and "Not Accurate At All". However, these three items have a Cronbach's Alpha of 0.49, suggesting that they are not highly internally consistent, and they will be analyzed separately in future analyses.

After respondents review the first resume and provide their interview recommendation and ratings of competence, warmth, and social skills, they are shown the second resume and asked the same set of questions. Once respondents have reviewed and answered this initial set of questions for both resumes, they were shown the resumes again (in the same order) and asked an additional set of questions. In this second set of questions, respondents were asked to rate the specific scholarship(s) on the resume(s) they saw in terms of selectivity (if one of the resumes the respondent reviewed did not have a scholarship on it, this question was skipped for that resume). They were asked, "How difficult do you think it is to get the [Scholarship Name]?". The response option was five-point scale with "Extremely difficult" and "Not difficult at all" as the endpoints.

Then, regardless of which scholarship condition they had been randomly assigned to, all respondents were asked to speculate about the sociodemographic characteristics of the applicant based on their first impressions. They were asked to choose a category that they thought most accurately described the applicant's socioeconomic background ("Working-class background", "Middle-class background", or "Upper-class background"), primary racial identity ("White","Black or African American","Hispanic or Latino(a)", "Asian or Asian American",or "Other"; "Hispanic" and "Asian", were collapsed into the "Other" category), and sexual orientation ("Heterosexual or Straight", "Bisexual", or "Gay and Lesbian"; "Bisexual" and "Gay and Lesbian" were combined into one "LGB" category). I asked this set of questions about perceived demographics after respondents made hiring recommendations for both candidates to minimize social desirability bias in the first set of questions.

Finally, respondents provided information about their own demographic characteristics and the size, industry, and geographic location of their company. For ease of interpretation and to minimize the number of sparse categories, I collapsed respondents' self-reported race to three categories ("White", "Black", and "Other") and their educational attainment into two ("Less than BA" and "BA or higher").

3.3 Results

3.3.1 Recommendation for Interview

First, I examine how results from the survey experiment align with findings from the field experiment. How do different types of merit scholarships affect hiring decision-makers' evaluations of Black and White applicants? Figure 3.1 shows the probability that an employer would be "extremely likely" to recommend an applicant for an interview at their company or organization.

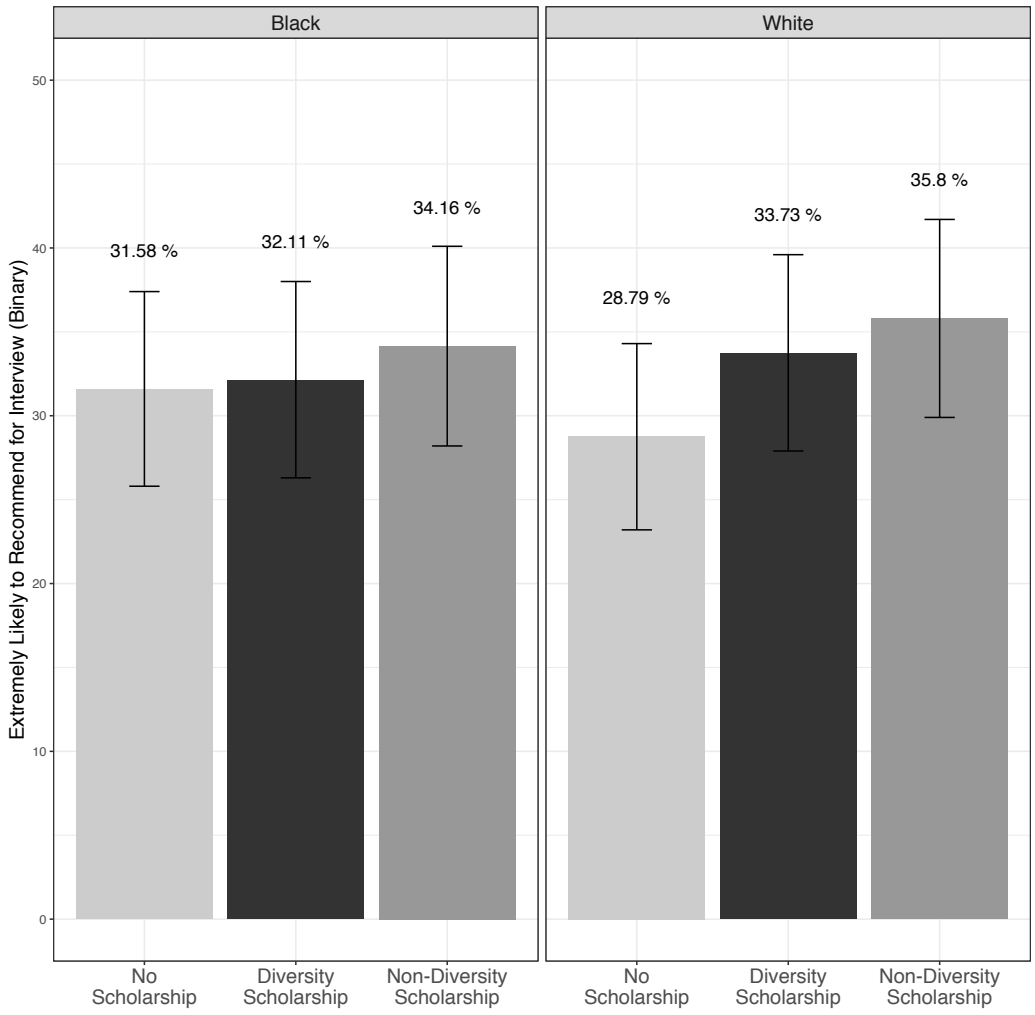


Figure 3.1: Probability of being "Extremely Likely" to be recommended for an interview by experimental condition

In terms of within-race scholarship effects and how they differ for Black and White applicants, the results from the survey experiment broadly echo the results from the field experiment in two ways. First, both Black and White applicants are more likely to be considered for an interview when they have a non-diversity scholarship, but the positive signal of that type of award appears to be larger for White applicants than it is for Black applicants. Second, the diversity scholarship increases White applicants' chances of moving on to the next stage of the hiring process more than it does for Black applicants. In both the field and survey experiment, White applicants with diversity scholarships are treated more similarly to White applicants with non-diversity scholarships than White applicants without scholarships, while Black applicants with diversity scholarships are treated more similarly to Black applicants without scholarships than Black applicants with other types of merit scholarships. However, while these results descriptively echo the field experimental findings in these two respects, the effects are not statistically significant in t-tests.

This lack of statistical significance could be related to the fact that the scholarship effects are smaller in the survey setting than they were in the field experiment. This is not because the absolute percentage point differences across scholarship conditions are smaller – in fact, the percentage point differences in the survey experiment are very similar to the percentage point differences observed in the field experiment – but because the base rate for interview recommendations is higher in the survey setting. Whereas the field experiment had an overall callback rate of 11%, 32.7% of respondents in the survey experiment sample said that they were "extremely likely" to recommend an applicant for an interview. This difference in base interview likelihood might be due to the different restrictions and incentive structures in field and lab settings – for example, in the "real world", employers often review more than two resumes for an open position and do not get to call back every candidate they like for an interview, but for a hypothetical situation presented in a survey, it is fairly

costless for them to say that they are "extremely likely" to invite an applicant for an interview.

Additionally, the reputational effect of the diversity scholarship for Black applicants also differs in field and survey settings: in the field experiment, Black diversity scholarship winners' probability of receiving a callback was about 0.5 percentage points lower than that of Black applicants without scholarships, but in the survey experiment, Black diversity scholarship winner's probability of being recommended for an interview is about 0.5 percentage points higher than Black applicants without scholarships. Given that the effect is small and statistically insignificant in both studies, this discrepancy may just be due to statistical noise.

Finally, as expected, the effect of applicants' racial cues do not replicate in a survey experimental setting. The field experiment found evidence of discrimination against Black applicants – on average, callback rates were noticeably higher for White applicants, and White applicants with both diversity and non-diversity scholarships were significantly more likely to receive a callback than their Black counterparts. In the survey experiment, Black and White applicants' chances of being "extremely likely" to be recommended for an interview were extremely similar (about 32.6% for Black applicants and 32.8% for White applicants), and Black applicants without scholarships are actually more likely to receive this kind of endorsement from hiring decision-makers than White applicants without scholarships. This survey-field discrepancy is consistent with prior research racial discrimination and social desirability bias in employer surveys (Pager and Quillian 2005).

To account for other sources of variability, I also estimate regression models that control for order effects (if the resume being rated was presented first or second) and respondent characteristics (respondents' race, gender, age, education, employment status, the number of people their company employs, and the region their company is located in). Since each respondent reviewed two resumes, I also

clustered standard errors by respondent ID. Tables 3.2 and 3.3 display these adjusted estimates of the effects of diversity and non-diversity scholarships on interview recommendation for Black applicants and White applicants, respectively. In both tables, Model 1 shows the simple bivariate relationship between the scholarship treatment and decision-makers' recommendations with respondent-adjusted clustered standard errors, while Model 2 also includes controls for order and respondent characteristics.

When order and respondent characteristics are accounted for, the effect of the non-diversity scholarship is statistically significant for White applicants and increases the chances of receiving the highest recommendation for an interview by about 47%. The effect of the non-diversity scholarship is also positive for Black applicants, but not statistically significant. Similarly, the effect of the diversity scholarship is positive for both Black and White applicants, but not statistically significant. Models accounting for experimental order show that on average, resumes viewed second are more likely to receive the highest recommendation for interview than resumes viewed first, but causal mediation analysis suggest that that order effects do not *explain* the effect of the non-diversity scholarship on interview recommendations (see Supplemental Materials).

Thus, the overall pattern of results from the survey experiment raise similar questions about mechanisms as the field experiment – why are diversity merit scholarships devalued relative to non-diversity merit scholarships, especially among Black applicants?

Table 3.2: Log Odds of being "Extremely Likely" to be Recommended for Interview (Black Applicants)

	Base Model (1)	With Controls (2)
Scholarship Status (ref. No Scholarship)		
Diversity Scholarship	0.025	0.060
	(0.174)	(0.181)
Non-Diversity Scholarship	0.117	0.149
	(0.183)	(0.187)
Order in Experiment (ref. Shown First)		
Shown Second		0.339*
		(0.133)
Race (ref. White)		
Black		0.431
		(0.272)
Other		−0.610*
		(0.266)
Gender (ref. Man)		
Woman		0.101
		(0.195)
Other		−13.005***
		(0.874)
Age		−0.004
		(0.010)
Education (ref. Has BA or Higher)		
Does not have Bachelor's Degree		−0.046
		(0.228)
Employment Status (ref. Paid Employee)		
Self-Employed		−0.018
		(0.326)
Between Jobs		−0.713
		(0.453)

(Continued on Next Page...)

Table 3.2: "Extremely Likely"to be Recommended for Interview (Black Applicants) *(continued)*

	Base Model (1)	With Controls (2)
Firm Size (ref. Fewer than 15 employees)		
Between 16 and 99 employees		0.603
		(0.355)
Between 100 and 499 employees		0.118
		(0.382)
Between 500 and 1,500 employees		0.156
		(0.415)
Greater than 1,500 employees		0.127
		(0.411)
Region (ref. South)		
Northeast		-0.583^{*}
		(0.273)
Midwest		-0.022
		(0.255)
West		-0.400
		(0.272)
Constant	-0.773^{***}	-0.795
	(0.137)	(0.614)
Observations	736	736
Log Likelihood	-464.485	-443.220
Akaike Inf. Crit.	934.971	924.439
Note:	*p<0.05; **p<0.01; ***p<0.001	

Table 3.3: Log Odds of being "Extremely Likely" to be Recommended for Interview (White Applicants)

	Base Model (1)	With Controls (2)
Scholarship Status (ref. No Scholarship)		
Diversity Scholarship	0.230	0.294
	(0.172)	(0.177)
Non-Diversity Scholarship	0.321	0.388*
	(0.175)	(0.181)
Order in Experiment (ref. Shown First)		
Shown Second		0.491***
		(0.128)
Race (ref. White)		
Black		0.178
		(0.280)
Other		0.204
		(0.222)
Gender (ref. Man)		
Woman		−0.044
		(0.196)
Other		14.875***
		(1.071)
Age		−0.012
		(0.009)
Education (ref. Has BA or Higher)		
Does not have Bachelor's Degree		0.133
		(0.204)
Employment Status (ref. Paid Employee)		
Self-Employed		−0.514*
		(0.256)
Between Jobs		−0.102
		(0.442)

(Continued on Next Page...)

Table 3.3: 'Extremely Likely" to be Recommended for Interview (White Applicants) *(continued)*

	Base Model (1)	With Controls (2)
Firm Size (ref. Fewer than 15 employees)		
Between 16 and 99 employees		−0.047
		(0.322)
Between 100 and 499 employees		0.065
		(0.313)
Between 500 and 1,500 employees		−0.019
		(0.339)
Greater than 1,500 employees		−0.381
		(0.388)
Region (ref. South)		
Northeast		0.064
		(0.288)
Midwest		−0.501
		(0.279)
West		−0.314
		(0.239)
Constant	−0.905***	−0.602
	(0.138)	(0.487)
Observations	766	766
Log Likelihood	−482.989	−466.362
Akaike Inf. Crit.	971.977	970.724
Note:	*p<0.05; **p<0.01; ***p<0.001	

3.3.2 Revisiting the Affirmative Action Stigma of Incompetence and Low Warmth

Do the reputational effects that merit scholarships have on interview likelihood operate through perceptions of competence and warmth? This next set of analyses uses causal mediation analysis (Imai et al. 2010) to examine if and how perceived competence and warmth explains the overall effect of diversity and non-diversity scholarships on respondents' interview recommendations[4]

First, I descriptively examine the causal effect of having a diversity scholarship on perceived competence and warmth, compared to other types of scholarships or not having a scholarship at all. Figure 3.2 shows that the average effects of merit scholarships on ratings of competence and warmth are small and not statistically significant for both Black and White applicants. Contrary to expectations from existing theories of affirmative action stigma, I do not find evidence that applicants with diversity scholarships are seen as less competent and less warm. Since perceptions of warmth and competence do not vary significant across scholarship conditions, perceived competence and warmth cannot causally explain any of the reputational effects of diversity and non-diversity merit scholarships.

[4]I conduct mediation analyses for both Black and White applicants despite the lack of a statistically significant direct effect for Black applicants, because significant indirect effects can still exist in the absence of a total or direct effect (Bollen 1989; Hayes 2017; Rucker et al. 2011)

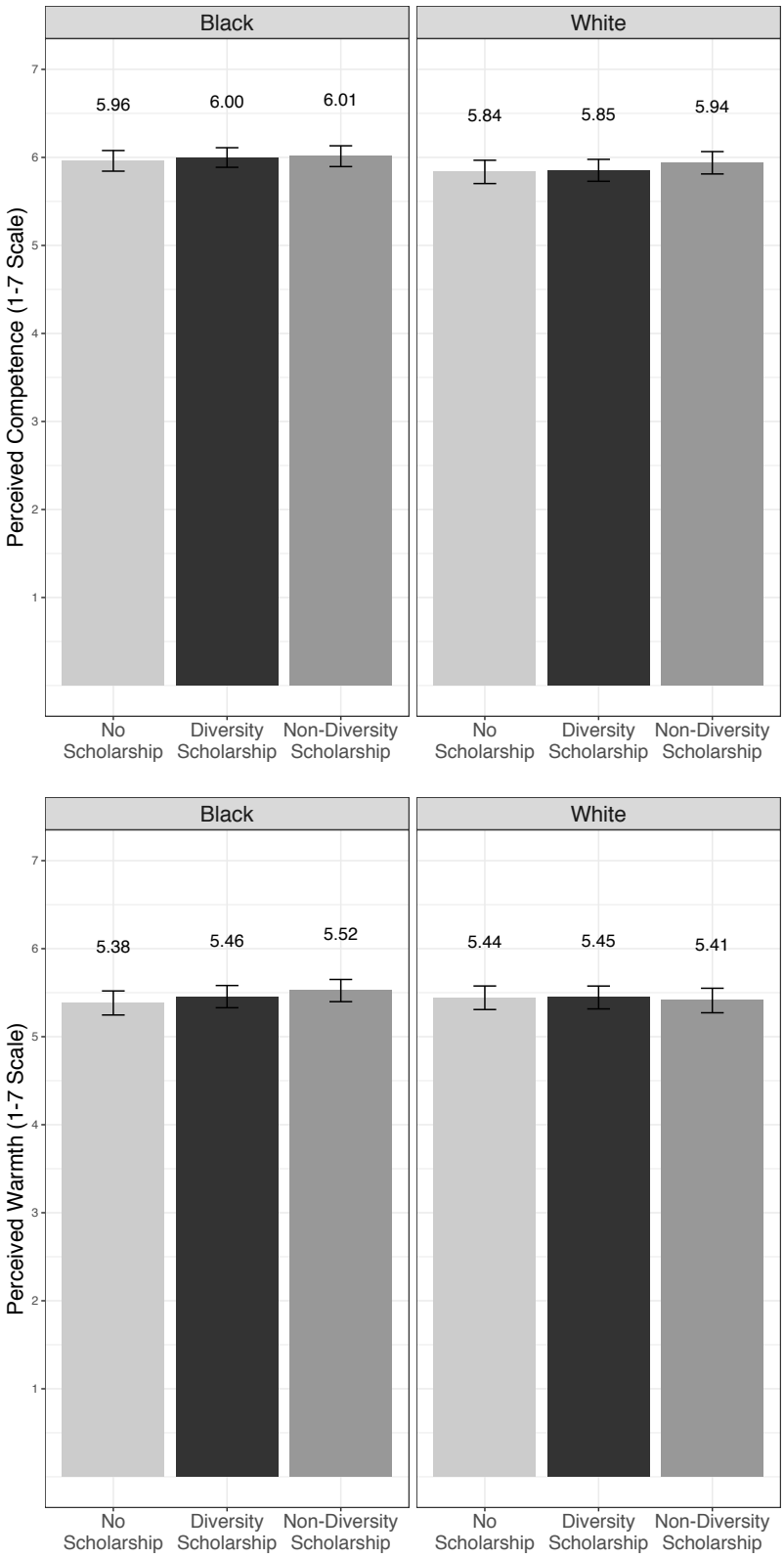


Figure 3.2: Perceptions of applicant competence (*top*) and warmth (*bottom*) by experimental condition

3.3.3 Perceived Selectivity and Signal Uncertainty

In the next set of analyses, I examine the mediating role of awards' perceived selectivity. Diversity merit scholarships might be devalued relative to non-diversity merit scholarships because people perceive their eligibility to be more restricted, so they appear easier to receive, which makes the informational value of their signal less useful.

Indeed, Figure 3.3 shows that non-diversity scholarships are seen as more difficult to receive than diversity scholarships are. The effects are statistically significant and almost identical for Black and White applicants ($p = 0.03$ in Student's t-test for both). The inter-quartile range of the rating is one point larger for the diversity scholarship than it is for the non-diversity scholarship, suggesting that people have slightly more varied impressions of the perceived selectivity of diversity scholarships compared to scholarships that don't mention diversity.

Next, I examine if perceived selectivity might explain why diversity and non-diversity scholarships have different reputational effects. The estimated average causal mediation effect (AMCE; also known as the indirect effect) of perceived selectivity, the average direct effect (ADE) of diversity and non-diversity scholarships, and the total effect are shown in Figure 3.4 with 95% confidence intervals. The blue estimates show the indirect and direct effects of a diversity scholarship on the probability of receiving a recommendation for interview, and the red estimates show the indirect and direct effects of a non-diversity scholarship on the probability of receiving a recommendation for interview.

Panel A in Figure 3.4 shows that among Black applicants, the indirect effect of perceived selectivity positively and significantly mediates the direct effect of the scholarship ($p < 0.05$), even though the direct effect is very small and not statistically significant. For Black applicants, having a non-diversity scholarship (as opposed to a diversity scholarship) increases the probability of getting a recommendation for an

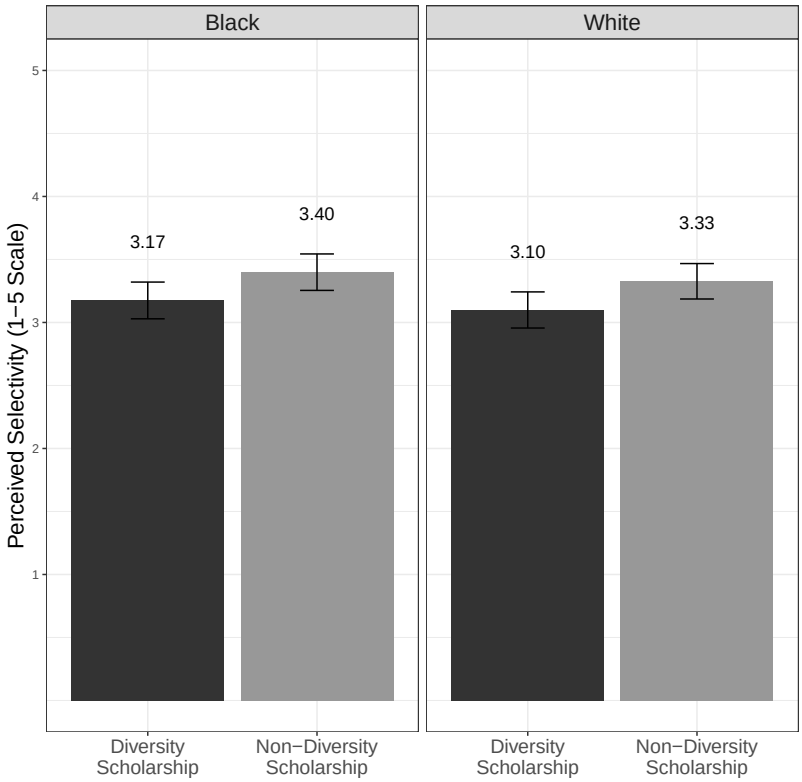


Figure 3.3: Perceptions of scholarship selectivity by experimental condition

interview by 0.018 points (or 1.8% percentage points) in total. 0.013 of those points can be attributed to the perceived selectivity of the scholarship. By contrast, the indirect effect of perceived selectivity does not significantly mediate the direct effect of the scholarship for White applicants ($p = 0.26$).

When the "no interaction" assumption is relaxed and the mediator is allowed to vary as a function of the treatment (Panel B in Figure 3.4), the mediating effect of perceived selectivity for Black applicants is only statistically significant for the diversity scholarship treatment. That is, for Black applicants, the overall signaling effect of a diversity scholarship on decision-makers' interview recommendation is explained by how selective the award seems. For White applicants and Black applicants with a non-diversity award, how selective the award seems does not explain the overall reputational boost conferred by a merit scholarship. This suggests that for Black applicants – but not White applicants – the potential reputational gains

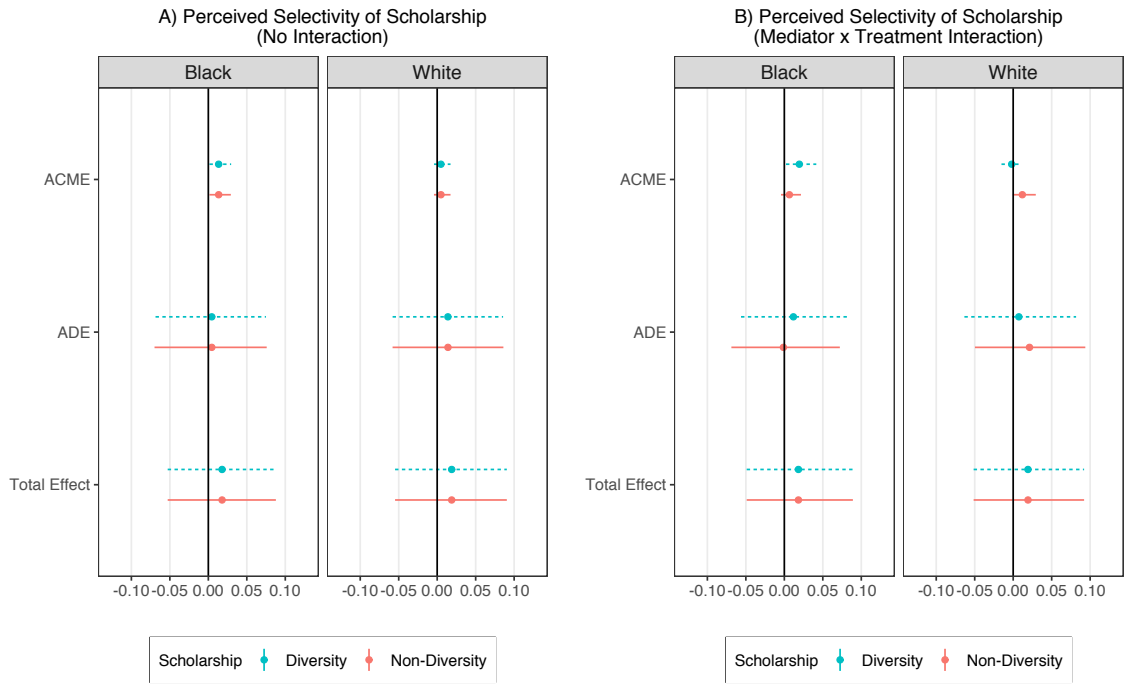


Figure 3.4: Average causal mediation effect of perceived scholarship selectivity on interview recommendation, without and with treatment interaction

of a diversity scholarship is conditional on how selective decision-makers perceive the award to be. The positive signal of merit scholarships that do not specifically mention diversity are not contingent on perceived selectivity for applicants of any race.

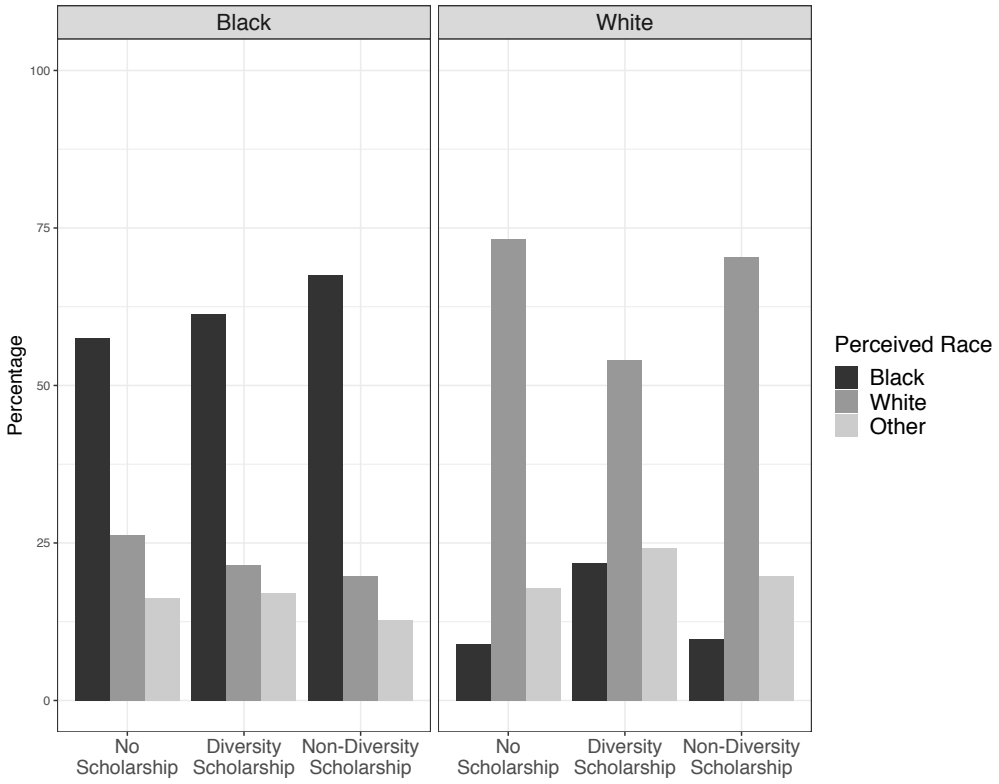
3.3.4 Alternative Explanation: Are Diversity Scholarships Altering Sociodemographic Cues?

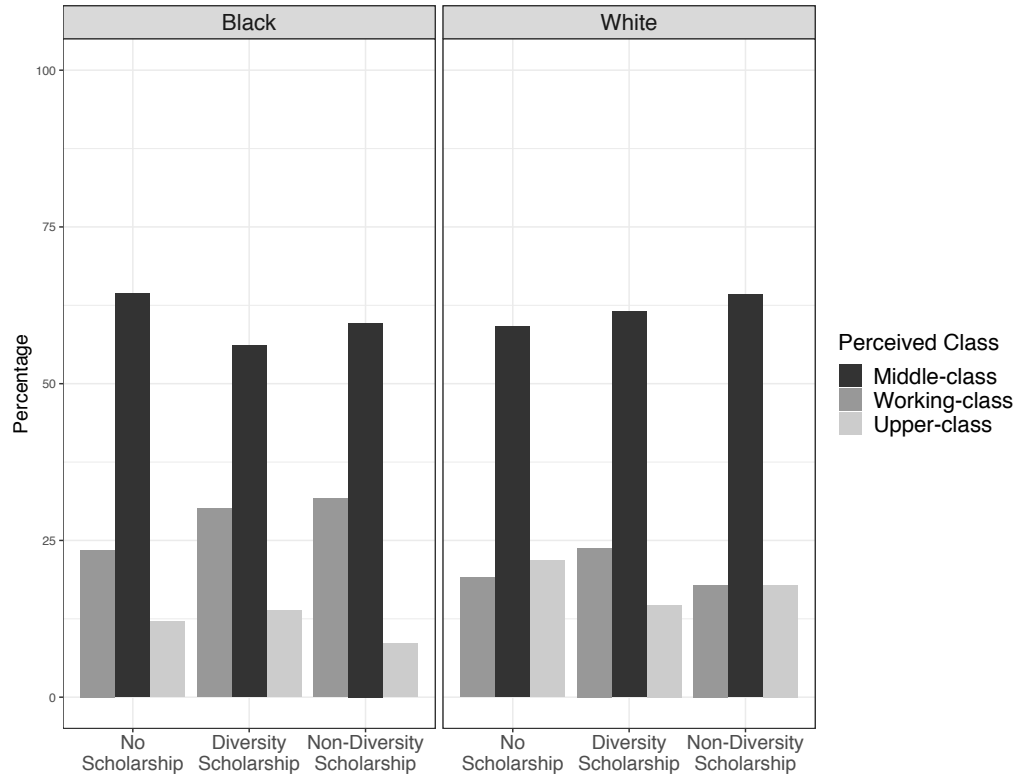
An alternative explanation is that diversity and non-diversity scholarships might alter how applicants are perceived in terms of race, socioeconomic class, and sexual orientation. For example, an employer who see an applicant with a diversity scholarship might wonder why that applicant received such an award and how they contributed to diversity on campus (e.g. by participating in activist organizations, being a racial or sexual minority, or being a first-generation college student). Employers may also have preconceived notions about the type of people who receive merit scholarships

and the type of people who are recognized for contributions to diversity. Although I pre-tested the applicant names to ensure that they would send the intended racial, class, and gender cues, those cues might change when presented in conjunction with different kinds of scholarships. Moreover, if diversity scholarships influence the odds that their recipients will be perceived as members of marginalized groups, subjecting them to additional discrimination, those perceived sociodemographic cues could be what's driving the overall signaling effect of the diversity and non-diversity scholarship.

Figure 3.4 shows respondents' perceptions of applicant race, socioeconomic class, and sexual orientation, respectively. Perceptions of applicants' race are less congruent with intended racial cues than they were in the pre-test, but they are still congruent more than 50% of the time in all experimental conditions. For applicants with names that were designed to signal Blackness, the diversity and non-diversity scholarships increases congruence slightly by raising the probability that they would be perceived as Black, but the difference is not statistically significant (Pearson's Chi-squared test; $p = 0.2$). However, for applicants with names that were designed to signal Whiteness, the diversity scholarship significantly decreases congruence by lowering the probability that they would be perceived as White (Pearson's Chi-squared test; $p < 0.001$).

Most applicants are perceived as middle class. Both diversity and non-diversity scholarships increase Black applicants' chances of being perceived as working-class, while only the diversity scholarship increases White applicants' chances of being perceived as working-class. However, these changes are small and not statistically significant (Pearson's Chi-squared test; $p = 0.1$ for Black applicants and $p = 0.2$ for White applicants).


Black
White
100
75
50
25
0
Percentage
No Scholarship
Diversity Scholarship
Non-Diversity Scholarship
Perceived Race
Black
White
Other


Black
White
100
75
50
25
0
Percentage
No Scholarship
Diversity Scholarship
Non-Diversity Scholarship
Perceived Class
Middle-class
Working-class
Upper-class

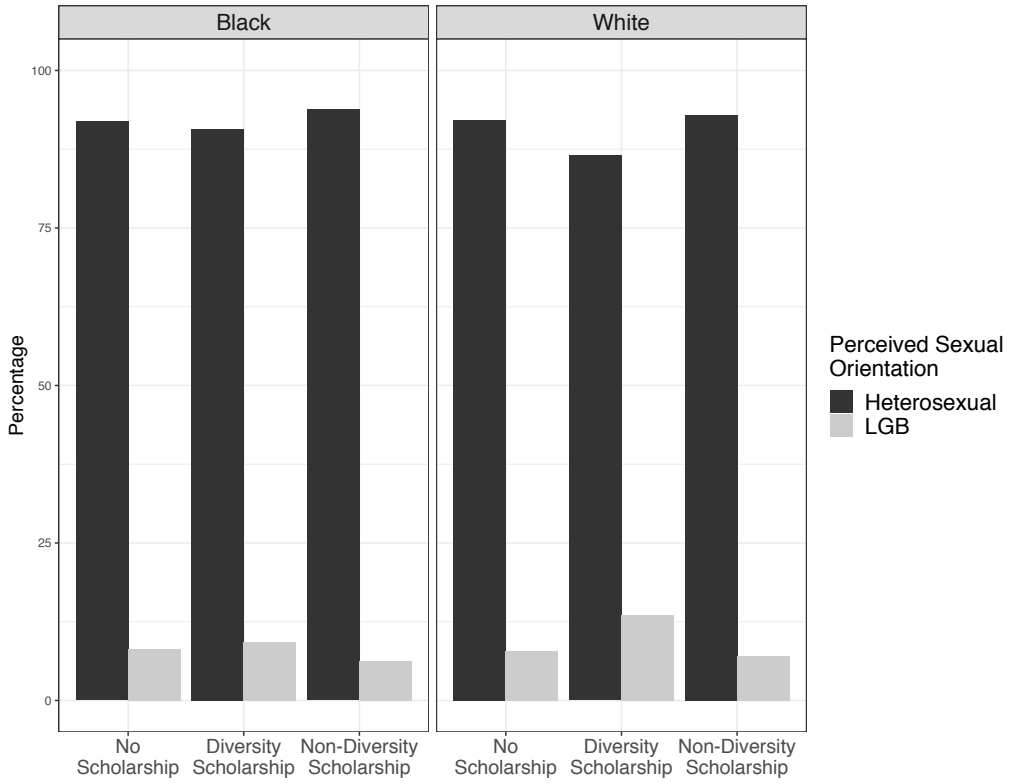


Figure 3.4: Perceived race, socioeconomic class, and sexual orientation by experimental condition

Finally, the diversity scholarship decreases the probability of being perceived as heterosexual for both Black and White applicants. However, the effect is larger and statistically significant for White applicants (Pearson's Chi-squared test; $p < 0.05$); it is not statistically significant for Black applicants (Pearson's Chi-squared test; $p = 0.4$).

The descriptive results in Figure 3.4 suggest that scholarships may have signaled information about applicants' racial identity and sexual orientation. I now turn to causal mediation analysis to examine if these changes in perceived demographic cues *explain* the awards' overall effect on hiring decision-makers' interview recommendations[5]. Figure 3.5 shows that the estimated ACMEs of perceived race and sexual orientation are all small in magnitude and include zero in their 95% confidence

[5]Since the merit scholarships do not affect perceived class, we can rule that out as a mediator.

intervals. The estimates are also similar when the "no-interaction" assumption is relaxed and mediation effects are allowed to vary as a function of treatment status (see Supplemental Materials). This suggests that the overall effect of different types of scholarships on interview recommendations are not due to respondent perceptions of applicants' sociodemographic traits. In other words, diversity and non-diversity merit scholarships may increase or decrease the congruence of applicants' racial cues or sway perceptions of their sexual orientation, but these signals do not explain why hiring decision-makers seem to prefer applicants with merit scholarships, or why these effects are larger for White applicants.

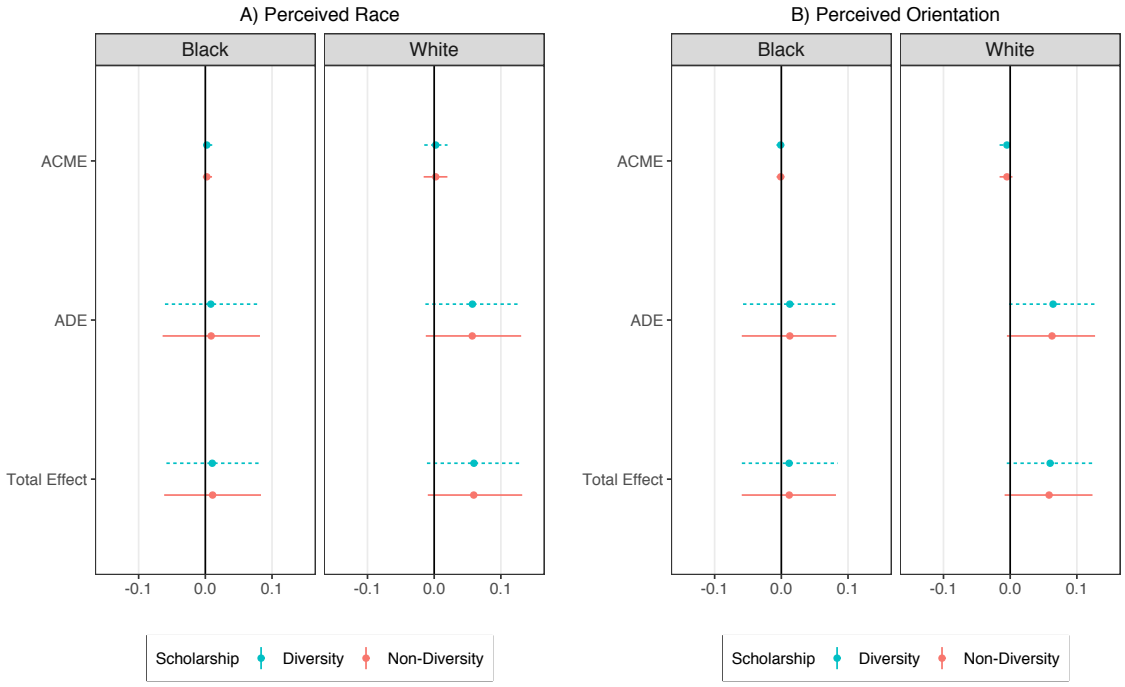


Figure 3.5: Average causal mediation effect of perceived race (*left*) and sexual orientation (*right*) on interview recommendation

3.4 Discussion

In national debates about affirmative action in the 1990s, concerns about stigmatization emerged as a major theme. Scholars worried that minorities who received "preferential treatment" on the basis of their race or gender would be seen as less qualified and competent. Policy recommendations consistently urged organizational leaders to emphasize that affirmative action beneficiaries are selected on the basis of qualifications first (see Reskin 1998 for example). Yet as modern affirmative action efforts – now depicted under the broader heading of "diversity management" – increasingly embraced this meritocratic framing, our understanding of how diversity program recipients are seen by others has not kept pace. Moreover, as organizations adopt more diffuse definitions of "contributions to diversity", diversity initiatives are no longer restricted to racial and gender minorities, raising new questions about how engagement with diversity is perceived.

In this chapter, I conducted a survey experiment on a sample of U.S.-based adults who have professional experience making hiring decisions to investigate how targets of modern diversity recruitment strategies, like merit scholarships, are perceived by employers. The results broadly echo findings from a similar field experiment: job applicants who received diversity-specific awards are not penalized relative to their same-race peers when the awards are portrayed as merit-based. However, these diversity scholarships were devalued relative to other types of merit scholarships when they were given to Black applicants, but not devalued when they were given to White applicants. Having a merit scholarship unrelated to diversity also increased interview likelihood more for White applicants than it did for Black applicants.

Older studies have found that perceptions of competence and warmth drive the stigmatizing effect of affirmative action. However, the results from descriptive and causal mediation analyses do not find support for this theoretical model. Having a diversity scholarship did not make applicants of either race appear less competent or

less warm. Both forms of merit scholarships had negligible impact on perceptions of competence and warmth for Black and White applicants, so perceived competence and warmth did not mediate the direct effect of the scholarships on interview likelihood.

The fact that there are no differences in perceived competence and warmth between diversity scholarship recipients and other applicants is perhaps not surprising in the absence of a negative direct effect. That is, given that this study has found no evidence of backlash towards diversity merit scholarship holders relative to their same-race peers who do not have any merit scholarships, it would make sense that these diversity scholarship recipients are seen as no less competent than their scholarship-less peers.

More puzzling is the fact that neither type of merit scholarship alter perceptions of competence or warmth, but seem to have at least a small direct effect on interview likelihood (though this is less pronounced in the survey context than it was in the field, and the direct effect in the survey experiment is only statistically significant for the non-diversity scholarship received by the White applicant). One possibility is that university awards do not necessarily offer prospective employers any additional informational value about competence or warmth - the similar college degrees, GPAs, and employment histories on the resumes all establish a baseline level of competence and warmth already.

Descriptive and causal mediation analyses on diversity and non-diversity scholarship's perceived selectivity may offer some insight about what these scholarships are actually signaling. Scholarships that do not specifically mention diversity are seen as more difficult to get than diversity-specific awards, and this difference in perceived selectivity is about the same regardless of the racial cue of the applicant. For Black applicants – but not for White applicants – perceived selectivity also causally mediates the effect of having a scholarship on interview likelihood. The effect of having a merit scholarship on interview likelihood is larger when the award

seems more difficult to get. Moreover, when indirect effects are allowed to vary with the scholarship treatment assignment, results suggest that perceived selectivity is a only causal mechanism in the case of the diversity scholarship. Taken together, these results suggest that the reputational effects of diversity scholarships for Black applicants are uniquely sensitive to how selective the award seems. This may explain why diversity scholarships and non-diversity scholarships have more similar effects for White applicants than Black applicants.

One alternative explanation for why diversity and non-diversity merit scholarships affect employers' evaluations is that they alter perceptions of the applicant's demographic characteristics rather than signaling anything about their non-demographic attributes. For example, if employers know that selective universities tend to use merit scholarships to recruit wealthier students, then applicants with merit scholarships might have higher chances of receiving job interviews because employers prefer applicants from higher socioeconomic backgrounds. While there is evidence that diversity scholarships do influence perceptions of race and sexual orientation for applicants with White names, causal mediation analyses suggest that these changes do not explain the scholarships' main effects.

3.4.1 Limitations and Questions for Future Research

There remain many unanswered questions about why different types of scholarships have different types of effects on getting a job for Black and White applicants. While this study was able to rule out some causal mechanisms (perceived competence and warmth; perceived race, socioeconomic background, and sexual orientation) and found suggestive evidence of others (perceived selectivity), it was not able to identify an explanatory mechanism that could explain the overall effect of merit scholarships for both Black and White applicants. More research – especially qualitative research – is needed to further unpack the informational value of of these signals. Since Merton

(1968), there has been a rich sociological tradition of examining how awards and status beget more awards and status (see Bol et al. (2018) for a recent example). Yet the precise signaling mechanisms may differ across cases, and have not been fully elucidated.

A second major question remaining for future research is why the positive reputational effects of the *non*-diversity scholarship is larger for White applicants than Black applicants in both the field and survey experiment. One potential explanation may lie in role congruity theory (Eagly and Karau 2002; Eagly and Diekman 2005), which proposes that groups are more positively evaluated when their characteristics seem to align with their typical social roles. If White men typically occupy the role of "award-winner", then White award-winners would be more role congruent than Black award-winners. (However, Black applicants are likely more role congruent with "diversity award winner", which then raises the puzzle of why White men with diversity scholarships are still more highly evaluated than their Black counterparts.) A more generalized approach to answering this question would also consider social groups beyond Black and White male college graduates. Further investigations of the racialized differential returns to awards and scholarships may also contribute to a better understanding of the mechanisms behind racial discrimination more generally. A major weakness of this study (and other studies of racial discrimination in a survey setting) is that it was unable to unearth and directly test between-race differences.

Additionally, while the materials and research design in this survey experiment hewed closely to that of the field experiment in the previous chapter, the findings may still not be strictly generalizable across the two. Not only does the survey and field settings present varying incentive structures and evaluation contexts, but the samples differ. The respondents in the survey experiment are drawn from a national opt-in sample, and their hiring experience covers a broader set of job types than the positions targeted in the field experiment. These extra sources of variation may have

introduced more noise, which would also explain the more muted effects in the survey setting.

Finally, the survey experiment presented here and the field experiment presented in the previous chapter both share some common theoretical scope conditions. First, the scholarship awards featured on the experimentally manipulated resumes were all smaller, institution-specific awards. As discussed in an earlier section, these awards might be more unclear signals of exceptionalism and selectivity than better-known awards. An employer might not know how prestigious a merit scholarship from a particular university is, but they are likely to know how prestigious a Rhodes scholarship is. However, the tradeoff to studying larger, better known awards is that they tend to be extremely rare, and thus in some sense less generalizable. Second, as discussed in the previous chapter, this study is only able to examine the reputational effects of scholarships. But scholarships, and other types of diversity programs, also have material benefits, such as money and mentorship opportunities. While my findings indicate that the symbolic dimension of diversity merit awards do not help job applicants racialized as Black during the hiring process, it does not mean that Black students and other minorities might not benefit materially from these types of awards at all.

3.4.2 Policy Implications

The policy recommendations from the 1990s seems to have been effective – when diversity initiatives emphasize merit and achievement, their beneficiaries are less likely to be burdened with a stigma of incompetence. One takeaway from this study is that diversity scholarships and pipeline programs should continue to emphasize the achievements and qualifications of their participants. Moreover, because my results indicate that perceived selectivity plays an important role, Black students and em-

ployees may benefit more when the selectivity of the program is made known and verifiable.

However, these older policy debates remain rooted in the idea that affirmative action programs are solely about opening up previously exclusionary organizations to minorities. Yet today, many organizations face a different issue with diversifying their membership: they are open, but have trouble attracting the talent they want. As the case of diversity merit scholarships shows, diversity programming also include targeted outreach and recruitment efforts. However, the results from this study show that minority candidates recruited through diversity-specific pathways might have more limited future labor market opportunities than candidates recruited through other means. This is not to suggest that organizations should stop doing targeted diversity outreach, which is one of the more effective diversity policies (Kalev et al. 2006), but to suggest better integration between different kinds of recruitment strategies. Recent research shows that the introduction of diversity awards may inadvertently siphon members of marginalized groups out of application pools for more lucrative, unrestricted awards (Germano et al. 2021). By recognizing people from historically excluded groups with diversity-specific awards or recruiting them to specific diversity pipeline programs rather than more generalized reward systems, organizations are amplifying this mechanism from the other side.

Chapter 4

Pipeline Solutions for Promoting Diversity: Lessons from Academia

The previous chapters of this dissertation focused on the symbolic aspects of diversity-related honorific programs by examining how merit scholarships from universities were perceived by people outside of those institutions. However, organizational diversity programs often provide material benefits as well, and some program designs consider mobility across positions more explicitly. This chapter offers a preliminary case study examination of one such program, the University of California President's Postdoctoral Fellowship Program (UCPPFP), which aims to increase faculty diversity in the University California (UC) system by recruiting outstanding recent doctoral recipients to well-mentored postdoctoral research positions across all fields of study.

After introducing the program and providing broader context, I leverage a natural experimental set-up to compare the career outcomes of one cohort of fellows to that of finalists from the same application cycle. Since the program is extremely competitive, and there is a large degree of randomness to how applicants are evaluated and selected, finalists are usually substantively comparable to fellowship recipients.

This comparison illustrates how the career paths of high-achieving Ph.D. students diverge during and after the entrance to this selective diversity postdoctoral program.

4.1 Background

4.1.1 Inequality in the Academy

Despite ongoing efforts to diversify the professoriate, tenure-track faculty in American acdemic institutions tend to be disproportionately White and male. A recent TIAA report found that colleges and universities have been hiring greater proportions of women, Black, Hispanic/Latino(a) and indigenous professors[1] over the past few decades, but those gains have been concentrated in non-tenure-track or part-time positions (Finkelstein et al. 2016). Between 1993 and 2013, the percentage of underrepresented minorities (URMs) in full-time tenured and tenure-track positions increased from 7.9% to just 10.6%. Women are better represented, increasing from 29.6% to 45.4%, although these gains are not evenly distributed across different fields (author's own calculations from the National Center for Higher Education Statistics' Integrated Postsecondary Education Data System as reported in Finkelstein et al. 2016).

Traditionally, research, policy, and public discourse have relied on the metaphor of the "leaky pipeline" to describe gender and racial inequalities in the transition from undergraduate education to faculty careers, especially in science, technology, engineering, and mathematics (STEM). In recent years, however, this paradigm has come under critique. First, some have challenged the very premise that academic careers follow a single linear "pipeline", pointing out that scholars have a variety of pathways into scientific fields and face a variety of challenges that can set them back

[1] Although Asian Americans experience marginalization in the academy and are not equally represented across different ethnic subgroups or in all disciplines (Poon et al. 2016), they are not usually considered underrepresented in the academy (Finkelstein et al. 2016; Griffin 2019)

(Gibbs 2014; Windsor and Crawford 2021). This perspective prioritizes expanding the number of ways that scholars can enter academy and calls for reforming support systems for junior researchers (such as better pay and family leave policies), rather than targeting the motivations of underrepresented scholars (Gibbs 2014; Goulden et al. 2011; Windsor and Crawford 2021).

Another perspective challenges the utility of the pipeline metaphor by debunking the idea that there are shortages of doctorates and an undersupply of scientists from underrepresented groups in the first place. (Gibbs et al. 2016; Miller and Wai 2015; Xue and Larson 2015). For example, by the 1990s, women who held bachelor's degrees in STEM were no less likely than their male counterparts to earn a STEM Ph.D. (Miller and Wai 2015), and the number of PhD graduates from URM backgrounds increased by a factor of 9.3 between 1980 and 2013, compared to a 2.6-fold increase in the number of PhD graduates from well-represented backgrounds (Gibbs et al. 2016). In medical schools in particular, there is actually no relation between the number of URM candidates and the number of URM faculty hired (Gibbs et al. 2016).

Despite the limitations of the leaky pipeline metaphor, the underrepresentation of women and URMs in the professoriate relative to their representation among Ph.D.s remains a stylized fact. Although part of this may be attributed to recent doctoral recipients' declining interest in academic careers (Roach and Sauermann 2017), different career goals cannot wholly explain persistent underrepresentation, as the number of Ph.D.s who hope to pursue academic careers continue to outpace the number of available tenure-track faculty positions National Academy of Sciences et al. (2014).

Thus, opportunity-enhancing initiatives like postdoctoral programs that help women and URM scholars who want academic careers transition into faculty positions are still useful interventions for diversifying the professoriate. For academic institu-

tions, they may also contribute towards recruiting goals. And although postdoctoral programs still require considerable institutional investment, in the context of broader structural issues like universities' increasing reliance on precarious adjunct labor and insufficient childcare policies, "pipeline" solutions are relatively easier to implement.

4.1.2 The Structural Role of Postdocs

The postdoctoral research position has become an established stage of professional training in many STEM fields and is becoming more common in the social sciences and humanities (National Academy of Sciences et al. 2014). Ideally, postdoctoral appointments provide recent doctoral recipients a temporary period of additional mentored training and research experience to prepare them for their chosen career paths. By bridging doctoral training with more permanent work, postdoctoral training is a critical juncture that is perhaps especially suitable for interventions concerning diversity and equal opportunity.

Despite the structural importance of this career stage in the faculty pipeline, however, postdocs comprise a poorly defined and hard to count population. There are no comprehensive data collected about them (National Academy of Sciences et al. 2014). Due to the lack of job title standardization, high mobility caused by the temporary nature of these positions, and the ad hoc nature of institutional postdoctoral administration and funding, postdoctoral outcomes are difficult to track (McConnell et al. 2018).

Even with the lack of comprehensive data on who postdoctoral researchers are and what they go on to do, we do know that the number of postdoctoral researchers is increasing in the United States. Between 2000 and 2012, the number of postdocs increased by 150% (Powell 2015). Compared to new U.S.-trained doctorates in 2000, recent doctorates in 2014 are less likely to have definite postgraduation plans – but

when they do, those plans are more likely to involve postdoctoral research positions (National Academy of Sciences et al. 2014).

According to the Survey of Earned Doctorates, 40% of all 2012 doctorate recipients who had postgraduate plans were committed to postdoctoral work (National Academy of Sciences et al. 2014). That number is even higher in STEM fields alone, and especially in life sciences. In a survey of STEM doctoral students from research-intensive U.S. universities, about three-quarters of life sciences doctoral students who graduated in 2013 took a postdoc, and about 46% of graduates in engineering and the physical sciences did so too (Sauermann and Roach 2016). The majority of postdoctoral research positions are also in the life sciences (National Academy of Sciences et al. 2014).

However, while a large number of doctoral students are becoming postdoctoral researchers, not very many postdoctoral researchers become faculty. One study estimates that about 65% of U.S. Ph.D.-holders continue to a postdoc, but only 15-20% of postdocs will become tenure-track faculty (Powell 2015).

This labor market imbalance is not new. Since the shift towards "big science" during World War II, principal investigators have needed ever larger groups of graduate students and postdoctoral researchers. While the demand for junior researchers has boomed, the number of faculty positions for them to move into has not grown commensurately over time. This is particularly stark in the biomedical sciences, where the value of postdoctoral researchers to the scientific enterprise remain clear, but the value of the training experience has become questionable for many postdoctoral researchers (Alberts et al. 2014; National Academy of Sciences et al. 2014).

Postdoctoral researchers also tend to be underpaid relative to their skills and training, and thus forgo considerable earnings (Stephan 2012). Additionally, funding streams for postdoctoral researchers have changed over time. The number of fellow-

ships and traineeships have basically remained the same since 1980, but the number of postdoctoral researcher positions created through federal research grants and non-federal funding sources have more than doubled (National Academy of Sciences et al. 2014).

Given that most postdoctoral research positions will not convert into tenure-track faculty positions, and median salaries are higher in non-postdoctoral research positions, why do so many recent Ph.D. graduates still take postdocs? Large surveys of postdocs suggest that the majority of them hope to pursue a career in academia, and they take postdoctoral positions because they think it will help them get a tenure-track faculty position (McConnell et al. 2018; Woolston 2020). However, the prevalence of postdocs cannot be attributed to this preference for academic careers alone. Sauermann and Roach (2016) find that a slight majority of STEM postdocs prefer other career paths when they are asked about their single most preferred career, but even then, they think that doing a postdoc will increase their chances of getting their desired jobs (despite there being no evidence that postdoctoral training helps with getting an industry job).

Thus, it seems like many new doctorates take postdoctoral positions because they are plentiful and have become a default of sorts. As a result, postdoc positions can end up being a sort of "holding pattern" rather than something that actually helps doctorates achieve their goals (Sauermann and Roach 2016).

Moreover, because of the difficulty of obtaining another position, postdocs on short-term contracts who can't find a permanent position often end up taking another postdoc position. According to Nature magazine's recent survey of postdocs around the world, more than 30% of respondents had already done at least two postdocs. A few reported completing as many as six postdoctoral appointments (Woolston 2020).

Finally, not only has the number of postdoctoral researchers have increased in recent years, but their demographic composition has as well. In the past 25 years,

the percentage of women in postdoctoral positions has grown by 10 percentage points in engineering and the physical sciences, and by 20 percentage points in the life sciences (National Academy of Sciences et al. 2014). However, historically underrepresented groups like Black and Latino Americans remain underrepresented, even relative to the population of doctoral degree recipients.

In short, while the postdoctoral research stage is a potentially powerful juncture for diversity and equal opportunity interventions, it is already a precarious position. Given its potential for exploitation, relatively low earning potential and low conversion rate into tenure-track faculty positions, creating more postdoctoral research positions might not necessarily create more opportunities for postdoctoral researchers.

4.2 Case Study: UCPPFP

The University of California President's Postdoctoral Fellowship Program is a strategic research site for examining the effects of selective "pipeline" oriented diversity programs. The study design offers the rare opportunity to observe not only the career outcomes of UCPPFP postdoctoral researchers, but also the career outcomes of similar scholars who pursued alternative career paths.

4.2.1 Overview and History

The University of California President's Postdoctoral Fellowship Program is the largest and one of the longest-running postdoctoral diversity programs in the United States. Founded in 1984, the program is designed to provide research opportunity and career development to scholars in all fields whose research, teaching, and service will "enhance the diversity of the academic community at the University of California" (UCPPFP Official Website). Currently, the program offers postdoctoral research fel-

lowships and training (renewable for up to two years), professional development, and faculty mentoring; fellows appointed after 1995 are also eligible for a salary hiring incentive at any of the University of California campuses. The program is designed to "give recent graduates an opportunity to become more competitive candidates for UC tenure-track appointments" (UCPPFP Official Website); it is not a research leave opportunity for scholars already on the tenure-track who intend to stay at their current institution.

UCPPFP oversees two types of fellowships: President's Postdoctoral Fellowships and Chancellor's Postdoctoral Fellowships. Both types of fellows can be appointed at any of the ten University of California campuses (Berkeley, Los Angeles, San Diego, Davis, Irvine, Santa Barbara, San Francisco, Santa Cruz, Riverside, and Merced). There is one centralized application for the two fellowships. In this application, applicants must select which campus they want to be hosted and obtain the sponsorship of a tenured faculty member at that campus who will then serve as their primary mentor. Applicants who are not selected for a President's Postdoctoral Fellowship are then automatically considered for a Chancellor's Postdoctoral Fellowship at the UC campus where their proposed faculty mentor is located[2]. Although the selection process and funding structure for the two types of fellowships are slightly different, President's Fellows and Chancellor's Fellows participate in UCPPFP in the same capacity, receiving the same benefits, mentoring, terms of appointment, and eligibility for hiring incentives for tenure-track faculty appointments[3].

In recent years, the program office has worked vigorously to maximize the number of fellowships they are able to offer every year. This has mostly been achieved

[2]The three UC National Laboratories can also host President's Fellows, but they do not host Chancellor's Fellows

[3]UCPPFP has also developed a national network of partnerships with postdoctoral diversity programs at other institutions, including University of Michigan, University of Colorado, University of Minnesota, Carnegie Mellon University, Georgia Tech, and others. Applicants can apply to those postdoctoral programs with the same application provided they submit separate mentor letters and research proposals. However, since these other programs are much smaller and I don't have the full universe of application data, they are beyond the scope of this paper.

by creating partial fellowships or fellowships without salary and permitting candidates to accept sources of support from other prestigious fellowships or grants (Lawson 2018). This type of cross-subsidization leverages external support and allows Fellows to further retain evidence of individual accomplishments – for example, a candidate who secured funding with a prestigious NIH grant is able to accept that award and use it to fund their fellowship, thereby freeing up funds for another candidate who does not have external support. However, as Figure 4.1 shows, the expansion of the program has not made it any less competitive, as the number of applications continue to increase each year[45]. In the past eight years, the "acceptance rate" for fellowships has never exceeded 8%.

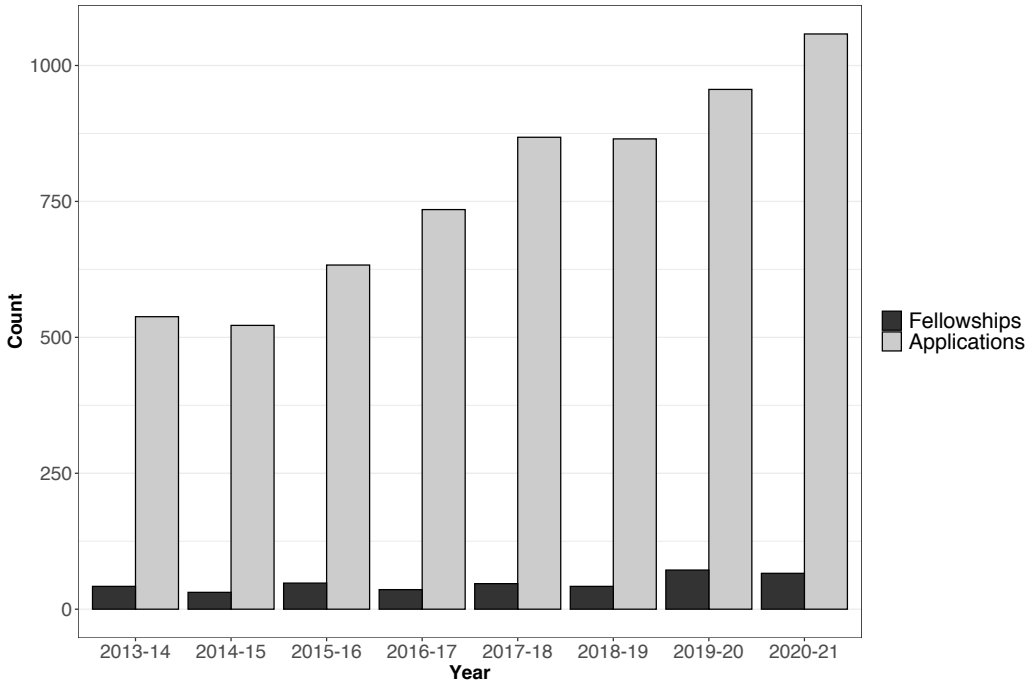


Figure 4.1: Number of applications and fellowship awards, 2013-2021

[4]This figure only counts complete applications that were sent out for review. Applications missing materials, including letters of recommendation from the proposed faculty sponsor/mentor or the Ph.D. advisor, are not considered complete. UC's public materials report the number of applications that are initiated in the application portal, not the number of complete applications, so the numbers I am reporting are lower than the ones publicly reported.

[5]There is no single reason why applications have grown at such a rapid pace. It's possible that new institutional partnerships with other universities made the application more attractive. The program office and board members may have done a better job marketing the program. The job market could also be getting more competitive.

A key feature of UCPPFP is the faculty hiring incentive program, which provides 85,000 dollars per year for five years to campuses that hire current or former President's or Chancellors' Fellows into ladder-rank faculty positions. Proper use of this initiative program is a major component of UC's formally stated best practices for increasing faculty diversity. Like other academic institutions, the share of women and URM professors in tenure-track positions at UC is not commensurate with minority representation in the student body, or even the faculty applicant pool. In fact, as a public memo from the chair of the academic council to the provost noted in 2016, underrepresented minority faculty hiring at UC has actually declined since 2006,"even as the availability pool has increased" (Hare 2016).

Under this faculty hiring incentive program, all fellows appointed after 1995 are eligible for the hiring incentive in support of their first UC ladder-rank position, regardless of previous rank or position (e.g. they are still eligible if they are already tenure-track faculty at a non-UC institution). When the faculty hiring incentive program began in 2003, the number of incentives awarded per year was restricted to 12 across all UC campuses and it was not applicable to the professional schools. In 2017, when UC President Janet Napolitano committed an additional five million dollars to the hiring incentive and initiate a new start-up hiring incentive for fellows in STEM fields, she also lifted the restrictions on incentives, removing both the annual limit and expanding eligibility to the professional schools. In the public letter announcing this development, Napolitano called the faculty hiring incentive program "our most successful program in diversifying the UC faculty" (Napolitano 2017). In personal correspondence, the program staff told me that before the restriction was lifted, the the number of hiring incentives claimed every year rarely met the 12-per-year cap, speculating that the perceived scarcity of the hiring incentive funds discouraged fellows and departments from thinking that they would be able to claim one of the 12 spots. However, the number of incentives claimed has risen in recent years.

4.2.2 Eligibility and Selection Process

To be eligible for UCPPFP, applicants must hold or receive a Ph.D. (or other terminal degree in their discipline, such as a JD, MD, MFA, EdD, etc.) from an accredited university and be legally authorized to work in the United States. The program is geared towards recent graduates preparing for careers in university teaching and research. Although candidates who received their degrees less recently are still eligible, applicants who have held postdoctoral appointments at any university for over five years are not[6]. Candidates that already have tenure-track appointments or offers at other universities, or have had significant postdoctoral experience typical of competitive tenure-track candidates in their field are "not favored in the selection process" unless there is a clear explanation of how candidates' participation in UCPPFP would align with its goal of diversifying UC faculty (UCPPFP official website).

There are no eligibility criteria related to race, ethnicity, or gender[7], but applicants are evaluated for their potential to contribute to diversity and equal opportunity through research, teaching, and service. UC recognizes a variety of ways to contribute to diversity and equal opportunity, including public service towards increasing equitable access for women and minorities in fields where they are under-represented, research that focuses on underserved populations or inequalities, or research perspectives that draw from understandings of groups historically underrepresented in higher education.

The application review and selection process for President's Postdoctoral Fellowships has two main stages: off-site review and on-site review. First, during off-site review, all completed applications are reviewed by at least two reviewers who are

[6]The University of California has a policy that limits total time of service in academic postdoctoral appointments to five years.

[7]Historically, UCPPFP was designed to encourage women and racial minorities into pursuing careers in the UC system. However, ever since Proposition 209 was approved in 1996, state governmental institutions in California have been prohibited from considering race, sex, ethnicity in public employment, public contracting, and public education.

in the applicants' own fields or in related fields. Applications that advance to on-site review are then evaluated by two or three new reviewers and discussed in disciplinary panels. Through group discussion, the panels will first select the top applicants to be recommended for a fellowship, then select a group of alternates and place them in ranked order, and then compile a list of unranked finalists.

Reviewers are ladder-rank faculty in the UC system who volunteer their service. Each faculty reviewer receives between 5 and 15 applications for the off-site round of review. Reviewers are allowed to evaluate applicants who has attended or is proposing a fellowship at their institution, but they are asked to refrain from evaluating applicants they have collaborated with or advised.

In addition to considering applicants' prior history of contributions towards diversity and equal opportunity and how they will use the fellowship tenure to expand these activities, faculty reviewers also evaluate applicants' academic accomplishments and the strength of their research proposal. Reviewers are asked to consider the proposed research's innovation and feasibility, the proposed mentor's plan for the applicant's scholarly development during the fellowship, recommendation letters' detailing of how the applicant's research contributes to their field, and the potential impact of the award on the applicant's career trajectory.

Notably, these criteria may be applied differently by disciplinary group. For example, the 2020-2021 expanded rubric given to reviewers during on-site review presented different considerations for research contribution by discipline. For research falling in the humanities, arts, social sciences and professions, reviewers were instructed to also "evaluate how the research contributes to our understanding of issues such as diversity, multiculturalism, or communities underserved by traditional academic research". For research in mathematics, engineering, life sciences, and physical sciences, reviewers were instructed to also "evaluate the demonstrated record of mentoring or outreach, or how the candidate intends to incorporate mentoring or

outreach activities into their work that will promote access and opportunity in U.S. higher education."

During off-site review, evaluators are instructed to give each application an A, B, or C rating[8]. Off-site reviewers do not discuss applications together or share their scores with each other. Off-site reviewers typically belong to the same broad disciplinary group (Arts and Humanities, Social Sciences and Professional Fields, Life Sciences, and Mathematics, Engineering, and Physical Sciences), but in more interdisciplinary fields, applications may be reviewed by reviewers from different disciplinary groups. If the application was evaluated by more than two reviewers, the two highest scores were recorded. Applications must receive at least one "A" score – i.e. an A-A score, an A-B score, or an A-C score– to advance to the next stage, the on-site review[9]. In addition to the letter scores, faculty reviewers are asked to write short statements explaining their rationales for their assigned scores; these comments can be seen by reviewers in future rounds of evaluation and may be provided to applicants in a redacted form.

During on-site review, faculty reviewers form panels based on their disciplinary group. Before the applications are read again, panels will first review the selection criteria together and discuss how they should be applied. To help panelists achieve mutual understanding of the evaluation criteria, panel chairs will present two files from previous application cycles for general discussion and evaluation. After this initial calibration, faculty reviewers will review the qualifying applications in their disciplinary group. Unlike the A-C scoring system, the on-site system uses a 1-5 numeric rating scale. The "1" rating is reserved for the top 10% of applicants, who are deemed outstanding and show clear potential for appointment to a UC professor position after participation in UCPPFP. Applications that receive a 1 rating would

[8]This A-C scale was expanded into an A-E scale starting in 2020

[9]After 2020, applicants were advanced to the next round if they received at least one A score *and* their two highest scores were less than three points apart. So, A-A, A-B, and A-C scores still qualify, but A-D and A-E scores do not.

have the highest recommendation for receiving a fellowship. The 2 rating is reserved for the top 11%-20% of applicants; applications that receive this rating would have a strong recommendation for being ranked as an alternate. Applications receiving 3 ratings are strong in all areas but not outstanding in any area, and applications receiving 4 or 5 ratings are deemed to have minor weaknesses in one or more areas.

Applications that received an A-A score during off-site review have priority during on-site review and will be rated by and commented on by at least two on-site reviewers. Applications that only received one A score (e.g. A-B scores and A-C scores) will be only be rated by one faculty reviewer. If that reviewer gives a 1 rating, then a second reviewer will also read the application. When two faculty reviewers have completed evaluations and there is a discrepancy between the two numeric ratings, the panel chairs may ask a third reviewer to evaluate the complete file. An applicant's final rating is the average of the two or three reviews. After the applications have been read at least twice and scored, the review panel convenes for group discussion and selects the President's Postdoctoral Fellowship receipts, ranked alternates, and unranked finalists. According to program staff, discrepant scoring – between two on-site reviewers or between the off-site and on-site reviewers – is fairly common. It is not unheard of for A-A applicants to receive lower scores during on-site review, and an applicant with an A-C score from the first round may receive 1 ratings during on-site review and ultimately be recommended for a fellowship. This level of disagreement reflects a certain degree of random variation in the ratings. Indeed, UCPPFP program staff told me that because the panels are choosing between applications of extremely similar quality by the end of the selection process and there are external constraints about the number of fellowships they can offer that application cycle, the differentiation between candidates who are recommended for fellowships and the other finalists is as good as random.

Finally, after the President's Fellowship recipients are selected, reviewers from each UC campus are invited to select the applicants to be recommended for Chancellor's Fellowships. Unlike the reviewer pool for the President's Fellowship, reviewers for Chancellor's Fellowships may include senior academic administrators (provosts, chancellors, etc.) or personnel from Diversity, Equity, and Inclusion offices in addition to faculty. Typically, reviewers will select fellows from the group of alternates or unranked finalists. However, before 2019, reviewers also had access to the top 40% of all online applications and they would occasionally dip back into this larger pool to select fellows. This type of behavior is more common at campuses that receive fewer applications, because it they had fewer candidates to choose from in the panel-curated pool of alternates and finalists. The fact that faculty and administrators selecting Chancellor's fellows would sometimes ignore the recommendations and screening provided by a previous set of reviewers further illustrates the amount of idiosyncrasy present in the evaluation process.

4.3 Methods and Data

4.3.1 Constructing the Dataset

The UCPPFP staff provided me with data on all applications initiated on the online application portal from fall 2012 to fall 2019 and reviewer scores and comments for all completed applications during the same time period. The application data included applicants' names, the UC campus they were applying to, the discipline and sub-discipline of the research proposal, the proposed faculty mentor(s) and what UC campuses they are associated with, the applicants' Ph.D. granting institution, the Ph.D. completion date, the discipline and sub-discipline of the Ph.D., the applicants' disability status, gender, and race/ethnicity, and the name, department, and institution of the applicants' thesis advisor. The reviews data includes the off-site

review scores and comments, the highest two scores from off-site review, and the ultimate status of the application(e.g. Selected for PPFP/CPF, Alternate or Finalist, Not selected in first round, Not selected in second round, Withdrawn)[10]. For the 2017-2018 cohort and after, the reviews data also includes on-site reviewer scores and comments. However, the on-site reviewer scores and comments for prior cohorts were not recorded digitally; the physical paper copies are stored in the UCPPFP office in Oakland, California. Thus, the data presented in this chapter from the 2014-2015 cohort does not include on-site reviewer scores.

In the 2014-2015 application cycle, 642 applications were initiated in the application system. 30 were never submitted and 90 were never sent out for review because letters from the proposed mentor or the thesis advisor were missing. I omitted these 120 applications from the total applicant pool because they were never read by reviewers and their applicants were never actually considered for a fellowship. Additionally, 12 applicants withdrew (two withdrew before off-site reviews were complete and 10 withdrew after off-site reviews were submitted), and one application never received a second off-site review. I kept these 13 observations in the dataset because these applicants were still considered for the fellowship at some point in the evaluation and selection process.

Among the 522 applicants whose complete applications were evaluated, 19 received President's Fellowships, 12 received Chancellor's Fellowships, and 34 applicants were selected as alternates or unranked finalists (only about 12% of all applicants became fellows or finalists). Unfortunately, the program office has not yet shared more detailed spreadsheets that specify the ranking of the alternates with me, which prevents me from disaggregating the alternates from the unranked finalists. For the purposes of this analyses, I treat these two pools of applicants as one group and

[10]UCPPFP maintains a public directory of all President's and Chancellor's fellows going back to the program's inception on their official website. I matched the names on this directory to the names in the application data to validate that the status associated with those applications were properly identified.

refer to them as "Finalists" throughout the rest of the chapter. Without these more detailed spreadsheets, I am also unable to observe if anyone was offered a fellowship but declined it; however, UCPPFP program staff told me that this is extremely rare, and they do not believe anyone declined an offer that year.

I used a two-step process to identify finalists' current positions and the positions they held during the fellowship year. First, I searched scholars by name in the Academic Analytics Database, an online directory of over 275,000 faculty members at over 400 U.S.-based research universities. These public data were compiled with web-scraping tools by Academic Analytics, an information services company that offers software tools to administrators at academic institutions. If I am able to find a scholar's current academic appointment in the Academic Analytics Database, I would proceed to search for the position they held in the 2014-2015 academic year (e.g. what they did instead of participating in UCPPFP). I would usually be able to find this information on their Curriculum Vitaes (CVs) when available. If I could not find a CV, I would look for their publications in the Academic Analytics Database, on Google Scholar, Academia.edu, or ResearchGate to track authors' institutional affiliations by year.

When I was not able to find the scholar in the Academic Analytics Database, I entered their name (sometimes in conjunction with the discipline they received their Ph.D. in or the name of their Ph.D.-granting institution) into the Google search engine. I was often able to find a personal website, a webpage hosted by an academic institution, a public CV, and/or a public LinkedIn page[11].Finally, I also looked up the current positions of 2014-2015 President's and Chancellor's Fellows. Although the UCPPFP office also keeps records on fellows' academic appointments and positions after they leave the program, this information is several years old for most cohorts.

[11]On LinkedIn, there is often more information available to users who are logged into the platform. However, information that is visible to a logged in user is not always necessarily visible to "the public". To avoid seeing anything that is restricted to logged-in users, I stayed logged out of my personal LinkedIn account and I opened a new private browser window for every search.

Overall, I collected the following information for the 34 finalists from the 2014-2015 fellowship cycle: current position, current institution, current department, fellowship year position, fellowship year institution, and fellowship year department. I also collected updated information on current position, current institution, and current department for the 19 President's Fellows and 12 Chancellor's Fellows. From this information, I created indicator variables identifying whether or not the current and fellowship year institutions are colleges or universities, whether or not the current and fellowship year positions are tenure-track faculty positions[12], and if the current or fellowship year institutions are colleges or universities, whether or not they were UC campuses.

For finalists, I also created indicator variables identifying whether or not their fellowship year position was another postdoctoral appointment, whether or not the finalist spent another year in graduate school during the fellowship year, and whether or not their fellowship year position was funded by another diversity-related initiative (such as a diversity postdoctoral program at another institution, or specific federal grants dedicated to training a diverse scientific workforce)[13].

4.3.2 Overview of the 2014-2015 Cohort

Table 4.1 shows the characteristics of the President's Fellows, the Chancellor's Fellows, and finalists in comparison to the overall applicant pool. The first set of rows show the distribution of fellows, finalists, and applicants across the ten UC campuses. UC

[12]For scholars at U.S.-based institutions, I define tenure-track faculty position as "assistant professor" or "associate professor", unless otherwise specified. For example, in one instance, a scholar was a "Tenure-Track Assistant Teaching Professor" at an institution where assistant teaching professors are ladder-faculty. For scholars at non-U.S. based institutions, I used the positions commensurate with the local academic rank system (e.g. "lecturer" in Israel).

[13]This coding is based on whatever information the scholar has chosen to self-report and publicly announce on their CVs or personal websites. There is likely some degree of measurement error because scholars might not always report the full information associated with a position. For example, a scholar might simply list that they were a postdoctoral associate in a particular lab without specifying what grant they were funded by. However, because diversity-related federal grants and postdoctoral fellowships tend to be highly competitive and prestigious, recipients are likely motivated to report them.

Berkeley received the highest number of fellowship applications, hosted the most fellows, and was the most common choice among the finalists. Almost 70% of all applications went to just four UC campuses: Berkeley, LA, San Diego, and Davis. These four schools also hosted about 80% of all fellows.

The next section of Table 4.1 shows that applications and fellowships are also not evenly distributed across the four major disciplines (the disciplines shown here are the categories that applicants identified their research proposal with, not the discipline that their Ph.D. is in; see Supplemental Materials for more a detailed discussion of disciplines and instances of interdisciplinarity). The social sciences and professional fields receive the most number of fellowship applications, accounting for over one-third of all applications. In fact, the number of applications in the social sciences is almost triple the number of applications in mathematics, engineering, and the physical sciences. However, because fellowship selections are made separately for each discipline, the fellows are more disciplinarily balanced than the overall applicant pool. For example, there are actually more fellows in mathematics, engineering, and the physical sciences than there are in social sciences and professional fields (9 compared to 7). This means that overall, fellowships in the social sciences, and to a lesser extent, the arts and humanities, are more selective than they are in the "hard" sciences.

Given the program's focus on diversity and equal opportunity, it is perhaps unsurprising that the majority of applicants, fellows, and finalists are women or racial minorities. Women were still slightly underrepresented among fellows compared to applicants – they made up 72% of the applicant pool, but only about 61% of the fellows that year were women. White, African American, and Asian American scholars were

Table 4.1: Characteristics of Fellows, Finalists, and Applicants (2014-2015 Cohort)

	President's Fellows		Chancellor's Fellows		Finalists		All Applicants	
	Count	%	Count	%	Count	%	Count	%
Total	**19**		**12**		**34**		**522**	
Campus								
UC Berkeley	5	26.3	2	16.7	9	26.5	113	21.6
UC Los Angeles	2	10.5	-	-	8	23.5	97	18.6
UC San Diego	2	10.5	2	16.7	4	11.8	70	13.4
UC Davis	4	21.1	4	33.3	4	11.8	75	14.4
UC Irvine	1	5.3	4	33.3	2	5.9	45	8.6
UC Santa Barbara	1	5.3	-	-	1	2.9	35	6.7
UC San Francisco	1	5.3	-	-	2	5.9	23	4.4
UC Santa Cruz	1	5.3	-	-	2	5.9	30	5.7
UC Riverside	2	10.5	-	-	1	2.9	28	5.4
UC Merced	-	-	-	-	1	2.9	6	1.1
Discipline								
Arts and Humanities	5	26.3	3	25.0	9	26.5	155	29.7
Social Sciences	5	26.3	2	16.7	11	32.4	191	36.6
Life Sciences	3	15.8	4	33.3	11	32.4	108	20.7
Physical Sciences	6	31.6	3	25.0	3	8.8	68	13.0
Gender								
Woman	13	68.4	6	50.0	23	67.6	374	71.6
Man	5	26.3	6	50.0	11	32.4	143	27.4
Declined to State	1	5.3	-	-	-	-	5	1.0
Race/Ethnicity								
White	3	15.8	4	33.3	7	20.6	145	27.8
Latino	8	42.1	6	50.0	14	41.2	129	24.7
African American	2	10.5	1	8.3	5	14.7	63	12.1
Asian American	-	-	1	8.3	6	17.6	112	21.5
American Indian	1	5.3	-	-	1	2.9	11	2.1
Other	3	15.8	-	-	1	2.9	47	9.0
Declined to State	2	10.5	-	-	-	-	15	2.9
PhD Completion Date								
2014 or Later	14	73.7	4	33.3	20	58.8	312	59.8
Before 2014	5	26.3	8	66.7	14	41.2	210	40.2
PhD Institution								
University of California	10	52.6	7	58.3	12	35.3	245	46.9
Other Public University	6	31.6	1	8.3	8	23.5	118	22.6
Other Private University	3	15.8	4	33.3	13	38.2	145	27.8
University Outside of U.S.	-	-	-	-	1	2.9	14	2.7

Note: Percentages may not add to 100 due to rounding

underrepresented among fellows relative to the applicant pool, while Latino[14] scholars were relatively over-represented among the fellows compared to the applicant pool.

Most of the applicants had not received their Ph.D.'s at the time of application, and President's Fellows were more likely to be Ph.D. candidates than the overall applicant pool. Moreover, fellows were slightly more likely to be Ph.D. graduates of the UC system than the applicant pool overall. A little under half of all applicants received or expected to receive their Ph.D.s from a UC school, while a little over half of all fellows are UC alums. The rest of the fellows are about evenly split between other public universities and other private universities. Very few applicants and no fellows received their Ph.D.'s from an university outside of the U.S.

The composition of applicants – in terms of gender, race, and time since Ph.D. – also vary considerably across disciplines[15]. Table 4.2 shows that the applicant pool in mathematics, engineering, and the physical sciences has a higher proportion of men than the applicant pools in the arts and humanities, social sciences and professional fields, and even the life sciences. Compared to applicants in the arts, humanities, social sciences, and professional fields, applicants in life sciences, math, engineering, and physical sciences are more likely to be White and less likely to be Asian American. Table 3 also shows that almost three in four applicants in the life sciences already received their doctorate degrees when they applied to the fellowship – a sharp contrast to the three other disciplinary fields, where applicants are far more likely to still be Ph.D. candidates at the time of application. This might reflect the unique pervasiveness of second postdoctoral appointments in the life sciences.

[14]The UC application form has three categories that typically fall under the "Hispanic/Latino" designation as defined by the Office of Management and Budget: "Latino", "Mexican American - Chicano/a", and "Puerto Rican". I aggregated these three categories into one broader category called "Latino" to be more consistent with the categories used by the Department of Education and other agencies. About 60% of scholars in this broader category self-identified as Chicano/a and 13% self-identified as Puerto Rican.

[15]There is likely some degree of heterogeneity within each disciplinary group as well. However, the finer-grain sub-discipline data is too sparse to disaggregate with just one year of data.

Table 4.2: Characteristics of Applicants by Discipline (2014-2015 Cohort)

	Arts and Humanities		Social Sciences		Life Sciences		Physical Sciences	
	Count	%	Count	%	Count	%	Count	%
Total	**155**	**29.7**	**191**	**36.6**	**108**	**20.7**	**68**	**13.0**
Campus								
UC Berkeley	34	21.9	42	22.0	16	14.8	21	30.9
UC Los Angeles	34	21.9	38	19.9	16	14.8	9	13.2
UC San Diego	8	5.2	26	13.6	26	24.1	10	14.7
UC Davis	24	15.5	28	14.7	13	12.0	10	14.7
UC Irvine	14	9.0	19	9.9	8	7.4	4	5.9
UC Santa Barbara	17	11.0	8	4.2	7	6.5	3	4.4
UC San Francisco	-	-	10	5.2	10	9.3	3	4.4
UC Santa Cruz	10	6.5	9	4.7	9	8.3	2	2.9
UC Riverside	13	8.4	8	4.2	3	2.8	4	5.9
UC Merced	1	0.6	3	1.6	-	-	2	2.9
Gender								
Woman	113	72.9	139	72.8	78	72.2	44	64.7
Man	38	24.5	52	27.2	30	27.8	23	33.8
Declined to State	4	2.6	-	-	-	-	1	1.5
Race/Ethnicity								
White	34	21.9	44	23.0	44	40.7	23	33.8
Latino	36	23.2	51	26.7	24	22.2	18	26.5
African American	17	11.0	30	15.7	8	7.4	8	11.8
Asian American	42	27.1	45	23.6	17	15.7	8	11.8
American Indian	3	1.9	5	2.6	3	2.8	-	-
Other	16	10.3	14	7.3	9	8.3	8	11.8
Declined to State	7	4.5	2	1.0	3	2.8	3	4.4
PhD Completion Date								
2014 or Later	96	61.9	138	72.3	29	26.9	49	72.1
Before 2014	59	38.1	53	27.7	79	73.1	19	27.9
PhD Institution								
University of California	76	49.0	107	56.0	34	31.5	28	41.2
Other Public University	31	20.0	33	17.3	37	34.3	17	25.0
Other Private University	44	28.4	45	23.6	34	31.5	22	32.4
University Outside of U.S.	4	2.6	6	3.1	3	2.8	1	1.5

Note: Percentages may not add to 100 due to rounding

Finally, before comparing the outcomes of the 2014-2015 cohort of fellows and finalists, I compared the off-site reviewer scores of President's Fellows, Chancellor's Fellows, and finalists to the score distribution of the overall applicant pool (Table 4.3). Although these scores from the first round of review do not strictly determine which applicants receive fellowships – recall that recommendations are made from on-site reviewer scores and panel discussions – they still provide some insight on how the applicants' work was perceived by others' in their field.

Table 4.3 shows that most President's and Chancellor's Fellows received A-A scores during off-site review, and President's Fellows were more likely than Chancellor's Fellows to have earned an A-A score during the first round of review. By contrast, although half of the alternates and unranked finalists also received A-A scores, a larger portion of them received A-B scores. However, the differences in score distribution between President's Fellows, Chancellor's Fellows, and Finalists pale in comparison to the differences between them and the overall applicant pool. In particular, it is worth noting that only 80 applicants out of 522 received A-A scores, and half of them were still not even selected as a finalist.

Table 4.3: Off-Site Scores for Fellows, Finalists, and Applicants (2014-2015 Cohort)

	President's Fellows		Chancellor's Fellows		Finalists		All Applicants	
First Round Scores	Count	%	Count	%	Count	%	Count	%
A-A	14	73.7	8	66.7	17	50.0	80	15.3
A-B	2	10.5	3	25.0	14	41.2	127	24.3
A-C	3	15.8	1	8.3	3	8.8	75	14.4
B-B	-	-	-	-	-	-	61	11.7
B-C	-	-	-	-	-	-	114	21.8
C-C	-	-	-	-	-	-	62	11.9
Review Not Complete	-	-	-	-	-	-	3	0.6

Note: Percentages may not add to 100 due to rounding

4.4 Results

4.4.1 Fellows and Finalists' Current Positions

Figure 4.2 shows the number of President's Fellows, Chancellor's Fellows, and finalists who are in tenure-track faculty positions, non-tenure track academic positions, and non-academic positions. 17 of the 19 former President's Fellows and 8 of the 12 former Chancellor's Fellows (about 89.5% and 66.7%, respectively) are currently in tenure-track faculty positions, compared to 21 out of 34 (about 61.8%) of finalists. Among those not who are not currently tenure-track faculty, most are in non-academic positions. A small number of former Chancellor's Fellows and finalists are in non-tenure track positions at colleges and universities (e.g. administration, lecturers or adjuncts, or research support).

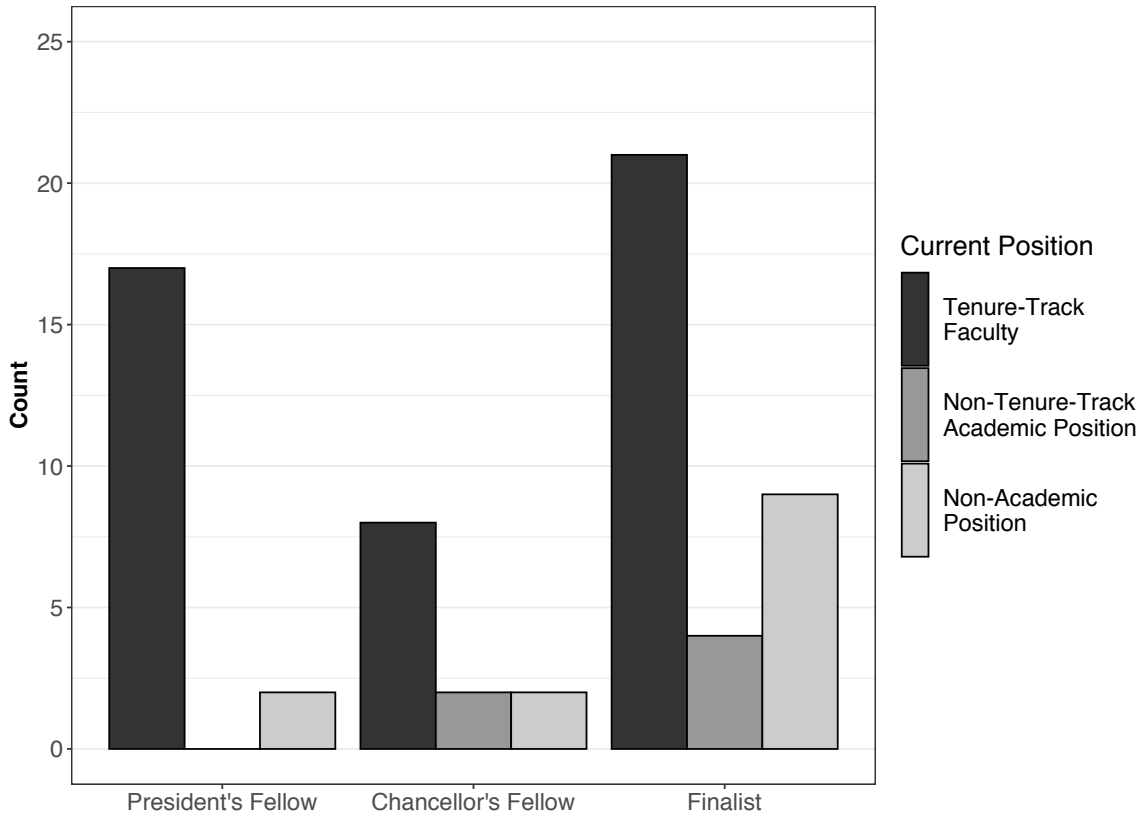


Figure 4.2: Current positions of fellows and finalists (2014-2015 cohort)

Figure 4.3 shows that these patterns differ by discipline, possibly due to the availability of job opportunities in industry. In the arts and humanities, the vast majority of finalists (8 out of 9) are currently in tenure-track faculty positions. Similar trends played out in social sciences and professional fields (8 of the 11 finalists are currently tenure-track faculty) and in math, engineering, and physical sciences (2 out of the 3 finalists became tenure-track faculty). In these three disciplinary groups, 2014-2015 finalists are more likely to be in tenure-track faculty positions than not, and at similar rates as the President's and Chancellor's Fellows from that cohort.

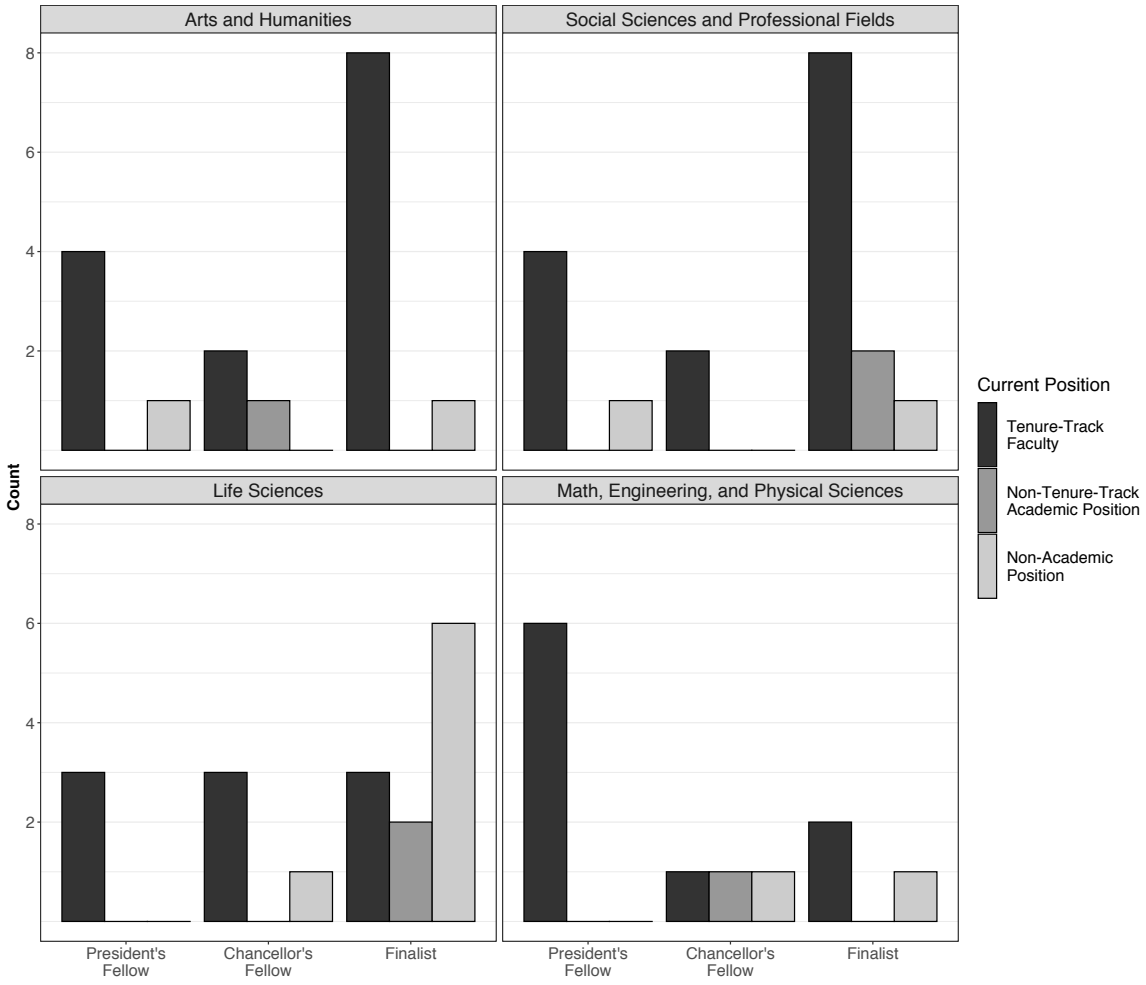


Figure 4.3: Current positions of fellows and finalists by discipline

By contrast, the career paths of fellows and finalists diverge in the life sciences. Like the fellows in social sciences, 6 of the 7 fellows in the life sciences are currently in tenure-track faculty positions. However, only 3 of the 11 finalists became tenure-track faculty; most life science finalists are no longer in the academy today.

Figure 4.4 shows that both men and women finalists are less likely to become tenure-track faculty than their counterparts who received fellowships. Among men, 9 out of 11 fellows are currently tenure-track faculty, while only 6 out of 11 finalists are. Among women, 15 out of 19 fellows are currently on the tenure-track, while 15 out of 23 finalists are.

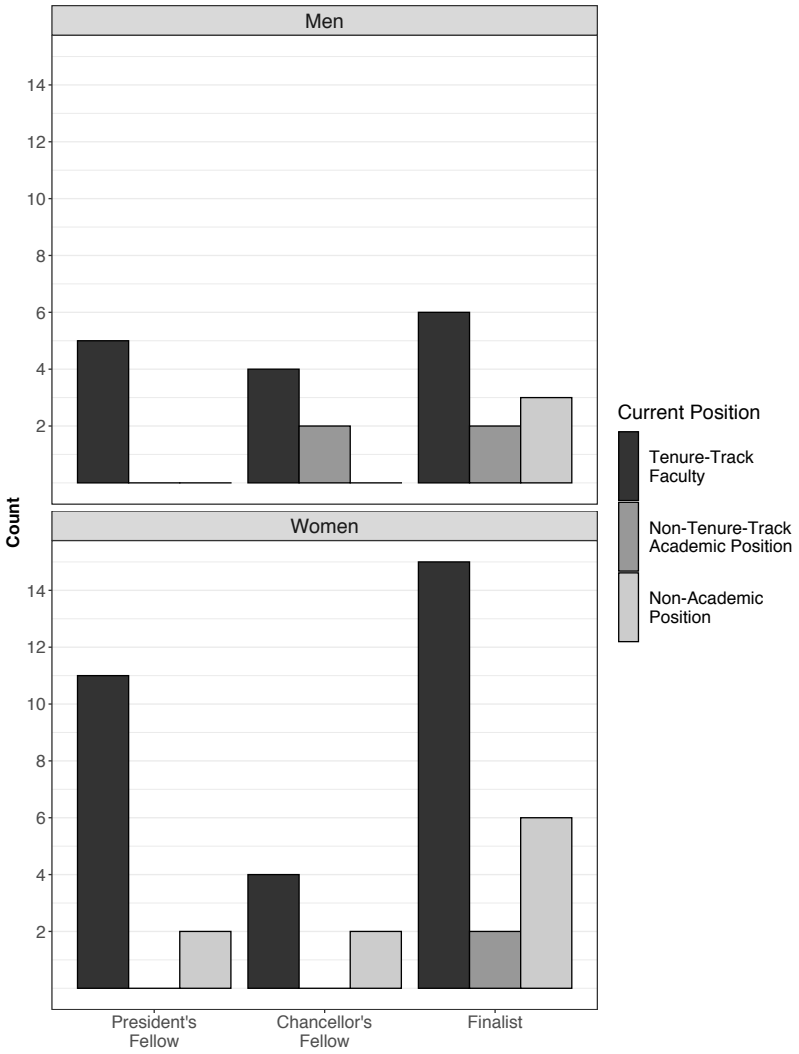


Figure 4.4: Current positions of fellows and finalists by gender

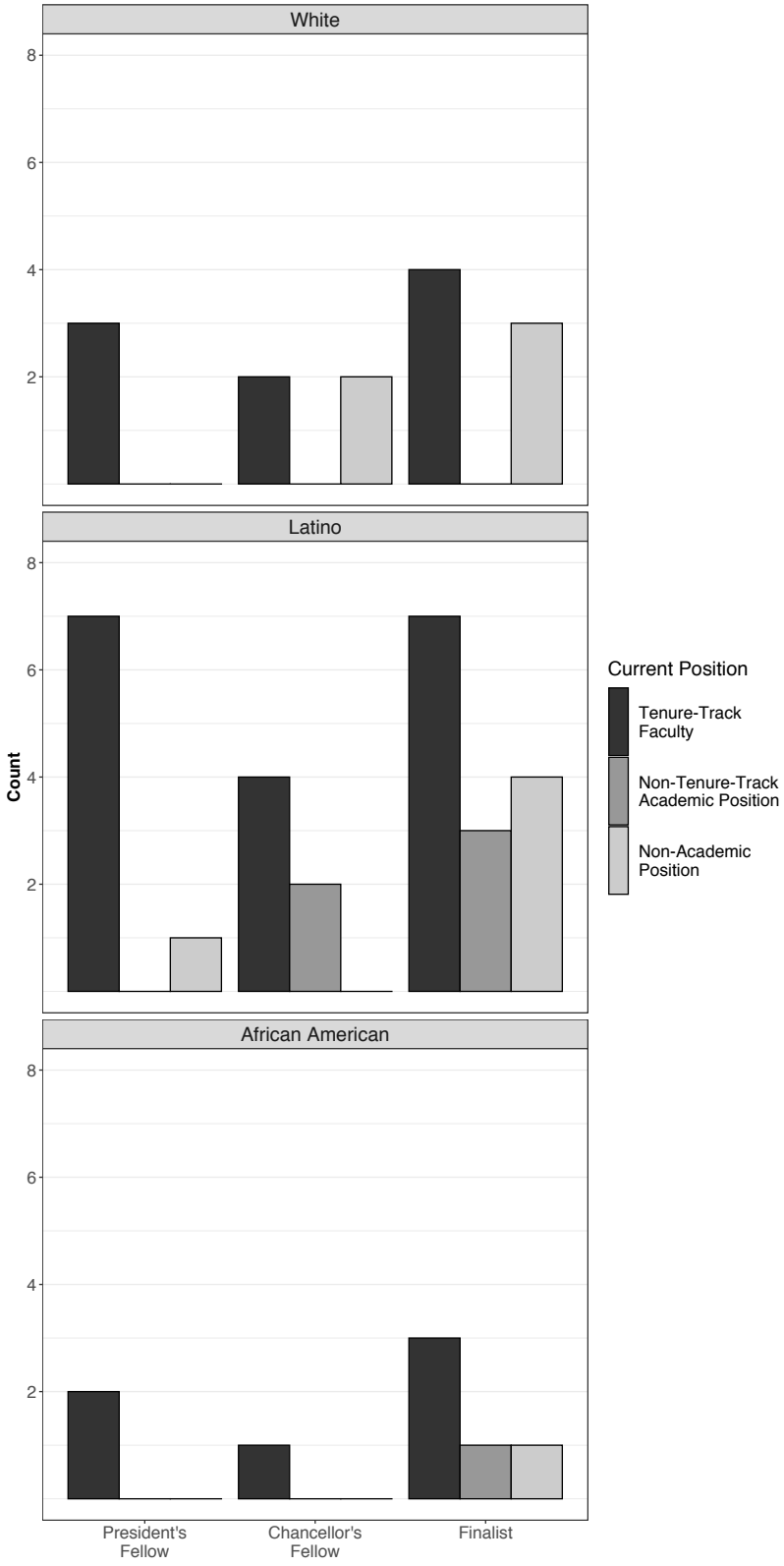


Figure 4.5: Current positions of fellows and finalists by race

Similarly, Figure 4.5 shows that finalists of all races are less likely to be tenure-track faculty than fellows. 5 out of 7 White fellows, 11 out of 14 Latino fellows, and all three African American fellows are currently tenure-track faculty. By contrast, only 4 out 7 White finalists, 7 out of 14 Latino finalists, and 3 out of 5 Black finalists hold tenure-track faculty positions today.

4.4.2 Retaining Fellows in the UC System

The results in the previous section suggest that UCPPFP fellows, and especially President's Fellows and scholars in the life sciences, are more likely to be in tenure-track faculty positions today. However, the stated goal of UCPPFP is not just to increase diversity in the professoriate, but to specifically increase faculty diversity in the University of California system. Figure 4.6 shows the number of fellows and finalists who are currently tenure-track faculty in the UC system and tenure-track faculty at other academic institutions.

Currently, 7 out of 17 former President's Fellows and 6 out of 8 former Chancellor's Fellows are faculty at a UC campus (4 of the fellows are faculty at the same institutions that hosted their postdoctoral fellowships). Taken together, that means that just a little over half of all UCPPFP fellows who became tenure-track faculty stayed in the UC system. Overall, about 40% of the 2014-2015 fellowship cohort became UC faculty. However, this proportion is much higher than it is for finalists – of the 21 finalists who are currently tenure-track faculty, only 5 teach at a UC school (approximately 1 in 4).

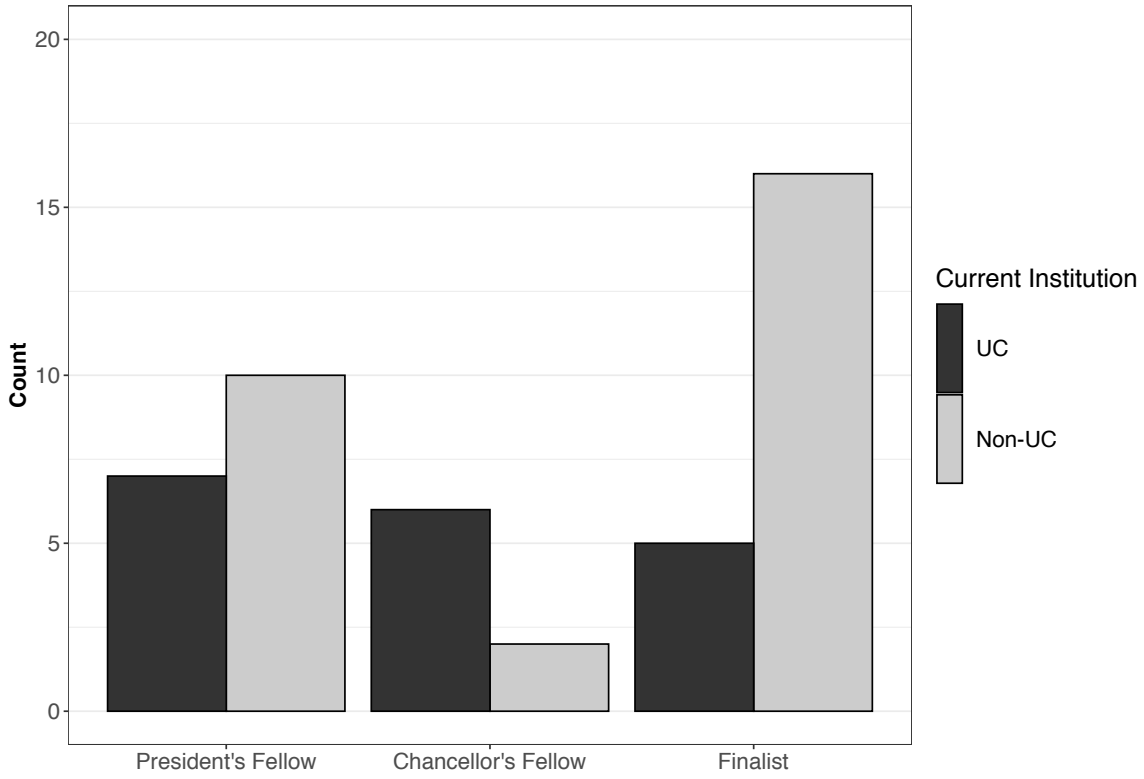


Figure 4.6: Current institutional affiliations of fellows and finalists who are tenure-track faculty

4.4.3 Alternative Paths

What did finalists end up doing in the 2014-2015 academic year instead? Figure 4.7 shows that about half of them (18 out of 34) began a different postdoctoral appointment[16]. Almost half of those other postdoctoral appointments (8 out of 18) were still in the UC system. Three of the 34 finalists stayed in graduate school, four took on non-tenure-track academic jobs (e.g. adjucting), and only two took non-academic positions. Thus, even though many of these finalists left the academy eventually and are now in non-academic positions, most did not leave academia at this juncture.

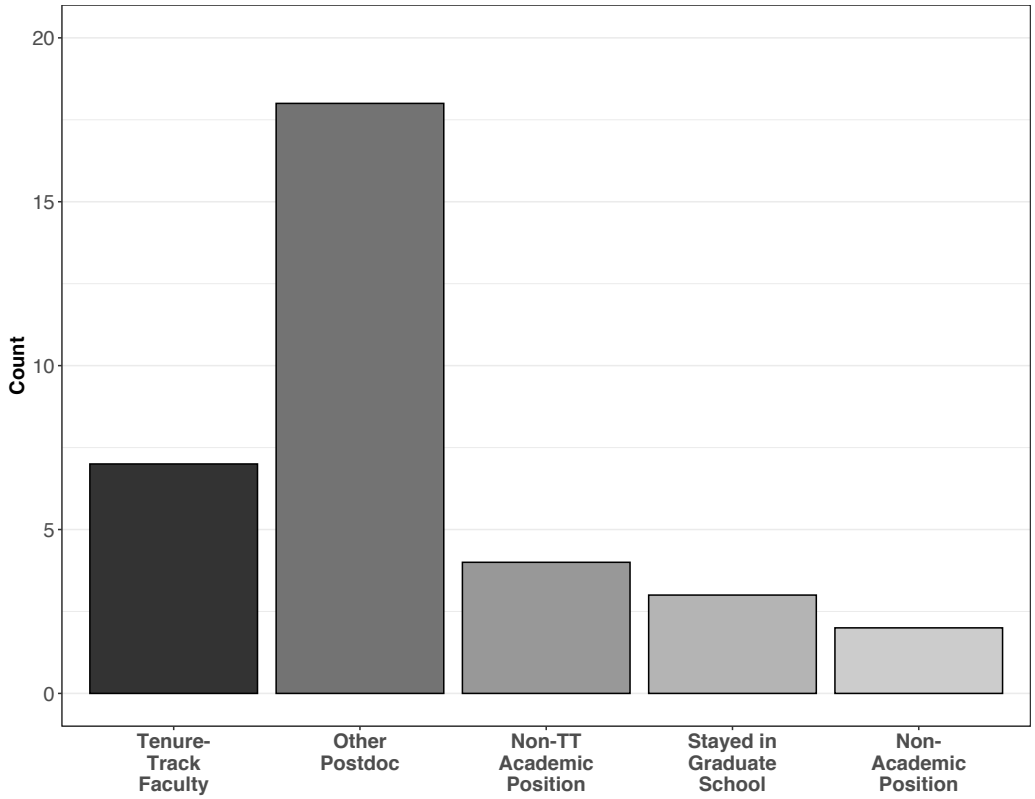


Figure 4.7: Finalists' positions during fellowship year

Additionally, 7 of the 34 finalists (about 20%) started a tenure-track job at a different institution. Most of these tenure-track faculty positions (5 out of 7) were at

[16]From what I could observe, at least one third (6 out of 18) of these alternative postdoc fellowship programs or funding sources were also explicitly diversity-oriented.

institutions with doctoral programs and very high research activity. One was actually at a UC campus – the same campus the applicant applied to for a fellowship.

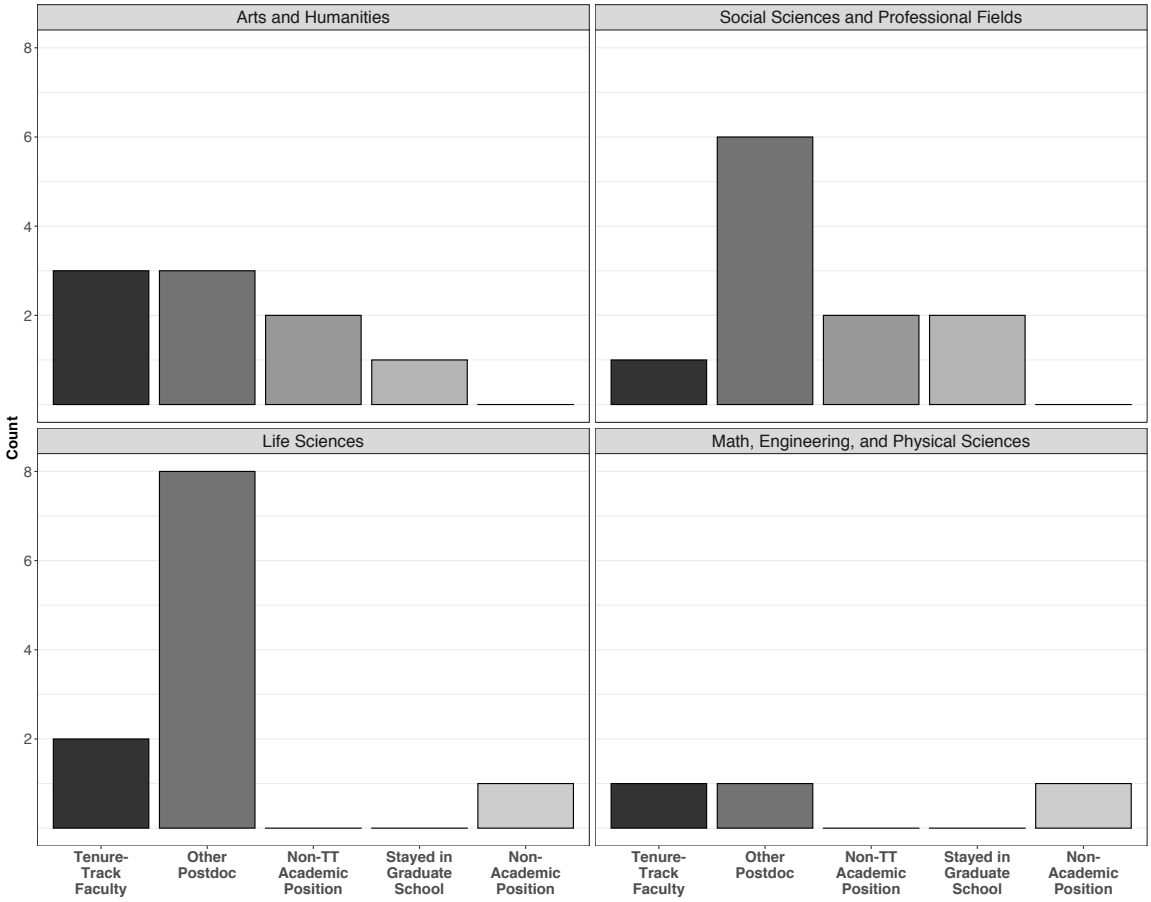


Figure 4.8: Finalists' positions during fellowship year by discipline

These alternative trajectories differ by discipline. Figure 4.8 shows that other postdoctoral fellowships or training positions were the most common pathway for finalists in the life sciences and social sciences. A solid majority of life scientists (8 out of 11 finalists) took another postdoctoral appointment, as did a little more than half of the social scientists (6 out of 11 finalists). By contrast, only one-third of finalists in the arts and humanities or in math, engineering, and physical sciences transitioned into other postdoctoral appointments. The lower rate of matriculation into a postdoctoral fellowship in the arts and humanities could be a reflection of the relative dearth of postdoctoral opportunities in those fields. However, lack of

postdoctoral opportunities are less likely to explain outcomes in math, engineering, and the physical sciences.

Another notable pattern in disciplinary differences is that some paths – like non-tenure-track academic positions and staying in graduate school for another year – are only taken by scholars in the arts and humanities or the social sciences and professional fields, while other paths – like going into industry – is only taken by scholars in traditional STEM fields. No finalists from life science, math, engineering, and physical science disciplines pursued non-tenure-track academic work or additional graduate training, and no finalists from the arts, humanities, social sciences and professional fields went into industry when they did not receive a UCPPFP fellowship.

4.5 Discussion

Seven years after their initial participation in the University of California President's Postdoctoral Fellowship Program, scholars from the 2014-2015 cohort are more likely to be in tenure-track faculty positions than finalists who were almost selected for the program. This is true for both men and women, and for members of both underrepresented and well-represented racial/ethnic groups. Fellows who became tenure-track faculty are more likely to be appointed at a UC campus than finalists who became tenure-track faculty, but about half of the fellows who entered the academy still ended up teaching in institutions outside the UC system.

Fellows' higher rates of entry into the academy are especially pronounced in the life sciences, where most of the fellows have gone on to faculty positions while most of the finalists are in non-academic positions. Fellows in the social sciences and in math, engineering, and physical sciences are only slightly more likely to become tenure-track faculty than finalists. Finally, because finalists in the arts and humanities

went on to tenure-track positions at such a high rate, their retention in the academy is comparable to that of the fellows' in those disciplines.

For the vast majority of finalists, not receiving one of the selective UCPPFP fellowships in 2014 did not push them out of the academy at that juncture. Most of the finalists ended up taking another postdoctoral research position in the 2014-2015 academic year, and this was especially true in the life sciences, where the postdoctoral research positions have proliferated. Notably, while finalists in the arts and humanities are less likely to take other postdocs, they – and to a lesser extent, finalists in the social sciences and professions – are more likely to take other types of non-permanent non-tenure-track academic positions, such as visiting professorships and lecturer positions. This suggests that some postdoctoral research positions in the natural sciences, and especially biomedical sciences, may be structurally equivalent to other forms of non-tenure-track academic labor and perhaps warrant more attention from those concerned about the adjunctification of the academy.

Seven of the 34 finalists (about 20%) went straight into tenure-track faculty positions in 2014-2015. Only one of these tenure-track positions was at an UC campus. Part of these dynamics may be explained by the program's consideration of who would benefit from the program – reviewer panels may decide to pass on an applicant because they think they are ready for a tenure-track job now, and would not need a postdoctoral fellowship to be competitive for an UC faculty position. (For example, in reviewer comments across all eight years of data, there are comments along the lines of "I would hire this applicant for a tenure-track position today".) However, these "over-qualified" applicants passed over for the postdoctoral program may not actually be selected in regular faculty searches (or may not accept a position if offered one). From an organizational perspective, this means that UC schools are potentially passing up on a fair amount of URM talent. This pattern, coupled with the fact that a slight majority of UCPPFP do not actually stay in the UC system for tenure-

track faculty positions, further illustrates the limited utility of the "leaky pipeline" metaphor and talent shortages for explaining racial and gender underrepresentation.

4.5.1 Limitations

The research design of this study enabled an unprecedented comparison of the career trajectories of diversity fellowship recipients and similar individuals who pursued alternative career paths, but it is not without limitations. First, the generalizability of this study to the broader academic and scientific workforce is limited. Not only are the finalists and fellows selected from a much larger group of applicants, the applicants themselves are probably not representative of the broader population of Ph.D. recipients. For example, one survey found that scientists with more publications, fellowships from federal agencies, and come from higher ranked Ph.D. programs are more likely to plan to do a postdoc in the first place (Sauermann and Roach 2016). Thus, the present case is really an examination of a small (and elite) slice of the larger academic and scientific workforce.

Moreover, this one year of data could be an anomaly from larger patterns. Repeating this study on the other years of data will be key. However, expanding the sample to cover multiple cohorts also brings challenges, as the amount of time between the initial fellowship appointment and scholars' current positions is different for every cohort; these methodological challenges warrant further consideration in future extensions.

The ability to interpret the study results causally is rooted in a few assumptions that would benefit from further examination. For example, the assumption that the differentiation of fellows from finalist is "as good as random" is key to making causal inferences from this study. Though we know that academic evaluation can be idiosyncratic and involve a large degree of randomness (Lamont 2009; Posselt 2016), more data about the selection process – such as the scores from on-site evaluation,

interviews with reviewers, and more detailed tracking information about the rankings of the alternates – would increase confidence that fellowship allocation really is quasi-random. I plan to examine and incorporate this data in the future as they become available.

A second complicating factor is that the "treatment" might begin prior to getting the award, as applicants have to contact potential faculty mentors before they apply. Especially in natural sciences, PIs and directors of training programs have incentives to identify and mentor potential postdocs, as UCPPPFP provides another source of funding for the junior research positions in their labs. Thus, applicants who are ultimately successful could be embedded in more pertinent scientific networks than finalists who are not chosen for the fellowship. Additional qualitative research about the networks of applicants might shed light on the role of social networks in obtaining one of these highly coveted fellowships.

Third, there is likely unobserved heterogeneity in the group of finalists: some of the applicants might have not been selected because they are seen as not needing the fellowship, while other applicants are not selected because their proposal and accomplishments are seen as not strong enough. Leveraging the more detailed reviewer comments, including comments from on-site review, may render this factor more visible.

Finally, the results of this study should be interpreted with nuanced consideration of what the ideal career outcomes are for the individual scholars. The organizational goals are clear: UCPPFP want the fellows to become faculty on UC campuses. But it's harder to judge if that is always the ideal outcome for the individual. For many scholars, non-academic careers can be more satisfying and better-paying. Scholars' dynamic career aspirations are not observed here – postdocs' career goals may change even during the course of the appointment – so the results should

not be interpreted to mean that people who did not get academic positions failed to achieve their goals.

4.5.2 Future research and extensions

If finalists who take other postdocs are less likely to become tenure-track faculty than UC President's and Chancellor's Fellows are, a major outstanding question is why. Two unique features of the UCPPFP program may offer insight. First, mentorship could be a key mechanism in this setting. In an ideal postdoctoral training scenario, the postdoc acquires new skills and receives mentorship from their principal investigator, which could help them achieve their long-term career goals (National Academy of Sciences et al. 2014). One study found that receiving influential career guidance from one's PI is a significant predictor of continued interest in academic career paths among postdocs in the biomedical sciences (Lambert et al. 2020). However, the reality is often far from the ideal (McConnell et al. 2018; National Academy of Sciences et al. 2014), and many postdocs do not receive such mentorship. Because the UCPPFP selection process also considers postdoctoral mentor fit during the evaluation process, UCPPFP fellows might receive better mentorship than scholars in other postdoctoral research positions.

A second mechanism worthy of future research is the role of monetary incentives. While many opportunity-enhancing diversity programs and policies offer monetary benefits to the individual recipients, it is rarer to offer explicit monetary incentives to the organization as well. Future research could use more years of data and the variation in the availability of the hiring incentives funds to potentially investigate the causal impact of the incentive. More broadly speaking, future research could examine how organizations view these types of "diversity monies" as a type of "special monies" (Zelizer 1989), and how this treatment impacts the rollout and outcomes of diversity initiatives.

Finally, this study compared the career trajectories of UCPPFP fellows and finalists, who are mostly women and/or underrepresented racial/ethnic minorities. However, this program is not the only channel through which women and URM scholars are hired at UC. Since UC is a public university system and personnel records are public, future research could also compare the career trajectories of UCPPFP fellows who eventually become faculty at UC campuses with women and URM faculty who did not get hired through UCPPFP.

Chapter 5

Conclusion

There is a tension in many modern organizational diversity initiatives. Are they policy holdovers from an earlier era of affirmative action, designed to help members of historically excluded groups who would not otherwise have access to the same educational and employment opportunities? Or are they selective honors and rewards systems, designed to give organizations more leverage to recruit top talent?

For social scientists, policy makers, and managers grappling with this tension, this dissertation offers at least three considerations. First, selective diversity programs often offer a bundle of benefits covering both material incentives and symbolic honors. These different components can have different effects. As Chapters 2 and 3 show, the purely reputational aspects of diversity-related honors and rewards do not always confer the esteem often associated with more generic honors. However, Chapter 4 shows that material benefits, like better mentorship or financial incentives, may be important mechanisms for policy effectiveness.

Second, this dissertation highlights the heterogeneity of diversity program beneficiaries, and shows that the effects of diversity initiatives are not experienced equally. For individuals who are presumed to be members of dominant groups, selective diversity-related awards and programs are recognized as legitimate honors.

However, diversity-related honors are devalued when they are given to members of historically excluded groups. Moreover, as the disciplinary differences in Chapter 5 shows, field-specific opportunity structures and mobility patterns also influences the ultimate role that diversity programs can play.

Finally, this dissertation emphasizes the analytical distinction between what diversity programs actually do and what people think diversity programs do. Over the past four decades, affirmative action has been de-fanged as it received legal scrutiny and lax federal enforcement. As a result, state and organizational actors have had resorted to using "the business case for diversity" to justify equal opportunity initiatives and use ambiguous criteria like "contributions to diversity" to implement diversity recruiting efforts. Yet it is not clear that members of the public, including professionals who play gatekeeping roles in organizations, understand these dynamics. For example, although preferential treatment by race is illegal and has not been practiced by organizations in decades, that is still the mental model that many people (and indeed, some researchers) associate with diversity policies.

There is still much to understand about contemporary organizational approaches to diversity. First, data on different types of diversity pipeline programs and who they impact is sorely lacking. A comprehensive typology of selective diversity initiatives would greatly benefit both researchers and practitioners. Another fruitful avenue for future research would involve greater attention to how managers and employees' folk understandings of diversity might impact policy implementation. Finally, while this dissertation focuses on linking meso-level organizational policies to micro-level outcomes, a more macro-approach that considers equilibrium effects would also be helpful. Selective diversity scholarships, fellowships, and pipeline programs create more opportunities in a broader ecosystem – understanding how these distinct openings fit in with existing career pathways and opportunity structures will be crucial to their success.

Appendix A

Supplemental Materials for Chapter 2

A.1 University Reputations and Scholarship Interpretations

In December 2019, I recruited 151 U.S.-based adults with hiring experience from Prolific Academic for a survey. I presented each respondent with the names of eight U.S.-based universities (the order was randomized) and asked them to rate how selective they thought the school was on a 5-point scale and how much they agreed with the statement that "Graduates of [X University] generally make good employees" on a 7-point scale. Five of the schools were universities that I was considering using for the study; three of the schools were used as "anchors" to assess the comparisons. One of the three anchor schools was much more selective than the five schools of interest and ranked extremely highly by U.S. News and World Report; the other two anchor schools were much less selective than the five schools of interest and ranked much lower by U.S. News and World Report. I picked the three universities that were

perceived most similarly in terms of selectivity and graduate employment (Figure A.1).

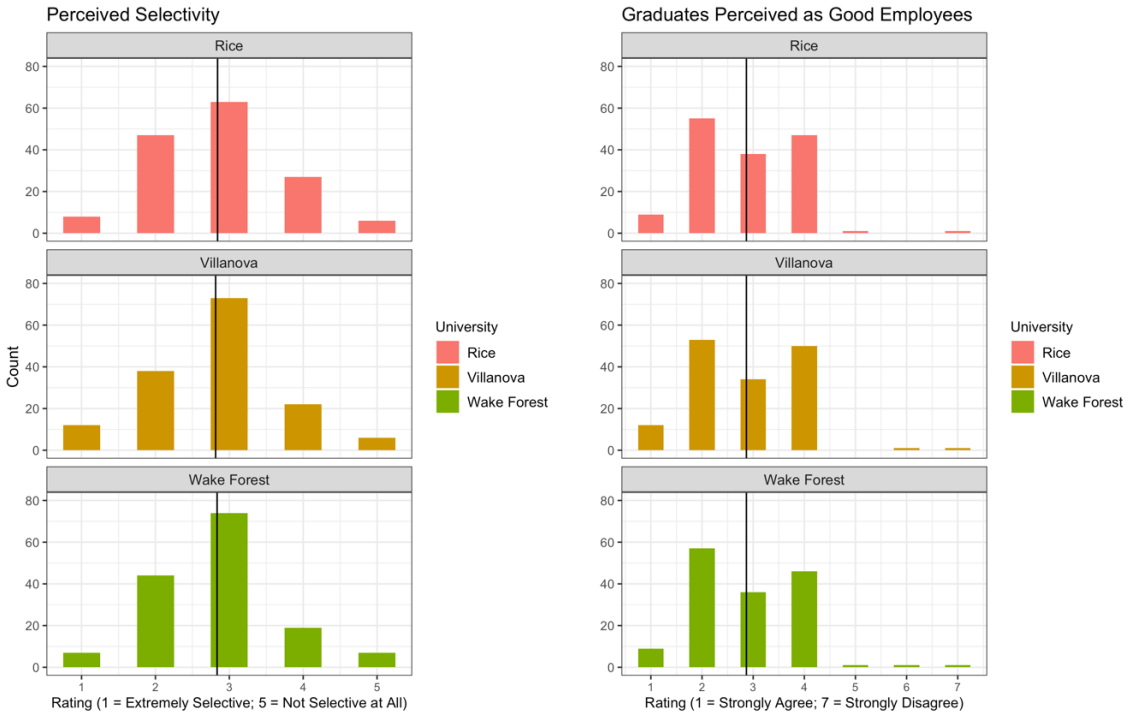


Figure A.1: University reputations; vertical lines show average ratings (N = 151 respondents with hiring experience recruited on Prolific Academic)

Then, each respondent saw the names and descriptions of a diversity scholarship associated with one of five universities of interest and a non-diversity scholarship associated with another one of the five universities (the order was randomized). For each scholarship, I asked respondents what criteria they thought the scholarship was based on by presenting them with a list of factors and asking them to check all that they though applied. The criteria included 1) academic achievements, like having good grades or standardized test scores, 2)Extracurricular achievements, like winning prizes for athletics or arts, 3) Holding leadership positions, 4) Doing a lot of community service, 5) Being an underrepresented minority of some sort, 6) Something else, and 7) Not sure (Figure A.2). Then I asked respondents to rate on a 1-7 scale how

likely they thought it was that someone who received the scholarship benefited from affirmative action (Figure A.3).

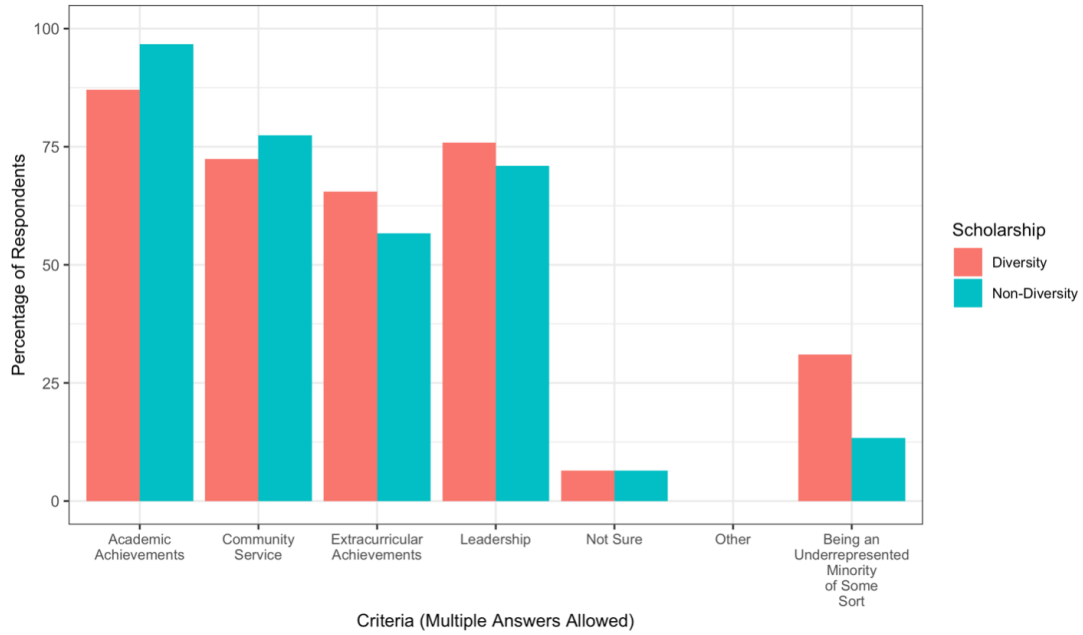


Figure A.2: Perceptions of scholarship criteria (N = 151 respondents with hiring experience recruited on Prolific Academic)

The average age of the respondent sample was 42 years. About 62% of the sample were men, approximately 81% were White, 8% were Black, 3% were Hispanic, 5% were Asian, 1% were American Indian/Alaskan Native, and 1% self-identified as some other race. About 30% had less than a Bachelor's degree (including those with associate degrees or completed some college), 44% had a Bachelor's degree, and 26% had an advanced degree. 77% of the respondents were working as salaried employees, 13% were self-employed, and the rest were retired or not working. The most common occupations were Engineering or IT (23%), Operations (16%), Customer service (15%), and Sale/Marketing/Business Development (11%). 58% of the respondents self-identified as Liberal, 16% as neither Liberal nor Conservative, and 27% as Conservative. About 21% were located in the Northeast, 19% in the Midwest, 35% in the South, and 25% in the West.

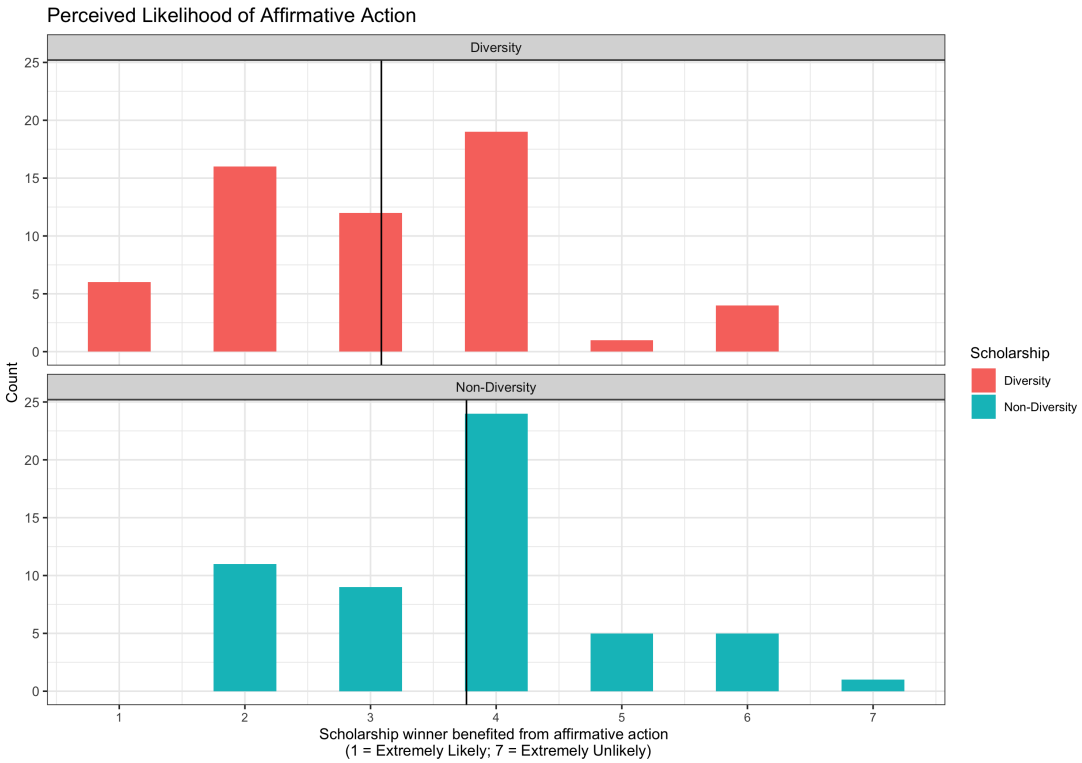


Figure A.3: Perceived likelihood of affirmative action (N = 151 respondents with hiring experience recruited on Prolific Academic)

A.2 Racial and Class Cues of Names

In April 2020, I recruited another 200 U.S.-based adults with hiring experience from Prolific Academic to assess the demographic cues of names. I pre-tested 28 male names (14 Black, 14 White). First names were drawn from NY state birth records 1994-2012 (college class of 2020 would fall in this range), as published in Gaddis (2017), and last names were taken from the 2000 Census (based on population-level occurrence of last names). All first names had at least 50 births per year associated with them. I picked the names that performed better than average in the Gaddis paper. I also avoided Judeo-Christian and Muslim origin names and names associated with mothers who were in the lowest quartile of educational attainment.

Each respondent saw a random subset of eight names and rated all eight names. Half of the sample were randomly assigned to always rate perceived socioeco-

nomic class before rating perceived race; the other half always assessed perceived race before perceived class. For each name, respondents were asked: 1) What gender do you most associate with the name [ ]? 2) What race(s) or ethnicity do you associate with the name [ ]? (Select all that applies) 3) What socioeconomic class do you most associate with the name [ ]?

All the names were highly likely to be perceived as male. I picked two names that were highly likely to be associated exclusively with being Black and middle-class, and two names that were highly likely to be associated exclusively with being White and middle-class (Figure A.4).

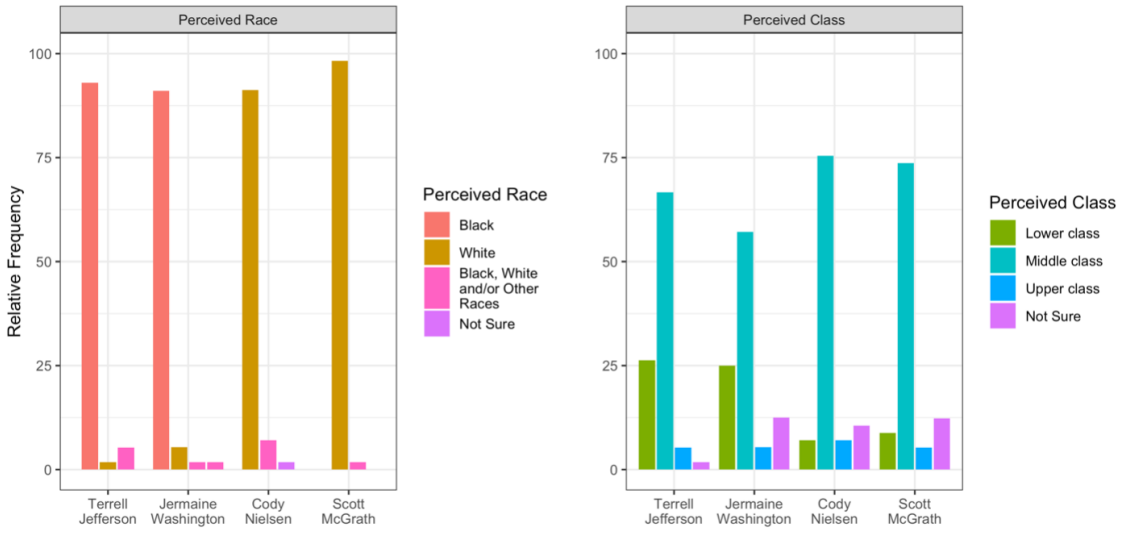


Figure A.4: Perceived race (left) and socioeconomic class (right) of names (N = 200 respondents with hiring experience recruited on Prolific Academic)

The average age in this sample was 41 years old. The sample was 48% male. This time, I allowed respondents to self-identify with multiple races. 82% self-identified as White, 6% as Black, 10% as Hispanic, 10% as Asian, and 2% as American Indian/Alaskan Native. About 34% had less than a Bachelor's Degree, 42% had a Bachelor's Degree, and 24% had an advanced degree. About 58% were working as salaried employees, 8% were temporarily not working (this was at the beginning of the COVID-19 pandemic, when employees in several sectors were furloughed), 13% were

self-employed, the rest were retired, or not working. The most common occupations were Customer service (19%), Engineering or IT (14%), Sale/Marketing/Business Development (13%), and Operations (12%). About 15% of the sample came from the Northeast, 22% from the Midwest, 40% from the South, and 23% from the West.

Appendix B

Supplemental Materials for Chapter 3

B.1 Causal Mediation Analyses of Order Effects

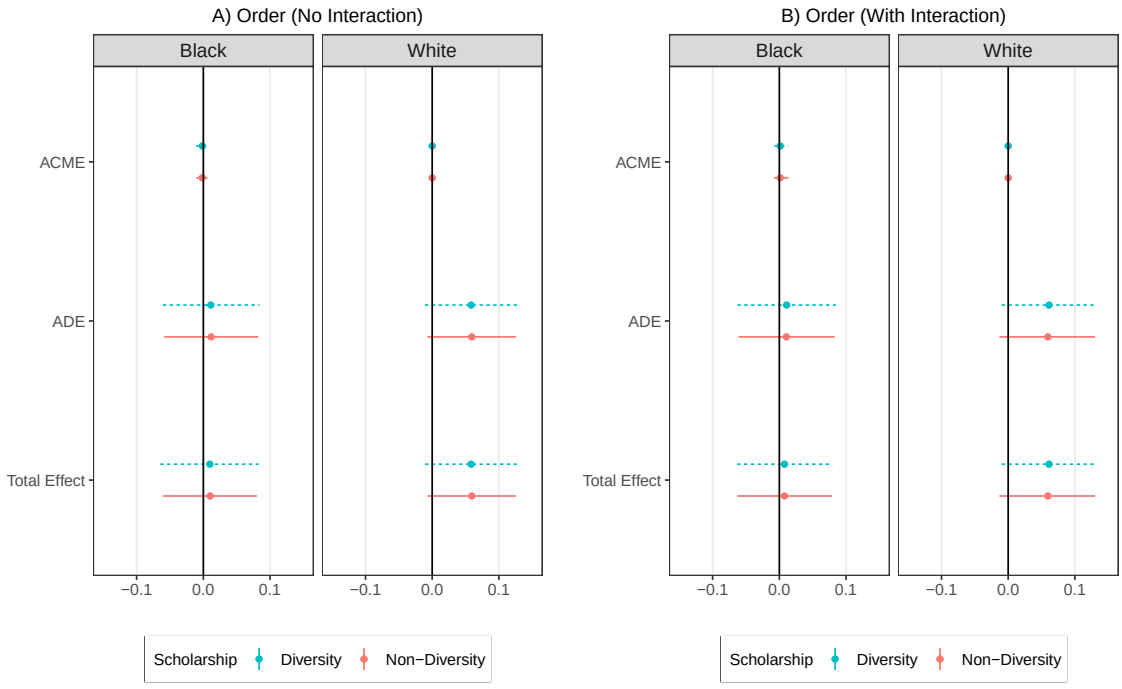


Figure B.1

B.2 Additional Causal Mediation Analyses of Perceived Race and Sexual Orientation

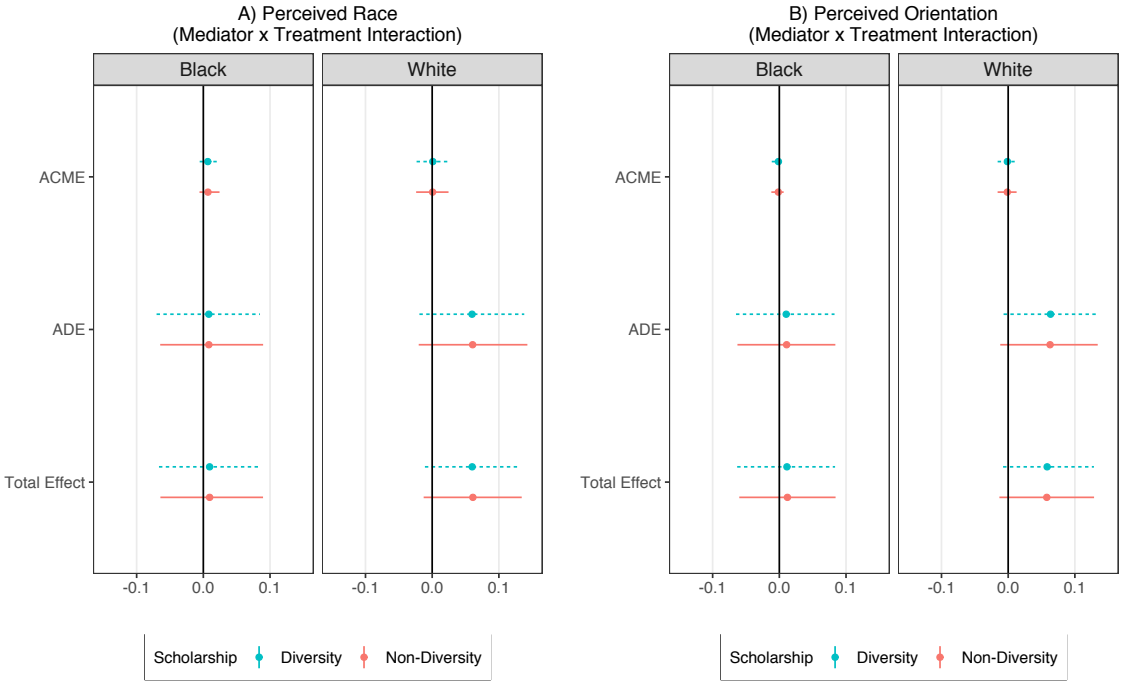


Figure B.2

Appendix C

Supplemental Materials for Chapter 4

C.1 Disciplines and Interdisciplinarity in the UCPPFP Evaluation Process

Applicants indicate what discipline (Arts and Humanities, Social Sciences and Professional Fields, Life Sciences, or Math, Engineering, and Physical Sciences) and sub-discipline (e.g. English, Political Science, Chemistry, etc.) their proposal is based in, and reviewers indicate their primary and secondary disciplines. Most sub-disciplines tend to be clearly associated with one disciplinary group, but a few sub-disciplines were inconsistently categorized. The most commonly inconsistently categorized sub-discipline is Ethnic Studies (14 applicants categorized it as arts and humanities and 14 applicants classified it as a social science). History was mostly categorized as arts and humanities, but a handful of historians categorized their proposals as social sciences. Most sub-disciplines that were inconsistently categorized did not receive many applications overall (usually between two and four); these include architecture,

linguistics, American studies, and environmental science. Table C.1 shows the five most common sub-disciplines for each disciplinary group.

Table C.1: Most Common Sub-Disciplines (2014-2015 Cohort)

Sub-Discipline	Count	%
Arts and Humanities		
English	32	20.6
History	27	17.4
Ethnic Studies	14	9.0
Languages/Literature	14	9.0
Women and Gender Studies	12	7.7
Social Sciences and Professional Fields		
Education	31	16.2
Sociology	27	14.1
Anthropology	19	9.9
Psychology	19	9.9
Ethnic Studies	13	6.8
Life Sciences		
Neurosciences	18	16.7
Biology	12	11.1
Ecology	12	11.1
Cell Biology	7	6.5
Marine Biology	6	5.6
Math, Engineering, and Physical Sciences		
Chemistry	11	16.2
Astronomy	8	11.8
Physics	8	11.8
Biomedical Engineering	7	10.3
Earth and Planetary Science	5	7.4

Most applicants' proposals were in the same discipline their Ph.D. is in. Table C.2 shows the frequencies of Ph.D. discipline and proposal discipline mismatch. The most common cases of discipline switching were arts and humanities Ph.D.s applying with social science proposals, and social science Ph.D.s applying with arts and humanities proposals. Switching between the life sciences and math, engineering, and physical sciences also occurred.

Table C.2: Applicants Reporting Different Disciplines for Ph.D. and Postdoc Proposal (2014-2015 Cohort)

Discipline for Ph.D.	Research Proposal Discipline			
	Arts and Humanities	Social Sciences	Life Sciences	Math/Eng/Phy Sci
Arts and Humanities	-	9	-	-
Social Sciences	6	-	3	-
Life Sciences	-	1	-	4
Math/Eng/Phy Sci	-	1	5	-

Because some proposals drew on expertise from multiple disciplines and sub-disciplines, they were sometimes assigned to reviewers from different disciplines during off-site review. In 2014-2015, 86 applications (about 16%) of were assigned to two reviewers from different disciplines during off-site review. Two reviewers may also share the same discipline but come from different subdisciplines (e.g. economics and political science), but I do not have data on reviewers' sub-disciplines.

Table C.3 shows counts of disciplinary homophily and heterophily among off-site reviewer pairs. The counts on the diagonal axis show the number of applications that were reviewed by two reviewers from the same discipline. The most common interdisciplinary pair was a reviewer from the arts and humanities and a reviewer from the social sciences or professional fields. In fact, all but four of the interdisciplinary reviewer pairs involved a reviewer from the social sciences or professional fields. Interdisciplinary reviewer pairs were no more or less likely to give out qualifying scores (A-A, A-B, or A-C) than same-discipline reviewer pairs (about 54-55% for both). Same-discipline reviewer pairs are slightly more likely to give A-A scores.

Table C.3: Reviewers' Disciplines

Reviewer 1	Reviewer 2			
	Arts and Humanities	Social Sciences	Life Sciences	Math/Eng/Phy Sci
Arts and Humanities	129	39	1	-
Social Sciences	10	133	10	4
Life Sciences	-	5	108	2
Math/Eng/Phy Sci	-	14	1	64

Bibliography

Alberts, B., Kirschner, M. W., Tilghman, S., and Varmus, H. (2014). Rescuing US biomedical research from its systemic flaws. *Proceedings of the National Academy of Sciences*, 111(16):5773–5777.

Berrey, E. (2015). *The Enigma of Diversity*. University of Chicago Press.

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *American Economic Review*, 94(4):991–1013.

Bobo, L. and Kluegel, J. R. (1993). Opposition to Race-Targeting: Self-Interest, Stratification Ideology, or Racial Attitudes? *American Sociological Review*, 58(4):443–464.

Bol, T., de Vaan, M., and van de Rijt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19):4887–4890.

Bollen, K. A. (1989). *Structural Equations with Latent Variables*. Wiley-Interscience, New York, 1st edition edition.

Brown, R. P., Charnsangavej, T., Keough, K. A., Newman, M. L., and Rentfrow, P. J. (2000). Putting the "affirm" into affirmative action: preferential selection and academic performance. *Journal of Personality and Social Psychology*, 79(5):736–747.

Connelly, B. L., Certo, S. T., Ireland, R. D., and Reutzel, C. R. (2011). Signaling Theory: A Review and Assessment. *Journal of Management*, 37(1):39–67.

Cuddy, A. J. C., Fiske, S. T., and Glick, P. (2008). Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map. In *Advances in Experimental Social Psychology*, volume 40, pages 61–149. Academic Press.

Dobbin, F. (2009). *Inventing Equal Opportunity*. Princeton University Press, Princeton, UNITED STATES.

Eagly, A. H. and Diekman, A. B. (2005). What is the Problem? Prejudice as an Attitude-in-Context. In *On the nature of prejudice: Fifty years after Allport*, pages 19–35. Blackwell Publishing, Malden.

Eagly, A. H. and Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3):573–598.

Eberhardt, J. L. and Fiske, S. T. (1994). Affirmative Action in Theory and Practice: issues of Power, Ambiguity, and Gender Versus Race. *Basic and Applied Social Psychology*, 15(1-2):201–220.

Edelman, L., Fuller, S., and Mara-Drita, I. (2001). Diversity Rhetoric and the Managerialization of Law. *American Journal of Sociology*, 106(6):1589–1641.

Embrick, D. G. (2011). The Diversity Ideology in the Business World: A New Oppression for a New Age:. *Critical Sociology*.

Finkelstein, M. J., Conley, V. M., and Schuster, J. H. (2016). Taking the Measure of Faculty Diversity. Technical report, TIAA Institute.

Fiske, S. T., Cuddy, A. J. C., Glick, P., and Xu, J. (2002). A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived

status and competition. *Journal of Personality and Social Psychology*, 82(6):878–902.

Foley, M. and Williamson, S. (2019). Managerial Perspectives on Implicit Bias, Affirmative Action, and Merit. *Public Administration Review*, 79(1):35–45.

Gaddis, S. M. (2015). Discrimination in the Credential Society: An Audit Study of Race and College Selectivity in the Labor Market. *Social Forces*, 93(4):1451–1479.

Gaddis, S. M. (2017). How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies. *Sociological Science*, 4:469–489.

Gaddis, S. M. (2019). Understanding the "How" and "Why" Aspects of Racial-Ethnic Discrimination: A Multimethod Approach to Audit Studies. *Sociology of Race and Ethnicity*, 5(4):443–455.

Gardner, D. M. and Ryan, A. M. (2020). What's in it for you? Demographics and self-interest perceptions in diversity promotion. *Journal of Applied Psychology*, 105(9):1062–1072.

Germano, A. L. and Cheryan, S. (2021). Choosing a Scholarship - Award Status.

Germano, A. L., Ziegler, S. A., Banham, L., and Cheryan, S. (2021). Do Diversity Awards Discourage Applicants From Marginalized Groups From Pursuing More Lucrative Opportunities? *Psychological Science*.

Gibbs, K. (2014). Beyond "The Pipeline": Reframing Science's Diversity Challenge. *Scientific American*.

Gibbs, Jr, K. D., Basson, J., Xierali, I. M., and Broniatowski, D. A. (2016). Decoupling of the minority PhD talent pool and assistant professor hiring in medical school basic science departments in the US. *eLife*, 5:e21393.

Gilbert, J. A. and Stead, B. A. (1999). Stigmatization Revisited: Does Diversity Management Make a Difference in Applicant Success? *Group & Organization Management*, 24(2):239–256.

Goulden, M., Mason, M. A., and Frasch, K. (2011). Keeping Women in the Science Pipeline. *The ANNALS of the American Academy of Political and Social Science*, 638(1):141–162.

Griffin, K. (2019). Redoubling Our Efforts: How Institutions Can Affect Faculty Diversity. *Race and Ethnicity in Higher Education*.

Hare, J. D. (2016). Best Practices for Diversifying the UC Faculty. https://ppfp.ucop.edu/info/documents/Best%20Practices%20for%20Diversifying%20the%20UC%20Faculty.2.2.16.pdf.

Harper, S. and Reskin, B. (2005). Affirmative Action at School and on the Job. *Annual Review of Sociology*, 31(1):357–379.

Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis, Second Edition: A Regression-Based Approach*. The Guilford Press, New York, second edition edition.

Heilman, M. E., Block, C. J., and Lucas, J. A. (1992). Presumed incompetent? Stigmatization and affirmative action efforts. *Journal of Applied Psychology*, 77(4):536–544.

Heilman, M. E., Block, C. J., and Stathatos, P. (1997). The Affirmative Action Stigma of Incompetence: Effects of Performance Information Ambiguity. *The Academy of Management Journal*, 40(3):603–625.

Heilman, M. E., McCullough, W. F., and Gilbert, D. (1996). The other side of affirmative action: Reactions of nonbeneficiaries to sex-based preferential selection. *Journal of Applied Psychology*, 81(4):346–357.

Hekman, D. R., Johnson, S. K., Foo, M.-D., and Yang, W. (2016). Does Diversity-Valuing Behavior Result in Diminished Performance Ratings for Non-White and Female Leaders? *Academy of Management Journal*, 60(2):771–797.

Hirschman, D. and Berrey, E. (2017). The Partial Deinstitutionalization of Affirmative Action in U.S. Higher Education, 1988 to 2014. *Sociological Science*, 4:449–468.

Holzer, H. and Neumark, D. (2000). Assessing Affirmative Action. *Journal of Economic Literature*, 38(3):483–568.

Imai, K., Keele, L., and Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4):309–334.

Kalev, A., Dobbin, F., and Kelly, E. (2006). Best Practices or Best Guesses? Assessing the Efficacy of Corporate Affirmative Action and Diversity Policies. *American Sociological Review*, 71(4):589–617.

Kang, S. K., DeCelles, K. A., Tilcsik, A., and Jun, S. (2016). Whitened Résumés: Race and Self-Presentation in the Labor Market. *Administrative Science Quarterly*, 61(3):469–502.

Kantrowitz, M. (2011). The Distribution of Grants and Scholarships by Race. Technical report, Fastweb.com and FinAid.org.

Kellough, J. E., Coleman Selden, S., and Legge, J. S. (1997). Affirmative Action Under Fire: The Current Controversy and the Potential for State Policy Retrenchment*. *Review of Public Personnel Administration*, 17(4):52–74.

Kessler, J. B., Low, C., and Sullivan, C. D. (2019). Incentivized Resume Rating: Eliciting Employer Preferences without Deception. *American Economic Review*, 109(11):3713–3744.

Kirkland, A. and Hansen, B. B. (2011). "How Do I Bring Diversity?" Race and Class in the College Admissions Essay. *Law & Society Review*, 45(1):103–138.

Lambert, W. M., Wells, M. T., Cipriano, M. F., Sneva, J. N., Morris, J. A., and Golightly, L. M. (2020). Career choices of underrepresented and female postdocs in the biomedical sciences. *eLife*, 9:e48774.

Lamont, M. (2009). *How Professors Think*. Harvard University Press.

Lawson, M. A. (2018). University of California President's Postdoctoral Fellowship Program Annual Report 2017-2018. Technical report, University of California.

Lerma, V., Hamilton, L. T., and Nielsen, K. (2020). Racialized Equity Labor, University Appropriation and Student Resistance. *Social Problems*.

Leslie, L. M., Mayer, D. M., and Kravitz, D. A. (2014). The Stigma of Affirmative Action: A Stereotyping-Based Theory and Meta-Analytic Test of the Consequences for Performance. *Academy of Management Journal*, 57(4):964–989.

Mai, Q. D. (2021). Unclear Signals, Uncertain Prospects: The Labor Market Consequences of Freelancing in the New Economy. *Social Forces*, 99(3):895–920.

Maio, G. R. and Esses, V. M. (1998). The Social Consequences of Affirmative Action: Deleterious Effects on Perceptions of Groups. *Personality and Social Psychology Bulletin*, 24(1):65–74.

Major, B., Feinstein, J., and Crocker, J. (1994). Attributional ambiguity of affirmative action. *Basic and Applied Social Psychology*, 15(1-2):113–141.

McConnell, S. C., Westerman, E. L., Pierre, J. F., Heckler, E. J., and Schwartz, N. B. (2018). United States National Postdoc Survey results and the interaction of gender, career choice and mentor impact. *eLife*, 7:e40189.

Merton, R. K. (1968). The Matthew Effect in Science. *Science*, 159(3810):56–63.

Miller, D. I. and Wai, J. (2015). The bachelor's to Ph.D. STEM pipeline no longer leaks more women than men: a 30-year analysis. *Frontiers in Psychology*, 6:37.

Napolitano, J. (2017). Letter to UC Chancellors.

National Academy of Sciences, of Engineering, N. A., Institute of Medicine, Sciences, N. A. o., Engineering, N. A. o., and Medicine, I. o. (2014). *The Postdoctoral Experience Revisited*. National Academies Press, Washington, DC.

Norman, J. (2019). Americans' Support for Affirmative Action Programs Rises. *Gallup*.

Nunley, J. M., Pugh, A., Romero, N., and Seals, R. A. (2015). Racial Discrimination in the Labor Market for Recent College Graduates: Evidence from a Field Experiment. *The B.E. Journal of Economic Analysis & Policy*, 15(3):1093–1125.

Pager, D. and Karafin, D. (2009). Bayesian Bigot? Statistical Discrimination, Stereotypes, and Employer Decision Making. *The ANNALS of the American Academy of Political and Social Science*, 621(1):70–93.

Pager, D. and Quillian, L. (2005). Walking the Talk? What Employers Say Versus What They Do. *American Sociological Review*, 70(3):355–380.

Pedulla, D. S. (2016). Penalized or Protected? Gender and the Consequences of Nonstandard and Mismatched Employment Histories. *American Sociological Review*, 81(2):262–289.

Poon, O., Squire, D., Kodama, C., Byrd, A., Chan, J., Manzano, L., Furr, S., and Bishundat, D. (2016). A Critical Review of the Model Minority Myth in Selected Literature on Asian Americans and Pacific Islanders in Higher Education. *Review of Educational Research*, 86(2):469–502.

Portocarrero, S. (2020). Qué Vergüenza: How an Organization Led National Scholarship Recipients to Become Ashamed of What Had Once Made Them Most Proud. *Academy of Management Proceedings*, 2020(1):18017.

Posselt, J. R. (2016). *Inside Graduate Admissions*. Harvard University Press.

Powell, K. (2015). The future of the postdoc. *Nature*, 520(7546):144–147.

Quadlin, N. (2018). The Mark of a Woman's Record: Gender and Academic Performance in Hiring. *American Sociological Review*, 83(2):331–360.

Reskin, B. F. (1998). *The Realities of Affirmative Action in Employment*. The American Sociological Association, Washington, DC.

Rivera, L. A. (2012). Diversity within Reach: Recruitment versus Hiring in Elite Firms. *The ANNALS of the American Academy of Political and Social Science*, 639(1):71–90.

Roach, M. and Sauermann, H. (2017). The declining interest in an academic career. *PLOS ONE*, 12(9):e0184130.

Rucker, D. D., Preacher, K. J., Tormala, Z. L., and Petty, R. E. (2011). Mediation Analysis in Social Psychology: Current Practices and New Recommendations: Mediation Analysis in Social Psychology. *Social and Personality Psychology Compass*, 5(6):359–371.

Saguy, T., Fernández, S., Branscombe, N. R., and Shany, A. (2020). Justice Agents: Discriminated Group Members Are Perceived to be Highly Committed to Social Justice. *Personality and Social Psychology Bulletin*, 46(1):155–167.

Sauermann, H. and Roach, M. (2016). Why pursue the postdoc path? *Science*, 352(6286):663–664.

Scarborough, W. J., Lambouths, D. L., and Holbrook, A. L. (2019). Support of workplace diversity policies: The role of race, gender, and beliefs about inequality. *Social Science Research*, 79:194–210.

Scully, M. A. (2000). Meritocracy and the Legitimation of Inequality in Internal Labor Markets and Beyond. In Leana, C. R. and Rousseau, D. M., editors, *Relational Wealth: The Advantages of Stability in a Changing Economy*, pages 199–214. Oxford University Press.

Sen, M. and Wasow, O. (2016). Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics. *Annual Review of Political Science*, 19(1):499–522.

Sniderman, P. M. and Piazza, T. (1995). *The Scar of Race*. Harvard University Press.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, 87(3):355–374.

Spence, M. (2002). Signaling in Retrospect and the Informational Structure of Markets. *American Economic Review*, 92(3):434–459.

Stephan, P. (2012). Perverse incentives. *Nature*, 484(7392):29–31.

Taylor, M. C. (1994). Impact of Affirmative Action on Beneficiary Groups: Evidence From the 1990 General Social Survey. *Basic and Applied Social Psychology*, 15(1-2):143–178.

Thomas, D. and Horowitz, J. M. (2020). Support for Black Lives Matter has decreased since June but remains strong among Black Americans.

Thornhill, T. (2019). We Want Black Students, Just Not You: How White Admissions Counselors Screen Black Prospective Students. *Sociology of Race and Ethnicity*, 5(4):456–470.

Vuolo, M., Uggen, C., and Lageson, S. (2016). Statistical Power in Experimental Audit Studies: Cautions and Calculations for Matched Tests With Nominal Outcomes. *Sociological Methods & Research*, 45(2):260–303.

Weisshaar, K. (2018). From Opt Out to Blocked Out: The Challenges for Labor Market Re-entry after Family-Related Employment Lapses. *American Sociological Review*, 83(1):34–60.

Windsor, L. C. and Crawford, K. F. (2021). Academe Isn't a Leaky Pipeline. *The Chronicle of Higher Education*.

Wingfield, A. H. and Alston, R. S. (2014). Maintaining Hierarchies in Predominantly White Organizations: A Theory of Racial Tasks. *American Behavioral Scientist*, 58(2):274–287.

Woolston, C. (2020). Postdoc survey reveals disenchantment with working life. *Nature*, 587(7834):505–508.

Xue, Y. and Larson, R. C. (2015). STEM crisis or STEM surplus? Yes and yes. *Monthly Labor Review*.

Zelizer, V. A. (1989). The Social Meaning of Money: "Special Monies". *American Journal of Sociology*, 95(2):342–377.