NBER WORKING PAPER SERIES

DISCRIMINATION AT THE INTERSECTION OF AGE, RACE, AND GENDER: EVIDENCE FROM A LAB-IN-THE-FIELD EXPERIMENT

Joanna N. Lahey
Douglas R. Oxley

Working Paper 25357
http://www.nber.org/papers/w25357

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2018

We thank the Alfred P. Sloan Foundation for funding this research Grant #B2012-23. We thank Gerianne Alexander for use of her lab and other invaluable help. We also thank all of the undergraduate and graduate assistants who helped make this study successful. It was a joy to work with Molly Beck, Savannah Collins, Luke Franz, Stephanie Leal, Laura Lombardo, Joel Mendez, Megan Mumford, Matt Sanchez, Janet Saenz, and Ashley Yaugher. We especially wish to thank Ryan Beasley for volunteering his time and his technical expertise for several aspects of the study and David Figlio for names. Thanks for feedback from participants at APPAM, ASSA, Auburn University, University of California at Davis, University of California at Santa Cruz, University of Chicago, NBER Cohort Studies, NBER Summer Institute on Aging, University of North Carolina, Santa Clara University, SIEPR Working Longer, Syracuse University, Texas A&M University, University of Toronto, and Tulane and from helpful conversations with Lorenzo Casaburi, Kerwin Charles, Kathleen Christensen, Jeff Clemens, Joshua Gottlieb, Damon Jones, Matt Notowidigdo, Paul Oyer, Orie Shelef, Ebonya Washington, and John Yinger. All errors and opinions are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w25357.ack

Discrimination at the Intersection of Age, Race, and Gender: Evidence from a Lab-in-the-field Experiment
Joanna N. Lahey and Douglas R. Oxley
NBER Working Paper No. 25357
December 2018
JEL No. J14,J15,J23,J7,J70,J71

ABSTRACT

We use a laboratory experiment with randomized resumes and eyetracking to explore the effects of race on employment discrimination over the lifecycle. We show race discrimination against prime-age black job applicants that diminishes into middle age before re-emerging for older applicants. Screeners mechanically process black and white resumes similarly, but spend less time on younger black resumes, suggesting they use negative heuristics or taste-based discrimination. Screeners demonstrate levels-based statistical discrimination, believing that younger black applicants have worse computer skills and more gaps in their job histories. We find no evidence that screeners believe black applicants have worse previous experience. Screeners demonstrate variance-based statistical discrimination against black applicants of all ages, suggesting that screeners perceive the stronger history signals for white applicants, with this type of discrimination disproportionately affecting older applicants. We find suggestive evidence that the signal sent by high school attended is weaker for younger black applicants compared to younger white applicants, and we find no evidence that the signal strength of the applicant's address varies by race. Evidence from the CPS and an additional study supports the external validity of our experiment, particularly for female job applicants. Results are robust to different controls and specification choices.

Joanna N. Lahey
The Bush School
Texas A&M University
Mailstop 4220
College Station, TX 77843
and NBER
jlahey@nber.org

Douglas R. Oxley
4102 Amber Trace Ct
Sugarland, TX 77479
dougoxley@gmail.com

I. Introduction

Research on race discrimination[2] has primarily focused specifically on younger ages or has pooled together all individuals of working age (see Lang and Lehmann 2012 for an extensive literature review). Far less has been written about how labor market differences by race change across the life-cycle. However, people have full working lives; with the average person in the NLSY79 reporting ~12 jobs by age 50, compared to ~8 by age 30 (authors' calculations), it is clear that not all job seekers are young. In addition, the average difference in labor market outcomes between black and white people has changed over time both by age and by cohort.[3]

The idea that demographic characteristics interact to produce different outcomes is termed "intersectionality" and was coined in the legal context by Crenshaw (1989). Audit studies have shown racial discrimination in terms of who is offered a job interview in resume audit studies, at least for young entry-level applicants (e.g. Bertrand and Mullainathan 2004, Nunley et al. 2015; see also Rich 2014 and Quillian et al. 2017 for literature reviews). Less is known about how interview outcomes for race interact with other demographic characteristics such as age, ethnicity, or socioeconomic status. Much of the focus on intersectional labor market outcomes by race has been looking at race*gender, and these studies have found mixed results on whether discrimination by race is worse for men or women (e.g. Bertrand and Mullainathan 2004, Nunley et al. 2015). Despite extensive reading across several social sciences, we have found very little that looks at age*race intersectionality in employment outcomes.[4] One exception is a recent audit study in the UK (Drydakis et al. 2018) that compared black and white applicants to low skilled private sector jobs who were either age 22 or age 50 and found that the black 50 year olds were disadvantaged

[2] The term "discrimination" will be used throughout the paper to refer to "differential treatment by demographic characteristic(s)." It does not refer solely animus or taste-based discrimination.

[3] Appendix Figure 1 provides cohort information for black-white differences across age using ACS/Census data for ln(Income) for all Americans and Lahey (2018) provides such differences for black and white women for several outcomes and education levels.

[4] Indeed, a recent meta-analysis by Jones et al. (2017) laments this lack of intersectional research within the field of psychology. Intersectionality with age, race, and ethnicity is, however, studied in more depth in conjunction with health outcomes (e.g. Firebaugh et al. 2014, Kirby and Kaneda 2010, Masters et al. 2014, Pais 2014). Additionally, age is correlated with experience and Bound and Freeman (1992) and Oyer and Schaefer (2002) study differential effects of experience by race.

compared to black 22 year olds or white 50 year olds. Our experiment instead focuses on ages after the bulk of education decisions have been made and applicants have had a chance to build up work experience, allowing us to better separate age from experience. We also look at a continuous distribution of ages to allow for a full picture of how age and race interact over the adult working lifecycle.

Little is also known about the mechanics of hiring discrimination. If screeners see a black-sounding name or an early date of high school graduation and immediately move onto the next resume, there is not much that potential job seekers can change other than their names. If, instead, screeners view the resume more intensively for items that could contradict negative stereotypes, then that would indicate that black job seekers could mitigate the effects of those negative stereotypes by including these items on their resumes. Depending on how screeners view minority compared to non-minority resumes, employers who wish to reduce discrimination in hiring could implement structural changes asking all applicants to provide specific information or forcing screeners to view all parts of all resumes.

In addition, it is not known how screeners mentally process resumes. Heuristics (also known as rules-of-thumb) regarding different demographic groups could cause screeners to automatically process resumes either positively or negatively. Similarly, justifying choices or counteracting negative stereotypes requires cognitive engagement. If screeners are discriminating without cognitive engagement, it is possible that they can be trained to counteract implicit biases (Bezrukova et al. 2016),[5] whereas if they are already cognitively engaging, they may be justifying their discriminatory choices and thus implicit bias training may be ineffective (Devine et al. 2012).

Finally, and importantly, although recent progress has been made on the reasons for labor market discrimination by race (e.g. Lang and Lehman 2012, Lang and Manove 2011, Lang et al. 2005, Nunley et al. 2015, etc.), there is still much work to be done. In economics, we often conceptualize discrimination in terms of taste-based discrimination (Becker 1971), levels-based statistical discrimination (Phelps 1972), and variance-based statistical discrimination (Aigner and Cain 1977). Taste-based discrimination, or animus, occurs when employers, employees, or customers gain disutility from interacting with a specific group of workers, for example, black

[5] Alternatively, they may be indulging in taste-based discrimination, in which case companies would need to increase the costs of discrimination in order to counteract it.

workers (or alternatively, they gain utility from the act of discrimination). Levels-based statistical discrimination occurs when one group has lower "quality" on average and these group characteristics influence how individuals are treated. Levels-based statistical discrimination can be partitioned into different stereotypes, for example, that black applicants had worse training on average. Although levels-based statistical discrimination theories generally assume that the stereotypes at play are correct on average, the results discussed in this paper will also apply to incorrect stereotypes. Variance-based statistical discrimination occurs when employers believe that quality signals are less informative for the minority group compared to the majority group. For example, employers might understand what graduating from a specific high school means for white applicants, but not for black applicants.

This paper uses a laboratory experiment to explore the black box of the hiring process. We randomly vary the content of resumes for an entry-level clerical position. We provide resumes with names that signal differences between races, genders, and socioeconomic status, and that also vary by date of high school graduation, indicating that the applicant is between the ages of 36 and 76. We then ask MBA, MPA, HR, and undergraduate business students to rate the resumes on a 1-7 Likert scale. While they are viewing the resumes, we use eye-tracking technology to track for how long and where on the resumes they are looking.

We find striking evidence of intersectionality for race by age. While we find, like most previous literature (see Lang and Lehman 2012 for a review), that younger white applicants are preferred to younger black applicants, this preference diminishes with age as white applicants become less attractive and black applicants become more attractive. Indeed, we find no preference for white compared to black applicants in their 50s, and black applicants are even preferred in some specifications. As applicants age, differences reemerge. We find a slight increase in the preferences for white applicants in their 60s and older as has been found in previous work (Lahey 2008), but at the same time, there is a steep decline in preferences for older black applicants, something that has not been studied in previous literature to our knowledge. We supplement these findings with patterns of hiring by age and race in the matched monthly CPS and from a smaller study using HR managers as participants. We do not find additional intersectional outcomes splitting by gender, although supplemental work suggests that there may be additional differences by gender in other markets. Results are unaffected by age of participant, participant hiring experience, controlling for the socioeconomic status of the name, and various specification choices.

In terms of mechanics, we find that screeners do look at the entire resume for each applicant and do not stop reading after seeing a name or graduation date. They do, however, spend less time on younger black resumes compared to younger white resumes, and the patterns of time spent on resumes by age are very similar to the patterns for ratings discussed in the previous paragraph. Similarly, viewers' gaze patterns do not differ by the ages and races indicated on the resumes. We also find evidence that screeners are not as cognitively engaged with younger black resumes compared to younger white resumes, suggesting that they are either using negative heuristics when screening these resumes or they are engaging in taste-based discrimination.

We cannot rule out taste-based discrimination in our study framework, although given our intersectionality results, we can rule out a simple model of taste-based discrimination against black applicants that does not differentiate by age. We find evidence for levels-based statistical discrimination suggesting that screeners believe that younger black applicants have worse computer skills and worse training than younger white applicants and are more likely to have had a gap in their work history (perhaps because of prison time). We find no evidence that screeners believe that black applicants have worse or less relevant experience. We also find compelling evidence for variance-based statistical discrimination against black applicants of all ages. Drilling down into the components of this type of discrimination, we find that screeners perceive the job history signal to be stronger for white applicants of all ages than for black applicants and suggestive evidence that the signal sent by the high school attended is stronger for younger white applicants compared to younger black applicants. We find no evidence of strength of signal for high school attended for older applicants by race, nor do we find any evidence that the signal provided by the applicant's address varies by race.

II. Operationalization of Theoretical Predictions

Our experiment provides a unique opportunity to explore these open questions about the existence of age*race intersectionality, about how screeners process resumes, and importantly, why there might be hiring discrimination. We randomly vary the content of resumes for an entry-level clerical position so the resumes contain different combinations of ages (indicated by date of high school graduation), race (indicated by first name), and other resume items. Our participants provide information when they rate the resumes on a 1-7 “Hireability” scale and from their eye movements and time spent viewing each resume and each part of the resume. This section discusses how the outcomes of theory derived from previous literature would be operationalized within our experimental framework.

Age*Race Intersectionality

In a regression context, intersectionality is generally operationalized as a simple interaction term including main effects. For example, if the coefficient of an interaction term of $Black * Female$ is negative and significant controlling for the main effects of *Black* and *Female*, then that would indicate that black women have negative outcomes beyond the effect of being black alone or being female alone. Exploring intersectionality in terms of *Age* is more complicated both because *Age* is a continuous variable and also because *Age* often has non-linear effects on outcomes. To explore the full effect of intersectionality of *Age***Black* we run the following regression:

$$(1) \qquad Hireability_r = \beta_1 * Age_r + \beta_2 * Age_r^2 + \beta_3 * Black_r + \beta_4 * Black_r * Age_r + \beta_5 * Black_r * Age_r^2 + \gamma_p + \alpha_r + \varepsilon_r$$

$Hireability_r$ is a Likert (1-7) score with 7 as the highest rating and 1 as the lowest rating. Age_r is the age of applicant on the resume. $Black_r$ is an indicator for having a “black” name. The main results are clustered on participant, and some robustness checks include participant fixed effects, γ_p. The equation ends with constant α_r and error term ε_r. A significant coefficient on β_4 indicates that age has different effects by the race of the applicant, while a significant effect on β_5 indicates that this interactive effect is non-linear. Again, because *Age* is non-linear, it is non-trivial to read the signs of β_4 and β_5 to get a single effect. Plotting out the quadratics by *Age*, including their confidence intervals, separately for *Black* and non-*Black* applicants provides a simple visualization. If estimated curves are parallel or their confidence intervals overlap throughout the distribution, then that suggests that *Black* and non-*Black* applicants are treated similarly by age and age*race intersectionality does not exist, at least for this functional form. Conversely, if the estimated curves diverge, then that provides evidence for intersectionality for race by age.

Mechanics of Resume Viewing

The mechanics of how screeners view resumes have both policy and theoretical implications. A primary question is whether or not screeners view the entire resume after seeing the name or date of high school graduation. This question can be addressed via eye-tracking—we know where screeners look after viewing the name and if they have viewed all sections of the resume.[6] If screeners do not view the entire resume for disadvantaged resumes, then applicants in those groups will be unable to improve their chances of being hired by improving their resumes. Alternatively, if after viewing the name or date of high school graduation screeners look at different parts of the resume based on race or age, then that will provide evidence about what specific information screeners find differentially important across groups. Other ways in which the mechanics of resume viewing combined with rating information can be used to advise applicants from disfavored groups (Johnson and Lahey 2011) are intertwined with levels-based statistical discrimination, and we discuss them in more detail in that section.

Because the eyetracking technology used in this study is still imperfect and does not catch the entire view-path for all participants, we also use time spent per resume and time spent in each area of interest (AOI), which is essentially a box around each section of the resume. We use a modified version of equation (1) with $Time\_spent_r$ on the entire resume in place of $Hireability_r$ to explore how long they view each resume based on age and race.[7] If the time spent viewing resumes varies by age and by race, then that means that the screeners are processing these resumes differently, even though the resumes for these groups are identical on average.

A final mechanic of how screeners process resumes is more theoretical. We may be interested in whether discriminatory and non-discriminatory processes are automatic or whether

[6] Eye-tracking technology has improved dramatically even in the past decade, and it is now possible for eyes to be tracked over a computer screen using a small box at the base of a monitor (Feng 2011). Eye-tracking use in economics has primarily been used to study marketing (Chandon et al. 2009, Lohse 1997, Maughan et al. 2007, Pieters et al. 1999, Pieters and Wedel 2004, Reutskaja et al. 2011, Russo and Leclerc 1994) and in a few notable neuro-economics papers (e.g. Knoepfle et al. 2009, Wang et al. 2010). For more information on the use of eye-tracking in the social sciences, see Lahey and Oxley (2016). Mouse-tracking is a similar technology that tracks where on a screen a mouse has pointed. Eye-tracking does a finer job of tracking where people are looking, while mouse-tracking has the benefit that it is less expensive and does not need to be done in person, as with the field experiments in Bartoš et al. (2016).

[7] We have also used the total number of fixations (or "count") as an outcome variable and it tracks the *Time_spent* variable closely enough that we consider these results to be redundant and only present the time spent outcomes. Count outcomes are available from the authors upon request.

they require cognitive engagement. Spending less time on resumes means either unconscious implicit biases are activated, or that there is an explicit overt rule that the resume screener is activating. Spending more time on resumes means that screeners like the resumes more, or that they have to think about their choices.[8] Figure 1 interacts these different possibilities with screeners' resume preferences by the race indicated on the resume.[9] Note that in this chart, for simplicity's sake we are not including age interactions, so one should think of having separate charts for younger applicants and for older applicants that could have findings that land in different boxes in the chart. If we find that screeners spend less time on black resumes and prefer white resumes to black resumes, then that would suggest that screeners are using negative heuristics against black applicants, or that they have taste-based discrimination against black applicants. If they spend less time on black resumes and prefer black resumes, then that could suggest that either they have positive heuristics that they are using in favor of black applicants or that they are using affirmative action heuristics to screen resumes. If, instead, we see that screeners spend more time on black resumes than white resumes and they prefer white resumes, then that suggests they may be spending time justifying their choice to not pick the black applicant (e.g. Lindner et al. 2014, Norton et al. 2004, Uhlmann and Cohen 2005). Note that being in this box does not preclude either taste-based discrimination or statistical discrimination against black candidates, just that screeners spend additional time justifying that decision. Finally, if screeners spend more time on black resumes and prefer black resumes, they could either be justifying the choice for black applicants based on their biases as they would for white resumes in the previous box, or they could be spending time counteracting their known bias.[10]

[8] Although it is beyond the scope of this paper, we recommend that researchers interested in the psychological debates about automatic vs. non-automatic processes read Sherman et al. (2014), which provides extensive discussions of dual process theories in psychology from the leading names in the field.

[9] A 3x3 version of this chart separates "prefer black resumes" from "no preference" and allows an additional column in which there is no difference in time spent by race indicated on resume. These cells were omitted for simplicity but are available in Appendix Figure 2. Note that the simplified chart presented is not symmetric because it implicitly assumes that taste-based discrimination against white applicants is unlikely. One can easily fill in symmetries relaxing that assumption.

[10] People spend more time viewing things that are unusual (see Holmqvist et al. (2011) for a literature review), which could push people from the left boxes to their corresponding right boxes, depending on their biases about what is unusual for different racial groups applying to a clerical position. This scenario is not problematic given our findings.

	Automatic less time on black resumes	Cognitively Engaged more time on black resumes
prefer white resumes	negative heuristics vs black or taste-based vs black	justifying choice vs. black
prefer black resumes or no preference	positive heuristics for black or affirmative action heuristics	counteracting neg. stereotypes or justifying choice for black

Figure 1: Preference and Cognitive Engagement

Why is there Discrimination?

Taste-based Discrimination

Absent of a participant vocalizing their biases, something that social desirability bias depresses (Sigall and Page 1971, Wicker 1969), there is little that can be done to directly measure taste-based discrimination.[11] Many early studies that report taste-based discrimination find it as the residual in a regression controlling for observables, often using a Oaxaca-Blinder framework (see Altonji and Blank 1999 for an early literature review), though notable exceptions include Ewens et al. (2014) and Yinger (1986) in housing markets and Nunley et al. (2015)'s audit study that finds evidence of customer taste-based discrimination against black recent college graduates. In our set-up, taste-based discrimination against black applicants could occur in either scenario in which white resumes are preferred over black, the difference being the amount of mental justification that the screener does. A model of simple taste-based discrimination would suggest that white resumes are preferred to black resumes at a constant rate across all ages. If we find intersectionality with race and age, as we do, that suggests either a more complicated model of taste-based discrimination or a combination of taste-based and statistical discrimination.[12] Unfortunately, as with most papers on discrimination, our paper will not be able to say much more about this type of discrimination.

Levels-based Statistical Discrimination

[11] But also see Charles and Guryan (2008)'s use of GSS questions in conjunction with black-white wage differentials.

[12] Although out of the scope of our framework, note that taste-based discrimination among management can lead to statistical differences in productivity as in Glover et al. (2017).

The simplest version of statistical discrimination is termed levels-based statistical discrimination (Phelps 1972). With this kind of discrimination, employers believe there are differences in the average quality of the two groups of applicants, for example black applicants and white applicants. Because screening for applicant quality can be expensive in the absence of easily accessible information, employers will expedite screening by attributing the average quality of the group to the individual candidate. These screening short-cuts can be operationalized via stereotypes. Common stereotypes about black workers compared to white workers often focus on pre-labor market differences, such as schooling quality (Altonji and Blank 1999, Card and Krueger 1992, Lang and Manove 2011, Margo 1986) or neighborhood quality (Dawkins et al. 2005, and see Altonji and Blank 1999 for a review). Additional beliefs may concern skills or previous labor market experiences (Holzer 1998, Lang et al. 2005, Smith and Welch 1989) or criminal history (Holzer et al. 2006).[13] In simple terms, if there is levels-based discrimination because of differences for specific attributes, screeners will assume that white applicants already have positive attributes if they are not specifically listed, while they will assume that black applicants do not.

Researchers have found evidence of general statistical discrimination against black applicants and workers. Autor and Scarborough (2008) finds evidence of statistical discrimination in hiring, although their measure of quality is a black box that cannot directly inform our experiment. Similarly, Arcidiacono et al. (2010) find evidence of statistical discrimination for black male high school graduates in wages. In contrast, Altonji and Pierret (2001) find no evidence of statistical discrimination in wages.

Several recent papers have found evidence of statistical discrimination for specific measures of potential productivity or costs. Agan and Starr (2018) and Doleac and Hansen (2016) use different methods with state law "ban the box" natural experiments and each find evidence of statistical discrimination against black workers based on the probability of having a criminal record. Doleac and Hansen (2016) additionally find that when the criminal history screen is disallowed, employment shifts from low-skilled younger black men to black women and older black men. While we will not be able to test the effect of criminal history in our study, we can test

[13] Screeners may also have different underlying beliefs about the job-related abilities of the two groups for which schooling can be a formal substitute for informal training; they may believe that black applicants have less job-related training, or that they lack computer skills.

the effects of job gaps, which could be coterminous with incarceration (Holzer et al. 2006), although job gaps also indicate unemployment and thus potential productivity differences. Testing another measure of productivity, Lang and Manove (2011) and Pinkston (2006) suggest that education is a clear signal for black workers; in our context their findings would predict that an investment in additional training will help black workers compared to white workers.[14]

The levels-based statistical discrimination model would have two important predictions in the context of our experiment. First, when information is made easily available and shows that the (younger) black candidate is of higher quality than the predicted average or has a specific skill that employers believe to be less common in (younger) black candidates than in (younger) white candidates, then that information should help (younger) black candidates more than it helps (younger) white candidates. (Conversely, when negative information is provided, then that should hurt the black candidate less than the white candidate.) To sum, positive information on the resume should help the discriminated group more than it does the non-discriminated group. Second, when our hypothetical employers know that this information may be available on resumes, then they should spend more time looking at resumes and in resume sections with this information for younger black resumes than they do for younger white, because they will assume on average that younger white applicants already have these skills but younger black applicants do not.

Equation (2) operationalizes this idea.

$$(2) \qquad Y_r = \beta_1 item_r * Black_r + \beta_2 item_r + \beta_3 Black_r + \boldsymbol{X}\beta_4 + \alpha + \varepsilon_r$$

For the first prediction, Y_r will be: $Hireability_r$. For the second prediction, Y_r will be $Time\_spent_r$, which is time spent on the resume overall and $Time\_spent\_AOI_{ar}$ which provides the time spent on a specific area of interest. The variable $item_r$ indicates if an item that we are using to indicate a specific quality measure such as computer training or previous clerical experience, is included on the resume. Alternatively, for non-binary signals, $item_r$ will be a continuous variable such as number of years of job gaps in the resume. A vector of controls $\boldsymbol{X}$

[14] Ritter and Taylor (2011) suggest that differences in communication skills and difficulty monitoring black compared to white workers may contribute to racial disparities in unemployment. We do not have a clear way to test for this possibility in our setup, although in results available from the authors, we do explore the effects of typos, number of words used in job history statements, "professionalism" of job history statements, and so on and find no differential effect of these items by race. However, these signals may be poor proxies for on-job communications skills.

include age and age^2 and participant fixed effects in some specifications. Other variables are as defined previously.[15] With $Hireability_r$ as the outcome of interest, if we code $item_r$ to have larger values provide positive information, a positive and significant β_1 will indicate that having the item on the resume helps black applicants more than white applicants, thus showing evidence of levels-based statistical discrimination based on that specific attribute. Similarly, when $Time\_spent_r$ is the outcome of interest, a positive and significant β_1 will indicate that the screener is spending more time overall on resumes with these attributes, thus providing mechanical evidence that statistical discrimination is occurring. Finally, when $Time\_spent\_AOI_{ar}$ is the outcome of interest, a positive and significant β_1 will indicate that the screener is spending more time on the part of the resume that has positive information about the stereotype for black resumes compared to white resumes, providing further evidence of levels-based statistical discrimination.

Variance Based Statistical Discrimination

Variance-based statistical discrimination is more complicated than levels-based (Aigner and Cain 1977). With this type of discrimination, it is not that black applicants necessarily have lower skills than white applicants, but that the signal for these skills is not as strong for blacks compared to whites. Here, a black applicant and a white applicant might show the same signal, for example, a high quality high school, but the signal would be more meaningful for white applicants compared to black applicants. Variance-based statistical discrimination would predict that at high levels of signaled quality, the group for whom the signal is better (generally the majority group, in this case white applicants) will be preferred, while at low levels of signaled quality, the group for whom the signal is worse (generally the minority group, in this case black applicants) will be preferred. Figure 2 provides a theoretical example of what a graph plotting signal quality vs. preferences might look like.

[15] Equation (1) can combined with equation (2), allowing effects to vary quadratically by age, but the results are more difficult to parse given the triple interaction with quadratic age terms. Results for these are available from the authors in regression or graphical form.

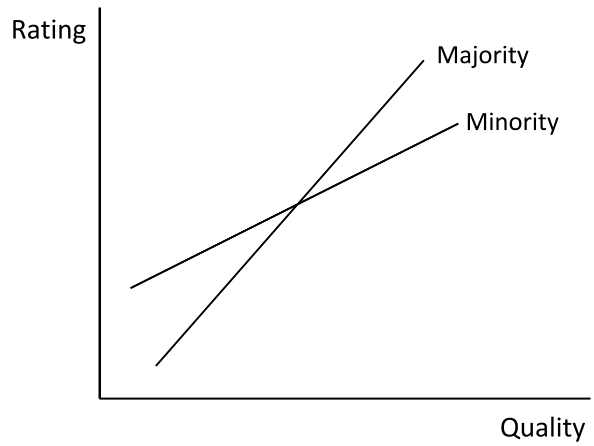


Figure 2: Theoretical Predictions of Variance-Based Statistical Discrimination

Many of the differences in pre-labor market conditions between black and white people could lead to either levels-based statistical discrimination, in which screeners believe that the level of an item is lower for black applicants, or variance-based statistical discrimination, in which screeners are less sure about the meaning of a signal for black applicants compared to white. For example, housing segregation leads to black children going to worse schools on average (Charles 2003), but discrimination within schools (Lhamon and Samuels 2014) could also lead to employers being less certain about the quality of education that black children get at mostly white schools. These differences have different predictions and are testable within our framework. Levels-based statistical discrimination suggests that providing a positive signal of quality would increase the rating of a black applicant more than it would the rating of a white applicant while variance-based statistical discrimination suggests the reverse.

Lang and Manove (2011) suggest that employers have a harder time observing productivity signals (other than education) for black applicants than for white.[16] In our study, this suggestion would operationalize as work histories providing a clearer signal for white applicants than for black, thus predicted quality measures based on work histories plotted against screeners' quality ratings should follow the same pattern as Figure 2. Pinkston (2006) similarly suggests that experience provides a less clear signal than does education for young black men. Another reason

[16] Fryer et al. (2005) provides a laboratory experiment suggesting that initial differences in investment costs could persist throughout the job history, which could be a reason for job histories providing a less clear signal of productivity.

screeners may discount the employment experience of black applicants compared to white is if they believe that affirmative action was a factor in black applicants' job histories.

In contrast to Lang and Manove (2011), differences in schooling quality may be an important reason why black resumes and white resumes may be treated differently. The quality of education in addition to the quality of education signals for black students have varied significantly over time with changing investment in education and desegregation for black students (e.g. Margo 1990). Because of these changes, it may be more difficult for screeners to understand the signal provided by high school graduation for black applicants compared to white. And, as mentioned previously, race discrimination within schools (Lhamon and Samuels 2014) could also make education signals stronger for majority white candidates compared to black.

The address of an applicant may be important for two reasons. First, address provides information about commute times and ability for the applicant to get to work on time, or, in a spatial mismatch model, their reservation wage (Dawkins et al. 2005, Immergluck 1998). In our set-up, screeners do not know the identity or location of the company for which they are doing the screening (which could lead to a test for levels-based statistical discrimination based on commute times), but they may still have beliefs about the general accessibility of different neighborhoods to employment on average and how that might vary with access to reliable transportation by race. If there is more variance in screeners' beliefs that a black applicant has, for example, a reliable car, then that could show up as variance-based statistical discrimination. Address also signals socioeconomic status and many things bundled with it such as primary school education quality (Ellen and Turner 1997). These signals could also result in levels-based or variance-based statistical discrimination depending on assumptions and clarity of the signals. While we will not be testing address in a levels-based context in this paper,[17] we will be testing address as a component of variance-based statistical discrimination.

The predictions of the variance-based discrimination model would be first that the graph of ratings vs. predicted ratings by group status would show the lines for black and white applicants crossing each other, with the white line above the black when predicted ratings are high and below when predicted rating are low. Second, screeners would spend less time looking at black compared

[17] The effects of socioeconomic status by race found using this experimental setup are non-linear and require a longer exploration than can be provided in this paper.

to white resumes with positive signals if those signals are stronger for whites. Third, reviewers will spend less time looking for such signals for black resumes compared to white resumes. Note that for variance-based statistical discrimination, unlike levels-based statistical discrimination, additional positive information for specific skills will not help black applicants more than white applicants because the signal will be trusted for white applicants with higher skill signals but not for similar black applicants.

To operationalize these tests for variance-based statistical discrimination, we first create a predicted quality measure by regressing the $Hireability_r$ measure on a set of controls that covers every resume input except those for race, ethnicity, and age. Each individual resume input (individual job histories, high schools, additional training, volunteering, statements about flexibility, home addresses, email providers) is included as a dummy and additional variables are included that combine job histories, such as having a gap in the job history, length of employment history (and length squared), and length at each job (and length squared). The results of this regression, shown in Appendix Table 1, are then used to estimate a predicted rating, or predicted "quality" for each resume.[18] This predicted quality measure is then graphed against the actual ratings of the resumes separately for each race. Variance-based statistical discrimination would predict that the line for whites will be higher than that of blacks at higher levels of resume quality but the line for blacks will be higher than that of whites at lower levels of resume quality.

To test the predictions that screeners spend comparatively less time on black resumes and AOI with positive signals, we return to equation (2). Here the predictions are different than they were for levels-based statistical discrimination. With $Time\_spent_r$ as the outcome of interest, a negative and significant β_1 will indicate that screeners spend relatively less time on the resume with positive item for black resumes compared to similar white resumes. Similarly with $Time\_spent\_AOI_{ar}$ as the outcome, a negative and significant β_1 will indicate that screeners spend relatively less time on area of interest with positive item for black resumes compared to similar white resumes.

One benefit to our set-up is that we can dig deeper into the black box of quality to explore which parts of our quality measure provide the crossover pattern from our predicted quality

[18] Results by race are robust to inclusion or exclusion of last names, inclusion of age dummies, and basing quality predictions off of white resumes only.

measure, thus indicating that they are noisier signals for black applicants compared to white. Using this method, we can test the effects of job history, high school, and address. We test for the signal quality of these multi-value resume components by using a method similar to our original test of variance-based statistical discrimination. To determine high school "quality" we regress our Likert (1-7) scale on all of the high school dummies in our sample and predict the Likert based on that regression. We perform identical procedures with address dummies and with job histories including job gaps and employment duration.[19] It is important to note that we are allowing the resume viewers to determine quality—we are specifically not using an objective measure of quality or a measurable feature of these items like socioeconomic status. It is the screeners' beliefs about quality that are important for these decisions, not the researchers'. We then graph our predicted "quality" measures against the actual Likert ratings of the resumes by race indicated on the resume. The black and white lines crossing again indicate variance-based statistical discrimination while parallel or identical lines do not.

III. The Experiment

Design

The study took place at the Brain and Gender Laboratory at a large southwestern university. Subjects were recruited via flyer and were restricted to MBA and MPA graduate students and human resources and business school students more generally. One hundred fifty-two participants participated in the study between January 2013 and January 2014 and each earned $20 for the session. Two participants were dropped for being non-native English speakers and one participant was dropped because of a diagnosed learning disability; these could affect eye-tracking (Holmqvist et al. 2011). Total time allotted to the study was one hour, but the majority of participants finished in less than 45 minutes.

Participants rated resumes for an open administrative assistant position. These resumes were created with Lahey and Beasley (2009)'s resume randomizer program and used a database

[19] We have repeated this chart without gaps or employment duration and results are very similar, but not quite as stark, suggesting that the bulk, but not all, of the effects found are from the job history dummies themselves. Results are available from the authors upon request.

of resume inputs drawn from actual resumes and from previous studies on discrimination. Variation included age (as indicated by date of high school graduation), gender and race (as indicated by first name), home address, email provider, high school attended, work experience,[20] additional training, volunteer experience, and a statement about flexibility. Fictional applicant names indicated race (Aura and Hess 2010, Bertrand and Mullainathan 2004, Figlio 2005, Fryer and Levitt 2004, Levitt and Dubner 2005, Lieberson and Bell 1992, Lieberson and Mikelson 1995), gender (http://www.babynamewizard.com/), and socioeconomic status (Figlio 2005, Levitt and Dubner 2005, Lieberson and Bell 1992, Mehrabian and Piercy 1993). Addresses were drawn from the nearest large metropolitan area and high schools were drawn from across the state. The perceived race, gender, ethnicity, and socioeconomic status of the names were checked in a separate study using 95 psychology undergraduates (see Barlow and Lahey 2018 for more details).[21]

We generated 40 unique resumes for each participant, for a total of 6,080 unique resumes, 5,960 of which were used after participants were screened for disabilities and English speaking. Resume line items were repeated across participants; however, each participant only saw each specific line item at most once. 50% of the resumes were given female names, 9% were given black names, and 13% were given Hispanic names. These percentages were chosen to approximate the composition of clerical workers in the US according to the 2012 ACS and are shown in Table 1.

In post-processing we divided resumes into specific "Areas of Interest" (AOI) in order to measure the amount of time spent on each resume section. These AOI are virtual boxes that surround the fixed parts of the resume and include Name, Address, Employment history, Years associated with employment history, Education, Year associated with high school graduation, Other (which includes items such as training, statements of flexibility, and volunteer work), and

[20] All resumes contain up to 10 years of recent work experience. As with previous work (Lahey 2008, Neumark et al forthcoming), we don't find any interactive effect on ratings between number of years worked and age of applicant, suggesting again that for entry level applicants, there should be no concerns with screeners expecting longer experience profiles for older workers in this setup.

[21] All results in the paper are robust to controlling for socioeconomic status of names. Similarly, we tested the perceived socioeconomic status of the addresses and high schools using this separate undergraduate sample. There are additionally interesting interactive effects between race and our three different measures of socioeconomic status. Because these effects vary non-linearly it makes sense to explore them in a separate paper in future work.

Outside (which includes everything on the page not in another AOI). An example of this partition can be seen in Appendix Figure 3.

Procedure

Upon entering the laboratory, participants read an informed consent form and provided consent. Participants' eyes were calibrated with eye tracking equipment, a D6 eye tracking system from Applied Science Laboratories (Bedford, MA) to observe where on the computer screen a participant was looking. Participants were told that the purpose of the research was to study how hiring managers make job interview decisions. They were given the description of a clerical position and asked to evaluate applicants for that position. Participants then viewed five sample resumes, and, following that, rated 40 candidates' resumes one at a time for a hypothetical clerical position using a Likert scale regarding the ability of the candidate to fulfill the position. Participants then rank ordered their top two resumes and their top one resume for fulfilling the position from a presentation of their top five most highly rated resumes (with the more recent resume presented in the case of rating ties). However, because not enough black resumes made it into this top five set, we will not be discussing the results of this part of the experiment in this paper. After rating the resumes, participants completed various psychological, political, and demographic questionnaires (Bogardus 1933, Greenwald et al. 2003, Nosek et al. 2007, Henkens 2005). After they completed the survey, participants were debriefed and paid.

The demographics of our participant sample reflected a variety of people affiliated with the university community, with an intended bias towards those from the business school. As shown in Table 1, 38% of participants were at the Masters level and 1% were PhD students, 38% were upper division undergraduates and 23% were lower division undergraduates.[22] 76% of participants studied business, 13% studied government, 6% studied humanities, and 5% studied other social sciences. The average age was 22 and 56% of the sample was female. The sample was 89% White, 7% Asian, and 5% Black or African American. 15% of participants reported that they identify as Hispanic or Latino.

IV. Results

[22] There is no significant difference between how graduate and undergraduate students rate resumes by race of the resume or race*age of the resume.

Table 1 provides baseline averages on ratings and time spent. The average Likert score (on a 1-7 scale, 7 high) for all resumes was 4.63, with a standard deviation of 1.39. On average, participants spent 16.24 seconds on each resume with a standard deviation of 10.17 seconds. This is slightly higher than, but comparable with, the estimate of 15 seconds often given by human resource professionals when asked (Lahey 2008).

The majority of experimental papers finding hiring discrimination against black applicants (if not all of them, to our knowledge) have focused on younger workers. Therefore, we present two simple t-tests in Table 2 to compare Likert outcomes for black applicants vs. white, first for a sample of our younger applicants and second for our entire sample. There is a significant preference for white resumes over black resumes when using a simple t-test to compare the average scores of black vs. white applicants under the age of 45. On the 7-point Likert scale, resumes with white names are preferred to those with black names with a significant difference of 0.28, which is 6% above the average rating for the entire sample of 4.63. However, when we run a t-test on the entire sample from age 36-76, the difference between black and white resumes declines to .03 and is no longer statistically significant.

Evidence of Age*Race Intersectionality

To get a deeper understanding of intersectional preferences by age and by race, we use a local weighted regression (using a lowess smoother) graph to plot how participants rate resumes by age for each race. It is important to note that no functional form has been imposed on these data other than that of the local weighted regression. Figure 3a demonstrates a small quadratic decline and increase in ratings by age for white resumes, similar to that found for women in Lahey (2008). However, the pattern for black resumes is strikingly different. In Figure 3a, the Likert ratings for black resumes start almost half of a standard deviation lower than those for white resumes, and their rating gradually increases with age until the mid-50s. At that point ratings decrease again to a value slightly above the starting point. Fitting these data to a quadratic in Figure 3b allows us to fit confidence intervals around the outcomes. This fit shows statistically significant differences in the age tails where white resumes are preferred to black as well as in the late 50s and early 60s when black resumes are preferred to white. Figures 3(c)- 3(f) break apart the sample by gender and show similar patterns, though for white men the pattern by age is more linear than quadratic while results for women do not show significant differences in 3(f) given the overlapping confidence intervals.

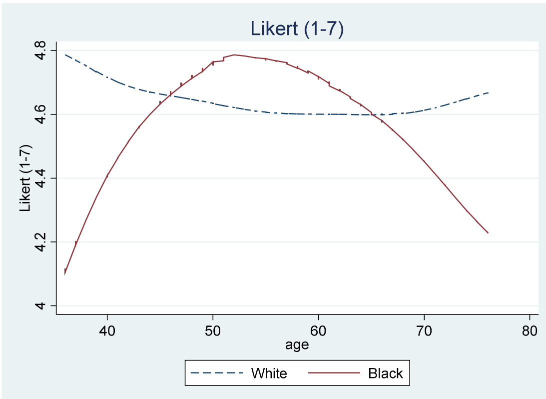


Figure 3a

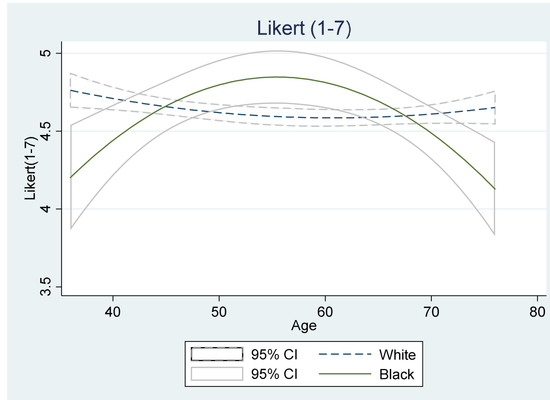


Figure 3b

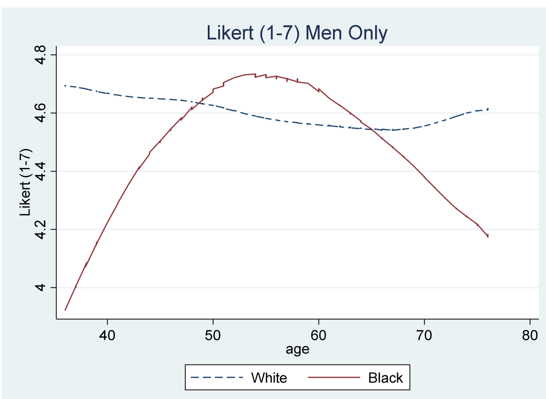


Figure 3c

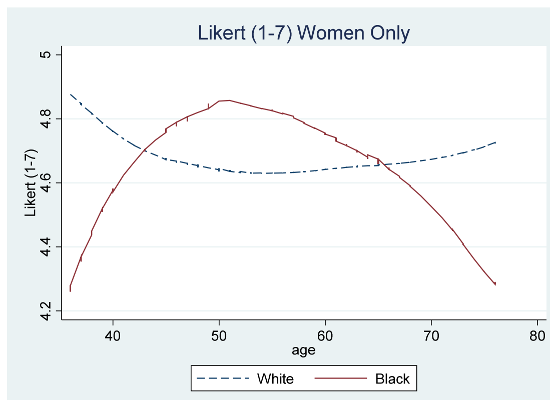


Figure 3d

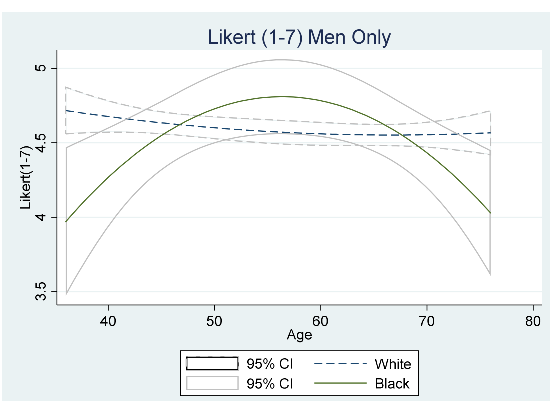


Figure 3e

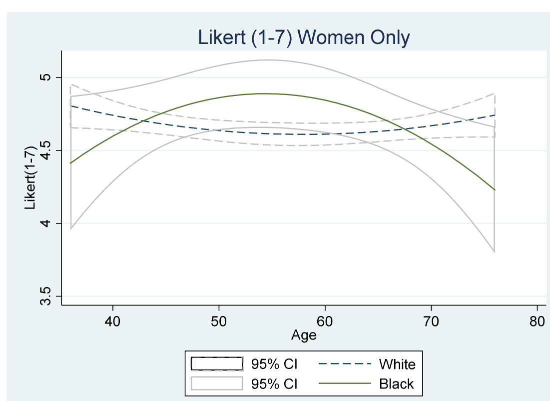


Figure 3f

Figures 3a-f: Lowess and Quadratic Fits of Likert Scale by Age of Applicant (as indicated by date of high school graduation on resume)

We use equation (1) to formalize these results in Table 3. Column 1 provides the results from equation (1) without the age*black interactions. Although the effect of having a black name on the Likert rating has a negative sign (-0.0292) in these regressions, it is not significant. However, when race is interacted with age as a quadratic as in Column (2), the main effects and interacted effects are significant at standard levels. This difference in results suggests that heterogeneity in hireability by age for race masks the effect of race by itself when this interaction is not taken into account. Similarly, columns (3) and (4) break apart the results by gender and are in line with the results from the lowess figures shown in figures 3(c)-3(f). To summarize, there is intersectionality in hireability preferences for race by age, and these differences in treatment vary quadratically with age.

Mechanics of Resume Viewing

We are interested in whether or not the viewing patterns for resumes differ once the name or date of high school graduation has been viewed. Table 4 provides transition matrices from an AOI in time t-1 (the rows) to an AOI in time t (the columns) for our four groups of resumes: young white, young black, older white, and older black. The numbers in each cell denote the probability of these movements such that the sum of all cells in each matrix is one. Neither an LR test (p-value .13) nor a Chi-Squared test (p-value .26) found significant differences between the four transition matrices, suggesting that patterns of viewing across the four groups are not different.[23] For the most part, people look at the resumes from top to bottom, thus the place people look after viewing the name AOI on the resume is the address information AOI, with 66 to 73% of transitions from the name AOI. Similarly, after viewing the age indicator, high school graduation year, they are most likely to move to the education AOI, with 31 to 57% of transitions from the grad year AOI to the education AOI. After viewing the last AOI on the resume, screeners are most likely to return to the employment history, with 32 to 42% of transitions from the other AOI to employment

[23] Markov homogeneity tests were performed using Python code available at http://pysal.readthedocs.io/en/latest/library/spatial_dynamics/markov.html#pysal.spatial_dynamics.markov.homogeneity . This code follows Bickenbach and Bode (2003).

history, or to the education information, with 17 to 32% of transitions going from other to education.

Another way to determine whether or not viewing patterns differ is to look at heatmaps that register the length of visual fixations across the resumes. We have created heatmaps for four separate groups, young white, young black, older white, and older black, in Appendix Figure 4. Viewing intensity is denoted by warmer colors (red, yellow) being areas with more visual fixations and cooler (green, blue) colors being areas with fewer visual fixations. Each heatmap represents the average of viewers in the group it represents, and so should not be affected by the fact that 90% of the resumes have white names while only 10% of the resumes have black names. While we have no statistical tests to compare heatmaps, visually it appears that the patterns of viewing intensity are the same for resumes across all groups, but the resumes with younger black applicants are viewed less intensively overall compared to the other three groups.

Finally, we can look directly at viewing time by age by graphing the time spent viewing each resume on age of the applicant as in Figure 4 using a lowess smoother. This chart shows that, as with the results in the hireability figures, screeners spend less time viewing resumes with black names at the youngest ages. The initial difference is over a second. Then as the age on the resume increases, screeners gradually spend more time on resumes with black names, eventually meeting and surpassing the line for resumes with white names shortly before age 50. At age 50, the amount of time spent on black resumes again decreases while the amount of time spent on white resumes remains flat or increases. For resumes of both races for people age 70 and older, the amount of time spent increases, possibly because people spend more time viewing things that appear unusual (Holmqvist et al. 2011).

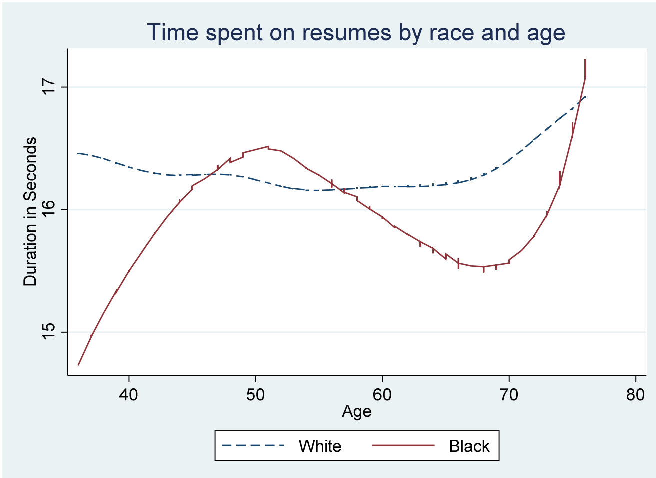


Figure 4: Number of Seconds Spent Viewing Resume as a Function of Age of Applicant (as indicated by date of high school graduation on resume)

To summarize, screeners look at the same parts of the resumes with the same transitions between parts regardless of race or age, but they spend longer on young white resumes than they do on young black resumes. Screeners do not just look at the name or date of high school graduation on a resume and pass on resumes, which means there is a potential for applicants to add things to their resumes to improve their chances in the hiring process. Returning to our simplified theory chart (Figure 1), these combined results suggest that we are in different boxes of the chart depending on the ages listed on the resume. Resumes with younger ages are in the top left box, suggesting either negative heuristics against younger black applicants, or taste-based discrimination that prefers white applicants or both.[24] For the rest of the sample, there is no statistical difference in ratings, though in general screeners spend less time on older black resumes compared to older white, suggesting that older resumes fall in the bottom left box, suggesting positive heuristics generally. As a caution, these results for the older resumes are not as stable as those for the younger resumes.

[24] These time results for the younger portion of our sample are similar to those found in Bartoš et al. (2016)'s comparison of time spent viewing resumes for hypothetical 30 year old white Czech applicants and similar Asian applicants in the Czech labor market. They differ, however, from the comparison between the white applicants and similar Roma applicants. One important difference in the Bartoš study compared to ours is that screeners must first make a choice to click on a hyperlink to each resume prior to viewing, so their time results are conditional on having chosen to open the resume. The external validity of each methodology differs based on how screeners are presented with resumes.

Because screeners view resumes in their entirety, even if they spend less time on resumes for younger black applicants, we can potentially use our methodology to determine more about theories of discrimination.

Theories of Discrimination

Given the differences in outcomes between the younger and older applicants, most of the discussion of theories of discrimination in this section will test these two groups separately. Most of the focus of the discussion will be on the "younger" group because previous literature on race differences in employment outcomes has generally focused on entry-level and "prime-aged" workers. The results shown split younger and older at age 45.[25]

Taste-based Discrimination

As noted before, we cannot say much about taste-based discrimination given our set-up. Because we are in the top left quadrant of Figure 1 for younger black candidates, we cannot rule out taste-based discrimination for these candidates. If only taste-based discrimination is responsible for preference differences by race, screeners would have to change their tastes for discrimination as applicants age.

Levels-based Statistical Discrimination

The first test of levels-based statistical discrimination uses equation (2) to test to see if different types of positive (negative) information help (hurt) black applicants more (less) than white applicants. The first three columns in Table 5 address this question separately for putting items on the resume indicating computer training, years of job gap, or having any clerical experience. Panel I tests these items for the younger applicants, while Panel II tests them for older applicants. Focusing on the Panel I results, Column (1) shows that compared to white applicants, black applicants who put having had computer training on their resumes have a Likert rating that is 1.04 points higher, a result that is even larger than the negative and significant effect of having a black name of -0.34 points. In results not shown, listing any training on the resume improves

[25] Additional regressions are available from the authors splitting the older group into middle-aged and age 65 and older. These two groups are combined for ease of presentation of the regression results because separated results are either insignificant at traditional levels, or when a result for middle-aged or older is significant, then the corresponding coefficient for the other has the same sign.

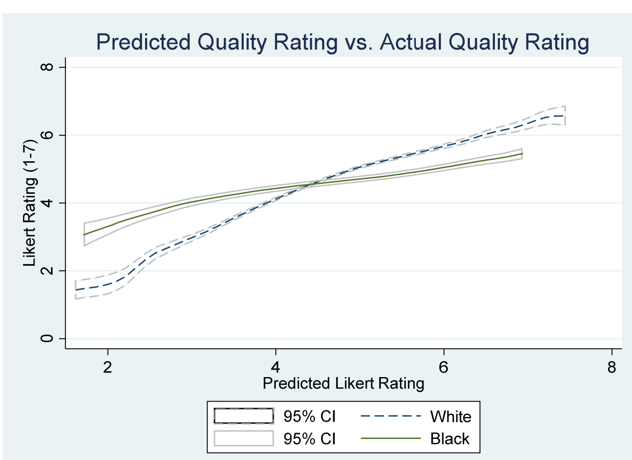
the ratings of black applicants compared to white applicants by 0.63 points, although this improvement is only significant at the 10% level, suggesting that it is computer training itself that is having the effect on outcomes and not just actively seeking training. Column (2) tests to see if more years of having a job gap hurts younger black applicants less than it does younger white applicants, suggesting that screeners expect gaps for black applicants but not for white applicants. Here we see that each additional year of job gap decreases the Likert rating for white applicants compared to black applicants by 0.15 points. These results are in line with recent findings on statistical discrimination against young black applicants based on beliefs about incarceration, though they could also be in line with beliefs about higher unemployment rates for these groups due to animus or similar obstacles. Column (3), on the other hand, does not show evidence of statistical discrimination based on beliefs about previous work experience. Here, previous clerical work experience helps young white applicants significantly more than it helps young black applicants by more than one point on the Likert scale, shown in Panel I, column (3). Thus it is not likely that employers view young black applicants as being less likely on average to have clerical experience compared to young white applicants. No significant interactive effects are found in Panel II for older black applicants compared to older white applicants for any of the items discussed above.

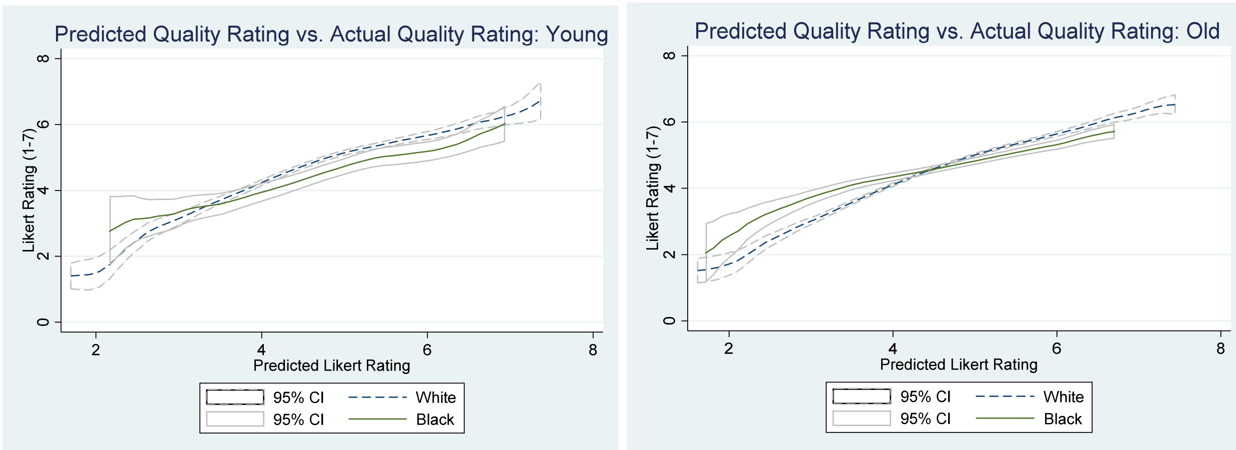
The next test of levels-based statistical discrimination uses eye-tracking to check if more time is spent viewing resumes with positive information indicated for computer training, years of job gap, and clerical experience. These results are shown in columns (4)-(6) of Table 5. Again focusing on the results for younger applicants in Panel I, we see that having computer experience on the resume is correlated with screeners spending 10 more seconds, a full standard deviation more, viewing black resumes compared to white resumes. Similarly, in results not shown, having any training is correlated with screeners spending 6 more seconds viewing black resumes compared to white. Additional years of job gap do not have a significantly different impact on time spent on the resume for young black applicants compared to young white applicants. Again, clerical experience has a negative interactive effect with race, suggesting that it does not cause screeners to spend additional time viewing black resumes. This odd-seeming result will be discussed again later when we turn to variance-based statistical discrimination. Although the signs of the interaction terms are the same for older applicants in Panel II, the magnitudes are smaller and none of them are significant.

Finally, we check to see if screeners spend more time looking at AOI on the resume that could contain information about the positive signal. For the results for younger applicants in Panel I, interactions are generally insignificant—screeners do not appear to spend more time looking at these AOI, although it may be that the eye-tracker used is not sensitive enough to pick up very small differences. Interestingly, there is some evidence in Panel II column (7) that screeners spend significantly longer looking at the training AOI for older white resumes with computer and training experience listed compared to similar older black resumes, although this difference is only around a third of a second and seems to have no effect on Likert ratings or time spent viewing the resume over all. However, these results are consistent with screeners having negative stereotypes about the computer skills of older white workers.

Variance-based Statistical Discrimination

The first test of variance-based statistical discrimination is to take our kitchen-sink predicted quality measure and graph it separately for black resumes and white resumes against the actual rating that each resume receives. We do that in Figure 5a for all resumes using an lpolyci function in Stata. Here we clearly see that higher quality signals, as given by the predicted Likert rating, result in higher actual Likert ratings for white applicants compared to black applicants, with a difference of 1.14 points (4.45 compared to 6.6) at the upper end of the graph. Lower predicted quality ratings hurt black applicants less than they hurt white applicants, with a difference of 1.6 (3.0 compared to 1.4) at the lower end of the graph. The crossover point occurs at 4.4 where the lines cross for black and white ratings.

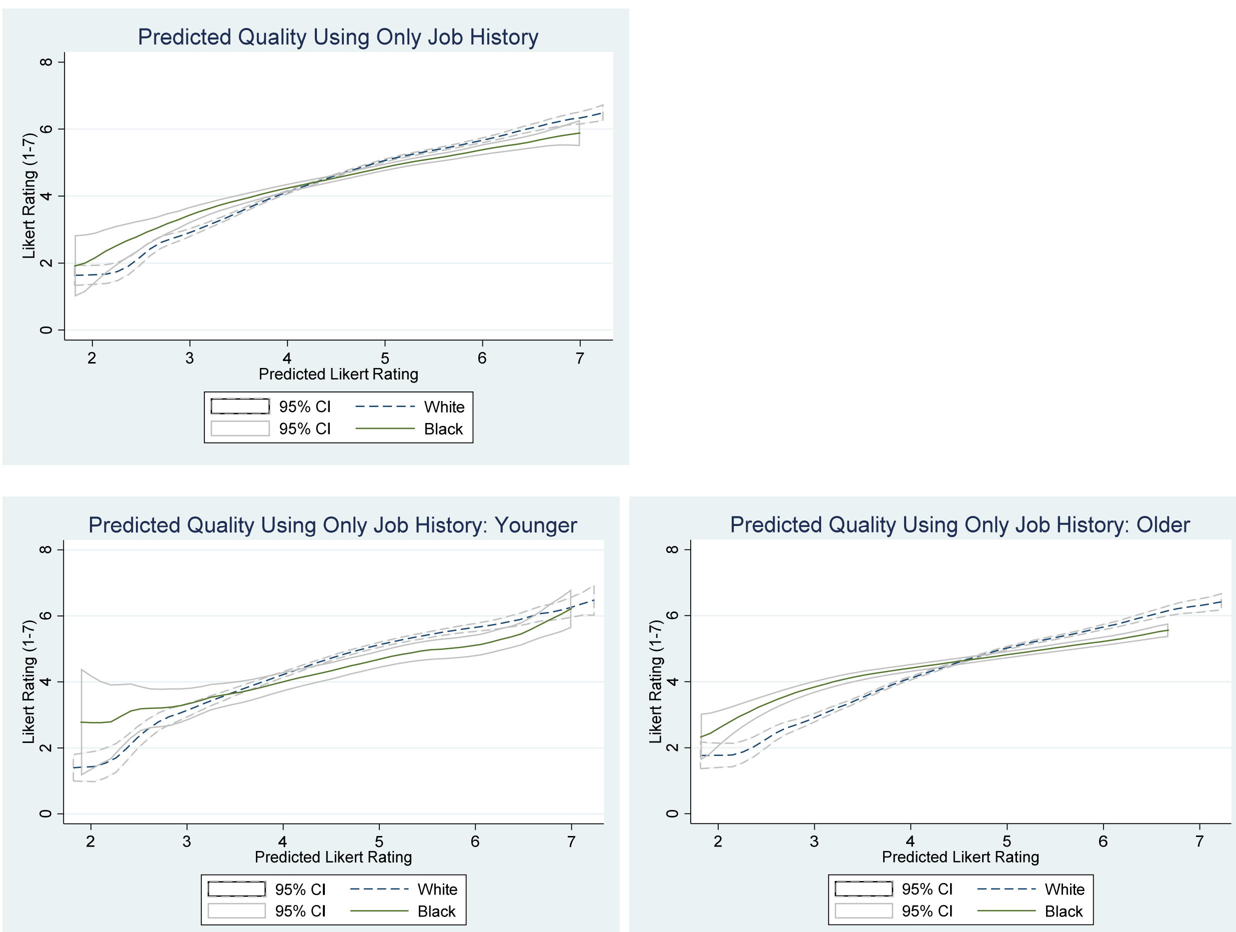




Figures 5a-5c: Screener's Rating of Each Resume as a Function of a Predicted Rating Using All "Legal" Resume Characteristics (See Appendix Table 1 for predicted rating inputs).

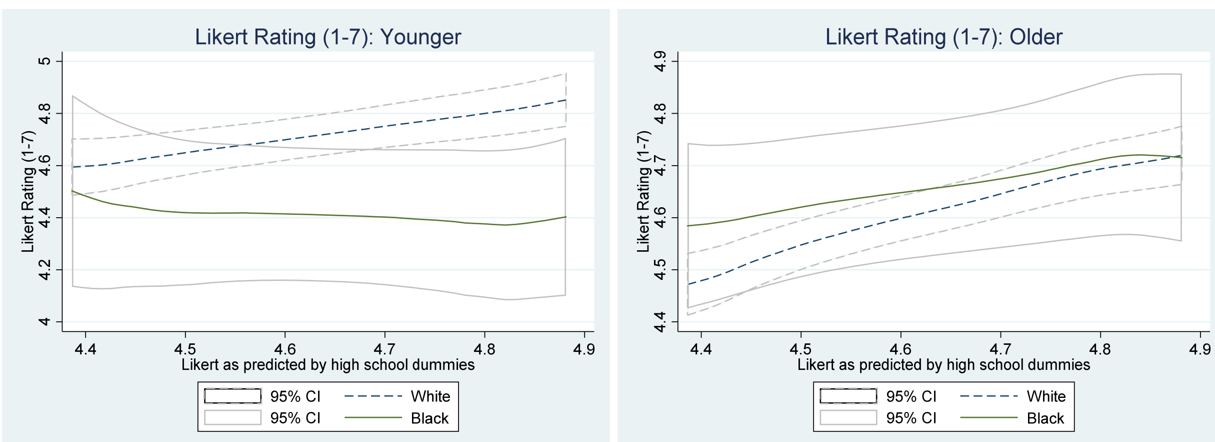
When we look at the relationship between predicted quality rating and actual rating separately for younger vs. older applicants, there are two important takeaways by age. First, the separation of the two lines is larger for the older applicants, suggesting that variance-based discrimination is stronger for older black applicants compared to younger black applicants. Indeed, given the overlap in confidence intervals, we cannot rule out no variance-based statistical discrimination for younger black applicants compared to younger applicants. Second, focusing on the lines without the confidence intervals, the crossover point is much earlier in the predicted quality distribution for the younger group. Given that we expect employers to hire the highest quality workers available (recall, this is "quality" as defined as "hireability" by the participants, not an objective measure of quality like education which could result in "over-qualified" applicants), employers are more likely to hire older applicants than younger applicants from the part of the quality distribution for which black applicants are preferred. For example, if the interview cutoff for worker quality is 4 in our study, they will interview some black applicants over equally qualified white applicants in the older group, but will always interview white applicants over equally qualified black applicants in the younger group.

Using our set-up we can also look into the components of predicted quality that could be driving variance-based statistical discrimination. Recall that having clerical experience on the resume helped young white applicants more than it helped young black applicants in Table 5, Panel I, column (3). Similarly, screeners spent more time viewing young white resumes with clerical experience than they did young black resumes in column (6). Finally, although the result is not significant, participants spent less time looking at the work history AOI for black applicants with clerical experience compared to white applicants with clerical experience in column (9). Taken together, these results suggest that our screeners believe that clerical experience is a stronger signal of quality for younger white applicants than it is for younger black applicants. We then repeat our initial variance-based statistical discrimination figure but this time we create a new predicted quality measure that only uses job history, including dummy variables for each possible job history item as well as quadratics for amount of time on each job and amount of time employed overall, rather than our kitchen sink measure, shown in Figures 6a-6c. Although this contrast for all workers is not as stark as that in our original kitchen sink figure, the pattern is still highly suggestive that job history is contributing to variance-based statistical discrimination. This finding supports earlier work by Pinkston (2006) and Lang and Manove (2011).



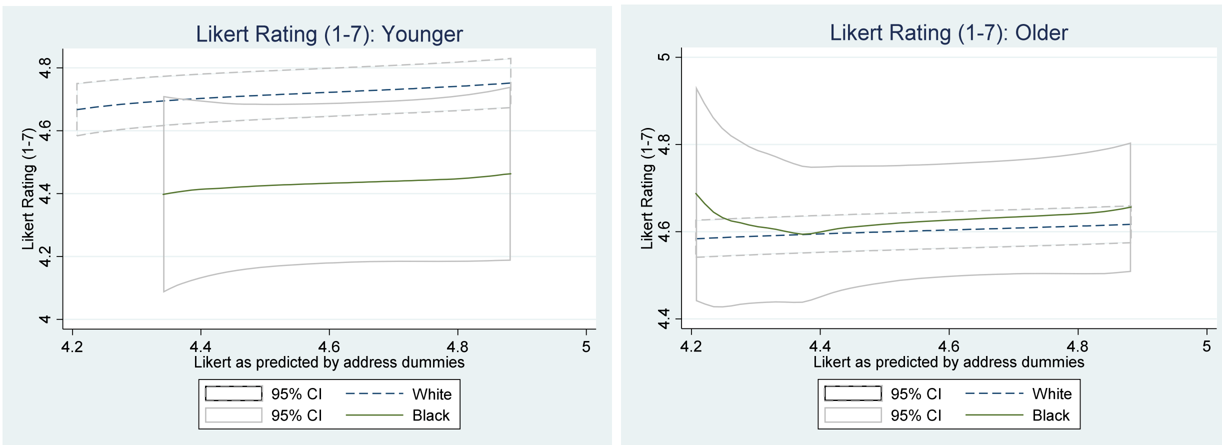
Figures 6a-6c: Screener's Rating of Each Resume as a Function of a Predicted Rating Using Only the Job History.

Other commonly suggested reasons for variance-based statistical discrimination against black applicants are those of school quality (does a high school degree mean the same thing for a black graduate as a white graduate?) and address. Figures 7a and 7b explore the effect of high school "quality" on Likert ratings. Graphing "quality" predicted by high school dummies against the actual Likert rating using a local polynomial with confidence intervals, we see what appears to be significant divergence at higher levels of "quality" for young blacks compared to young whites. Although we do not see a cross-over at lower levels of "quality", it may be that our high schools are not of low enough quality to show a crossover. This picture is consistent with variance-based statistical discrimination with high school not being as good an indicator of applicant quality for young black applicants as it is for young white applicants. For older applicants, the black line is higher than that of whites although it is never significantly different and is not consistent with variance-based statistical discrimination.



Figures 7a and 7b: Screener's Rating of Each Resume as a Function of a Predicted Rating Using Dummies for Each High School. Younger Resumes on Left, Older Resumes on Right.

Figures 8a and 8b repeat this exercise using address dummies in place of high school dummies. Here we see a constant difference (albeit a statistically insignificant one) between young blacks and whites at all points of address "quality" and the lines for older and younger blacks are very similar. These provide no evidence of variance-based statistical discrimination against black applicants. Note that the reasons posited for address-based discrimination against blacks are two-fold: first, employers may have worries about the neighborhood quality signaling something about the applicant, or second, employers may be worried about commute time and reliability based on where the job is situated compared to where the applicant lives. Our experiment does not do as good a job testing the latter compared to the former, as we did not provide an address for our hypothetical job. Address may still be important, possibly in terms of levels-based statistical discrimination, if employers are worried about commute time to work.



Figures 8a and 8b: Screener's Rating of Each Resume as a Function of a Predicted Rating Using Dummies for Each Home Address. Younger Resumes on Left, Older Resumes on Right.

External Validity and Robustness Checks

As with any laboratory experiment, our study has concerns about external validity. First, we would like to emphasize that our results are only valid for hiring for an entry-level clerical position and for applicants who have a high school education and possibly some college, but not any further degrees. While this population is an important one that we should care about given that it accounts for 52% of adults in this age range and 12% of this population is at or below the poverty line (compared to 4% for college graduates) in the 2012 ACS (authors' calculations) we want to stress that the results may not, and probably do not extrapolate to more educated populations and positions that require more targeted skills.

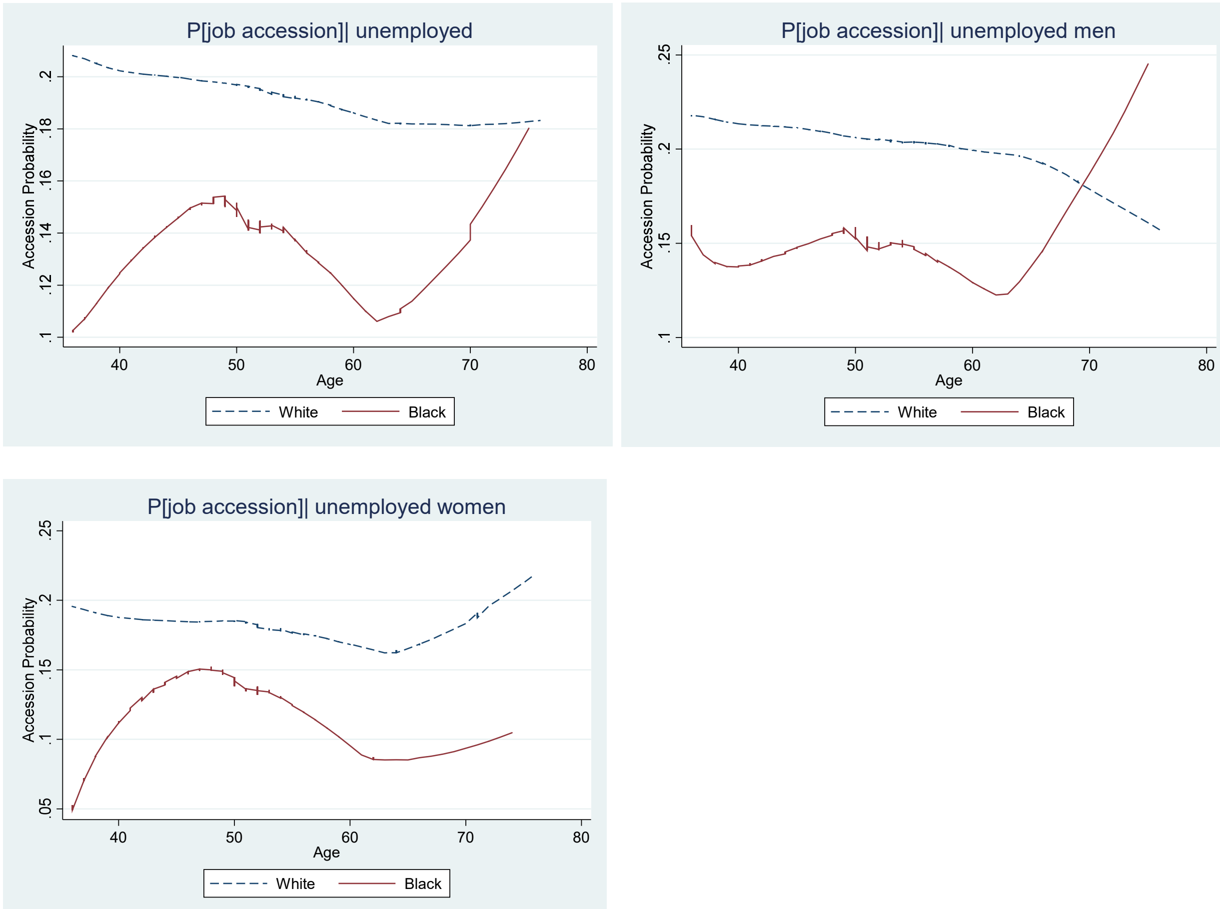
A second concern is that our participant population consists of students, and the highest age in our sample is 35. These groups may not be representative of the hiring population.[26] We were careful to limit our population to MBA/MPA/HR and business students, who may be likely to manage and hire workers in the future. In our sample, 24% of the screeners in our sample have

[26] Although we do not know the overall demographics of the workforce likely to be doing screening for entry-level clerical positions, several snapshots of hiring managers across different industries suggests that the mean age for these groups is somewhere in the 30s (e.g. https://blog.simplyhired.com/hiring-truths/, Giuliano et al. 2011, Giuliano and Ransom 2013, Tews et al. 2011), which is older than our participant population.

experience hiring workers and 11% have HR experience (ranging from 1-8 years). In Appendix Table 2, we compare the main results in Table 3 by age of screener and for screeners with and without hiring experience. We find no significant interactive effects for the main results by having hiring experience, having HR experience, or being age 25 or older.[27] Additionally, even if the student population does not behave the same way that fully trained and experienced HR people behave (which is unlikely to describe all people doing hiring screening), the results from this sample are still real behaviors from real people that give us insight into human behavior.

Another way we address the external validity question is by exploring whether or not the differences from our main results appear in the labor market. We cannot simply look at employment rates by race and age to determine how employers treat black and white applicants differentially by age, because these results will include the full general equilibrium effects of applicants' choices given their treatment in labor market. The closest that we can get to the hiring decision using a publicly available US population dataset is to look at job accessions, that is, becoming employed in month t+1 conditional on having been unemployed in month t, in the matched monthly CPS limiting to high school graduates (Madrian and Lefgrin 2000). This snapshot is still an imperfect replication because black and white candidates do not have the same observable characteristics on average in real life like they do in an experiment, because the measurement of job accession requires a month of unemployment (meaning it does not measure shorter unemployment spells or moving from job to job), and because this measure does not limit to hiring in clerical positions as our study does. That said, the lines for the probability of a job accession conditional on unemployment in the previous month, as shown in Figures 9a-c, show similar quadratic shapes for both black and white job seekers. Unlike the results in our experiment, the line for black job seekers never reaches that of white seekers, but recall that, unlike the hypothetical resumes from our experiment, the unemployed seekers in the CPS are not identical. Separating these results by gender, we can see that the quadratic patterns are more pronounced for female job seekers than for male job seekers.

[27] In an early paper, Cleveland and Berman (1987) also find that students and managers have the same perceptions of the age-type of different jobs, although they do not look at race. Race interaction results as in Appendix Table 2 are similarly insignificant separating grad vs. undergrad instead of by age 25, although undergrads give marginally significantly higher scores overall.



Figures 9a-9c: The Probability of being Employed in Month t Conditional on having been Unemployed in Month t-1 using the 2012 Matched Monthly CPS

We have additional evidence from a second study with a different focus that uses HR managers for its population sample (Lahey and Oxley mimeo). Although the sample size of black resumes from that study is too small to replicate the full study described in this paper,[28] we can replicate Figures 3 (a) (c) and (d) from this paper using that sample. Here we again see a similar pattern for all resumes and for female resumes, but black male resumes do not seem to follow the same pattern as our sample (see Appendix Figures 5). That suggests that this more experienced population treats black female resumes differently by age than they do black male resumes, at least for an entry-level clerical position, though we must caution that the sample sizes for this study are low enough that these results can only be suggestive. Because this newer, more expensive, study uses a different (portable) eyetracker, results cannot be combined with this study.

[28] There are 111 black male resumes and 88 black female resumes in the HR sample.

We provide additional robustness checks for our main results in Table 6. Results are robust to controlling for participant differences, for example including participant fixed effects as in column (2) or normalizing the "hireability" variable across each participant as in column (3).[29] Results are also robust and potentially larger when resumes for groups that may have heterogeneous effects are removed. For example, in column (4), dropping resumes with Hispanic last names from the sample, the magnitude of the black name main effect increases somewhat to -6.7 points, while the magnitudes of black name increases to -8.5 points and the interacted black variables all increase when resumes are limited to those with applicants age 65 and under in column (5). Finally, controlling for perceived name SES using the SES ratings from Barlow and Lahey (2018) in column (6), the results are virtually unchanged from the original results, with a slight increase in the black name variable to -6.1 and no real change in the other variables.

V. Discussion and Conclusion

This paper uses a laboratory experiment with MBA/MPA/HR and business students to explore how screeners for an entry-level clerical position interact with resumes indicating different races and ages. Randomized resumes combined with eye-tracking technology allow us to explore the mechanics of resume viewing by age and race in addition to theories of discrimination. We find strong evidence of intersectional discrimination by age and by race of applicant. Younger white applicants are preferred to younger black applicants, but these differences diminish and even disappear into middle age before reemerging. Screeners do view the entire resumes of less-preferred groups, and their viewing patterns do not vary by type of resume, but they spend less time on resumes for younger black resumes than for other groups. These viewing results suggest that screeners are either using negative heuristics or are engaging in taste-based discrimination against younger black applicants. External validity for these results is bolstered using job flows from unemployment to employment in the matched monthly CPS as well as a smaller laboratory experiment using HR managers as participants, although both of these studies suggest that effects may be stronger for female applicants compared to male applicants, at least for this entry-level clerical position. Although more research has been done on the labor force outcomes by race for

[29] Note that the coefficient is not measured in the same units as the main results because of this normalization, but the sign and significance can still be compared directly.

younger applicants, much more work needs to be done exploring race differences in employment outcomes by age.

It is clear from our work that there is not just one overriding reason for hiring discrimination against younger black applicants. While we cannot rule out taste-based discrimination as a reason for differential treatment by age and race, we can also not test for it in our current set-up. We find strong evidence of levels-based statistical discrimination against younger black applicants having worse computer skills and less training, as well as evidence that screeners are more likely to expect work gaps for younger black applicants compared to white. We also find compelling evidence of variance-based statistical discrimination against all black applicants. Job history signals and possibly high school signals are clearer for white applicants compared to black applicants.

The fact that screeners do view entire resumes even for less-preferred groups means that discriminated-against applicants can include items on their resumes that combat levels-based statistical discrimination. Other potential changes to decrease discrimination include employers requiring applicants to include information about their computer skills and similar items. Employers could also take the step of removing information such as the applicant's name, address, and graduation dates and other items that would signal that an applicant is from a disadvantaged group, a policy that has been gaining more traction recently.[30] Finally, employers could train their screeners about these potential types of discrimination in the hopes that employees will combat it, though such training has had mixed success empirically (Devine et al. 2012, Bezrukova et al. 2016).

We hope that this paper provides early evidence of the need for more research on intersectionality in labor markets overall. Most intersectional work has been narrowly focused, and even that literature is small. Very little is known about how labor market outcomes vary over the lifecycle for the non-white population. Similarly, while there is considerable work focused on hiring outcomes for entry-level workers and recent high school and college graduates, less is known about different parts of the education spectrum or how outcomes differ for experienced workers. Recent advances in working longer have added to this literature (e.g. Goldin and Katz

[30] See, for example https://www.fastcompany.com/3057631/how-blind-recruitment-works-and-why-you-should-consider *last accessed 11/13/18*.

2018, Farber et al. 2018) but there is room for much more work to be done. We hope that the results of this paper encourage future work in this area.

Table 1: Summary Statistics

	Mean	SD
Resume Characteristics		
Female	0.50	0.50
Black	0.09	0.29
Hispanic	0.13	0.34
Age	56.20	11.78
Participant Characteristics		
Female	0.56	
White	0.89	
Asian	0.07	
Black	0.05	
Hispanic	0.15	
MA student	0.38	
PhD student	0.01	
Upper division	0.38	
Lower division	0.23	
Business	0.76	
Government	0.13	
Social Science	0.05	
Humanities	0.06	
Age	21.98	2.84
Ratings		
Likert (1-7)	4.63	1.39
Eye-tracking		
Seconds spent: total	16.24	10.17
outside	3.03	3.78
employment history	4.87	5.72
name	0.17	0.52
high school	1.20	1.77
years employed	0.48	1.09
graduation year	0.02	0.14
other	0.22	0.55
education	0.21	0.46

Note: 5,960 resumes for the non-eyetracking statistics. 4,909 resumes for the eyetracking statistics other than seconds spent total which has 5,615 resumes.

TABLE 2: Likert Scale Differences by Race (t-test)

	Mean (1-7)	N	Difference	p (two-sided)
Age 45 and under				
White	4.72	1293		
Black	4.44	110	0.28	0.042
Entire Sample				
White	4.63	5425		
Black	4.60	535	0.03	0.616

Note: Entire sample includes ages 36-76.

Table 3: Effect of black names with and without age interactions

	Likert rating (1-7)			
	All		Female	Male
	(1)	(2)	(3)	(4)
black name	-0.0292	-6.0465***	-5.2598***	-6.9505***
	(0.0591)	(1.3866)	(1.9180)	(2.1606)
black*age		0.2241***	0.2006***	0.2515***
		(0.0500)	(0.0694)	(0.0786)
black*age squared		-0.0020***	-0.0018***	-0.0022***
		(0.0004)	(0.0006)	(0.0007)
age	-0.0145	-0.0349**	-0.0471*	-0.0232
	(0.0157)	(0.0170)	(0.0239)	(0.0225)
age squared	0.0001	0.0003*	0.0004*	0.0002
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Observations	5,960	5,960	2,982	2,978

Note: * sig at 10%, ** sig at 5% and *** sig at 1% level. Standard errors are clustered on participant. Female and Male refer to the gender indicated on the resume.

Table 4a: Younger White Applicants

Time t-1 \ Time t	Name	Address	Years Emp	Emp Hist	Grad Year	Education	Other
Name	.	0.68	0.02	0.26	0.00	0.02	0.02
Address	0.15	.	0.13	0.66	0.01	0.04	0.02
Years Emp	0.01	0.20	.	0.73	0.02	0.02	0.02
Emp Hist	0.04	0.38	0.39	.	0.01	0.11	0.08
Grad Year	0.00	0.18	0.18	0.19	.	0.31	0.13
Education	0.02	0.18	0.09	0.35	0.09	.	0.27
Other	0.01	0.24	0.09	0.34	0.04	0.27	.

Table 4c: Younger Black Applicants

Time t-1 \ Time t	Name	Address	Years Emp	Emp Hist	Grad Year	Education	Other
Name	.	0.73	0.07	0.20	0.00	0.00	0.00
Address	0.07	.	0.06	0.78	0.02	0.04	0.03
Years Emp	0.00	0.17	.	0.77	0.00	0.03	0.04
Emp Hist	0.02	0.39	0.38	.	0.02	0.11	0.07
Grad Year	0.00	0.14	0.00	0.14	.	0.57	0.14
Education	0.00	0.21	0.04	0.36	0.07	.	0.32
Other	0.00	0.23	0.14	0.32	0.00	0.32	.

Table 4b: Older White Applicants

Time t-1 \ Time t	Name	Address	Years Emp	Emp Hist	Grad Year	Education	Other
Name	.	0.69	0.04	0.22	0.00	0.03	0.02
Address	0.14	.	0.13	0.65	0.01	0.05	0.03
Years Emp	0.01	0.19	.	0.75	0.01	0.03	0.02
Emp Hist	0.04	0.35	0.41	.	0.01	0.12	0.08
Grad Year	0.02	0.16	0.19	0.18	.	0.38	0.07
Education	0.03	0.18	0.07	0.36	0.06	.	0.31
Other	0.03	0.22	0.07	0.39	0.02	0.27	.

Table 4d: Older Black Applicants

Time t-1 \ Time t	Name	Address	Years Emp	Emp Hist	Grad Year	Education	Other
Name	.	0.66	0.01	0.27	0.00	0.02	0.04
Address	0.12	.	0.13	0.66	0.01	0.05	0.03
Years Emp	0.02	0.18	.	0.77	0.00	0.02	0.02
Emp Hist	0.04	0.37	0.37	.	0.01	0.14	0.08
Grad Year	0.00	0.06	0.12	0.18	.	0.47	0.18
Education	0.01	0.16	0.07	0.47	0.04	.	0.24
Other	0.01	0.21	0.14	0.42	0.05	0.17	.

Table 5: Effect of resume items on Likert ratings, total time spent, and time spent on area of interest

	Likert ratings			Time spent viewing resume			Time spent on area of interest		
item =	Computer	Yrs Job Gap	Clerical Exp	Computer	Yrs Job Gap	Clerical Exp	Computer	Yrs Job Gap	Clerical Exp
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel I: Younger								
black*item	**1.0363*** **	**0.1535** **	**-1.2545*** **	**10.1643** **	**0.0664**	**-7.7303** **	**0.1457**	**-0.0084**	**-0.7980**
	(0.2629)	**(0.0707)**	**(0.4626)**	**(4.2919)**	**(0.3385)**	**(3.2963)**	**(0.5439)**	**(0.1490)**	**(1.9206)**
blackname	-0.3439***	-0.4335***	0.8437*	-1.6271*	-1.0248	5.8828*	-0.0112	-0.4335	0.2222
	(0.1318)	(0.1383)	(0.4372)	(0.8882)	(0.9460)	(3.2181)	(0.0320)	(0.5207)	(1.9024)
item	0.3441***	-0.1560***	1.5621***	1.5907*	-0.7068***	5.2938***	0.3109***	-0.3122***	1.9348***
	(0.1309)	(0.0207)	(0.1434)	(0.9488)	(0.1066)	(0.9107)	(0.0939)	(0.0687)	(0.5243)
Observations	1,403	1,403	1,403	1,332	1,332	1,332	1,183	1,183	1,183
	Panel II: Older								
black*item	**-0.0008**	**0.0425**	**0.1305**	**1.3934**	**0.0626**	**-2.1519**	**-0.3649*** **	**-0.0173**	**-1.4588**
	(0.2239)	**(0.0416)**	**(0.3325)**	**(1.9357)**	**(0.2189)**	**(2.0410)**	**(0.1062)**	**(0.1283)**	**(0.9357)**
blackname	0.0441	-0.0123	-0.1455	-0.2540	-0.2916	1.6181	-0.0165	-0.2205	1.0440
	(0.0699)	(0.0788)	(0.3229)	(0.4077)	(0.4641)	(1.9832)	(0.0231)	(0.2972)	(0.8677)
item	0.3025***	-0.1478***	1.4748***	1.8610***	-0.7844***	5.1359***	0.3453***	-0.3488***	2.2575***
	(0.0691)	(0.0119)	(0.0822)	(0.5345)	(0.0777)	(0.4666)	(0.0721)	(0.0473)	(0.2549)
Observations	4,557	4,557	4,557	4,283	4,283	4,283	3,726	3,726	3,726

Notes: * sig at 10%, ** sig at 5% and *** sig at 1% level. Controls include age and age squared. Standard errors clustered on participant. Younger includes ages 36-45. Observation numbers decrease because of lack of sensitivity in the eye-tracker.

Table 6: Robustness Checks for Likert Ratings in Table 3

	Original	Participant effects	Normalized Y	No Hispanics	Under 66	First Name SES
	(1)	(2)	(3)	(4)	(5)	(6)
black name	-6.0465***	-5.0385***	-3.8699***	-6.7031***	-8.4834***	-6.1227***
	(1.3866)	(1.3481)	(0.9971)	(1.4650)	(2.7015)	(1.3842)
black*age	0.2241***	0.1872***	0.1433***	0.2442***	0.3268***	0.2240***
	(0.0500)	(0.0497)	(0.0359)	(0.0529)	(0.1086)	(0.0499)
black*age squared	-0.0020***	-0.0017***	-0.0013***	-0.0021***	-0.0030***	-0.0020***
	(0.0004)	(0.0004)	(0.0003)	(0.0005)	(0.0011)	(0.0004)
age	-0.0349**	-0.0324**	-0.0262**	-0.0393**	-0.0271	-0.0349**
	(0.0170)	(0.0164)	(0.0127)	(0.0187)	(0.0348)	(0.0170)
age squared	0.0003*	0.0003*	0.0002*	0.0003**	0.0002	0.0003*
	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0003)	(0.0002)
Observations	5,960	5,960	5,960	5,170	4,349	5,960

Notes: * sig at 10%, ** sig at 5% and *** sig at 1% level. Column 2 includes participant fixed effects. Column 3 normalizes "respval" variable for each participant's ratings. Column 4 removes all Hispanic last names. Column 5 limits universe to ages under 66. Column 6 includes a variable for the perceived SES of the first name from Barlow and Lahey (2018).

Appendix Table 1: What goes into quality measure

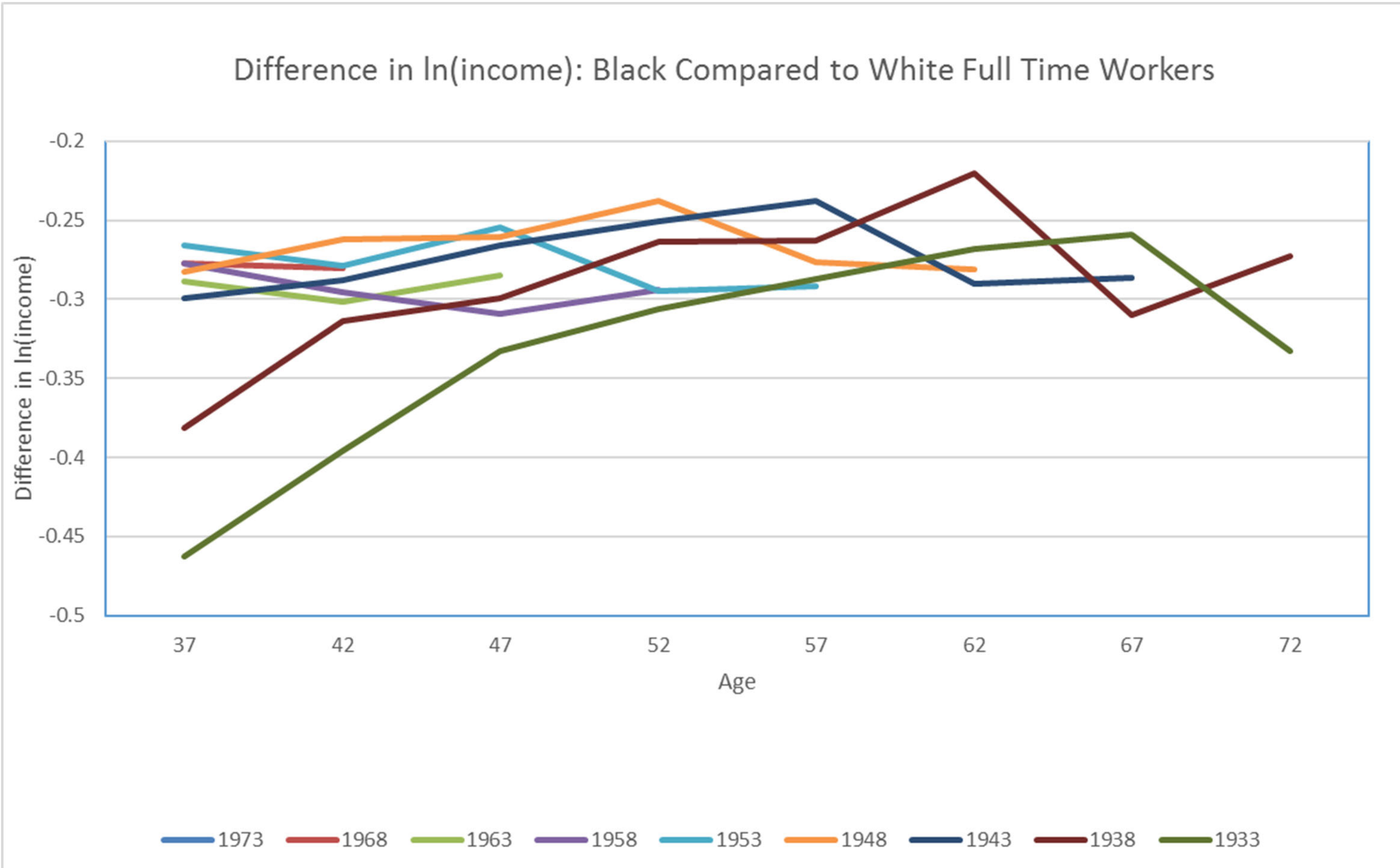
	Likert (1-7)
Work Gap	-0.1388
	(0.1277)
Basic army training	0.1263
	(0.1220)
Army training	0.1539
	(0.1054)
Phlebotomy training	0.0136
	(0.1083)
CNA/HHA	0.0210
	(0.1050)
Medical transcription	0.0677
	(0.0990)
Clerical Arts	0.6998***
	(0.0965)
Computer and Clerical	0.3304***
	(0.1098)
Manpower computer	0.1881*
	(0.1030)
Computer certificate	0.1616
	(0.1210)
Project Bread volunteer	0.2448**
	(0.0944)
Hope Worldwide volunteer	0.1937**
	(0.0852)
PTA volunteer	0.1298
	(0.0860)
Neighborhood Association	-0.0547
	(0.0923)
Flexible	0.1628**
	(0.0716)
Embrace change	0.0685
	(0.0717)
Hotmail	0.0587
	(0.0398)
Gmail	-0.0276
	(0.0393)
number of jobs	0.7492***
	(0.1788)
# of jobs squared	-0.0955***
	(0.0163)
years of employment	0.2644***
	(0.0584)
yrs employment squared	-0.0188***
	(0.0052)
R2	0.359
N	5,960

* sig at 10%, ** sig at 5% and *** sig at 1% level.

Additional controls include: 231 separate jobs, 46 separate high schools, 72 separate addresses.

Appendix Table 2: The Effect of Participant Characteristics on Likert Ratings (1-7)

Xistic =	Hiring Exper.	HR Exper.	Age 25 or older	Female	MPA	Hispanic
	(1)	(2)	(3)	(4)	(5)	(6)
black*Xistic	**3.0105**	**3.3196**	**-4.1880**	**-4.6644***	**3.3791**	**0.3453**
	(3.2689)	**(4.5421)**	**(3.2306)**	**(2.7941)**	**(3.4070)**	**(2.7655)**
black*age*Xistic	**-0.1019**	**-0.1444**	**0.1486**	**0.1758***	**-0.0953**	**-0.0155**
	(0.1174)	**(0.1608)**	**(0.1206)**	**(0.1009)**	**(0.1226)**	**(0.0985)**
black*age^2*Xistic	**0.0008**	**0.0014**	**-0.0013**	**-0.0015***	**0.0007**	**0.0003**
	(0.0010)	**(0.0014)**	**(0.0011)**	**(0.0009)**	**(0.0011)**	**(0.0008)**
Xistic*age	0.0126	0.0067	0.0036	-0.0231	0.0305	-0.0535
	(0.0422)	(0.0696)	(0.0446)	(0.0337)	(0.0562)	(0.0490)
Xistic*age^2	-0.0001	-0.0001	-0.0001	0.0002	-0.0003	0.0005
	(0.0004)	(0.0006)	(0.0004)	(0.0003)	(0.0005)	(0.0004)
Participant Xistic	-0.4803	-0.1441	0.0395	0.6763	-1.2593	1.4419
	(1.1674)	(1.9403)	(1.2727)	(0.9199)	(1.5191)	(1.2959)
blackname	-6.7593***	-6.4291***	-5.5225***	-3.4163	-6.4596***	-6.1875***
	(1.5996)	(1.4711)	(1.5425)	(2.2404)	(1.5281)	(1.5744)
black*age	0.2482***	0.2411***	0.2058***	0.1249	0.2351***	0.2301***
	(0.0576)	(0.0532)	(0.0554)	(0.0815)	(0.0551)	(0.0569)
black*age^2	-0.0022***	-0.0022***	-0.0018***	-0.0011	-0.0021***	-0.0021***
	(0.0005)	(0.0005)	(0.0005)	(0.0007)	(0.0005)	(0.0005)
age	-0.0379**	-0.0356**	-0.0358*	-0.0235	-0.0358**	-0.0269
	(0.0189)	(0.0175)	(0.0188)	(0.0244)	(0.0178)	(0.0182)
age^2	0.0003*	0.0003*	0.0003*	0.0002	0.0003*	0.0002
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Observations	5,960	5,960	5,960	5,960	5,960	5,960



Appendix Figure 1

	Automatic less time on black resumes	Cognitively Engaged more time on black resumes	same time
prefer white resumes	negative heuristics vs black or taste-based for white	justifying choice vs. black (could be taste-based or statistical)	implicit biases vs black
prefer black resumes	positive heuristics for black or affirmative action heuristics	counteracting neg. stereotypes or justifying choice for black	implicit biases for black
no preference	positive heuristics for black or affirmative action heuristics	counteracting neg. stereotypes or justifying choices	no discrimination

Appendix Figure 2

Patricia Edwards

8830 Maple Rapids Lane
Humble, TX 77338
(832) 894-9682
PatriciaEdwards86@hotmail.com

Employment History

2011–now	Clerical, Stage Stores, Inc., Jacksonville, Texas Possess years of clerical experience in all aspects of general office/customer service.
2005–2011	Accounting Principals, Data Entry Clerk, Houston, TX Entered timesheets and billing information. Assisted expense department with filing.
2003–2005	Security officer/first aid, Burns Securitas, Texas Secure premises, first aid responder, clerical

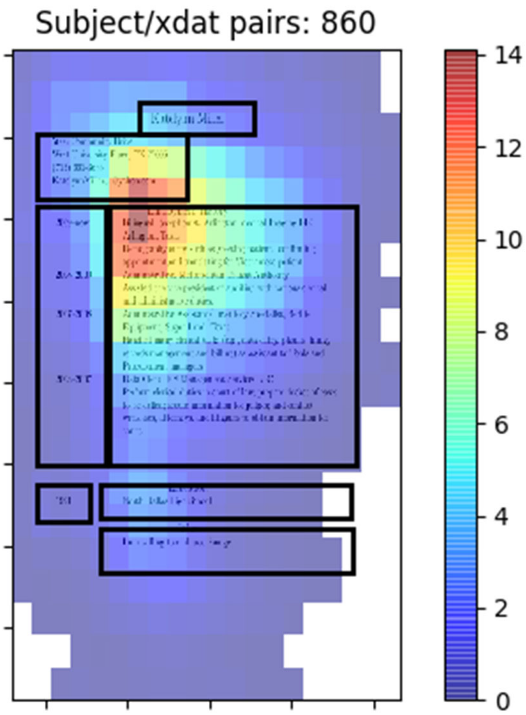
Education

1981	Pearland High School

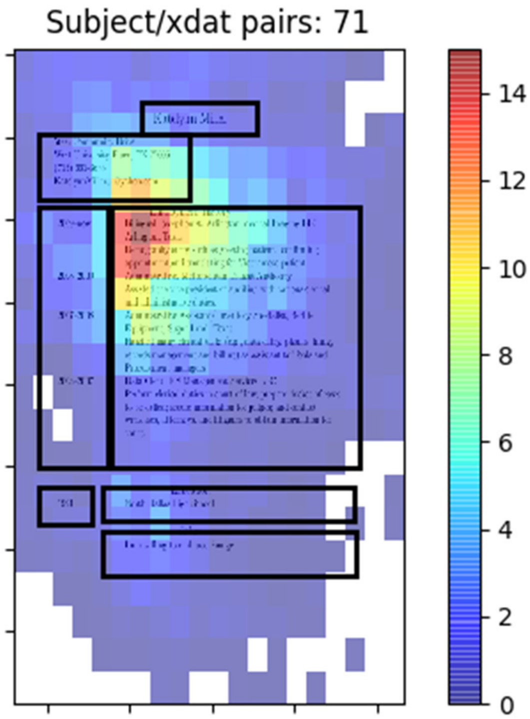
Other

I am willing to embrace change.

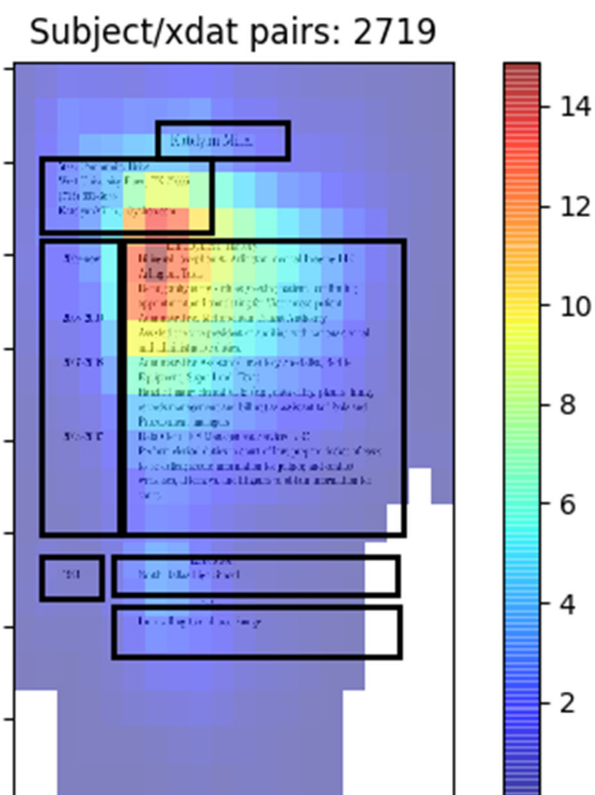
Appendix Figure 3



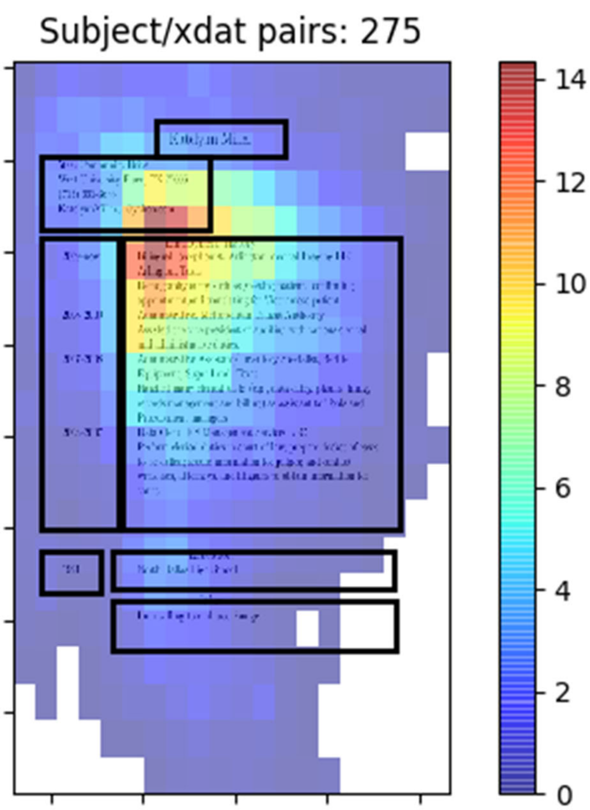
Younger White Resumes



Younger Black Resumes

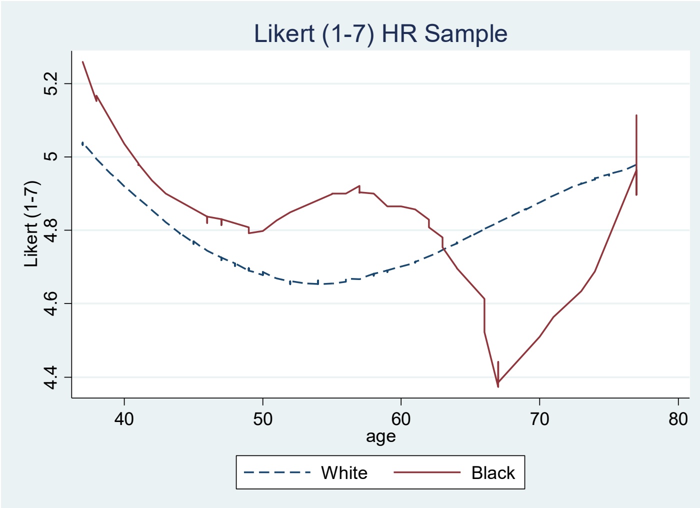


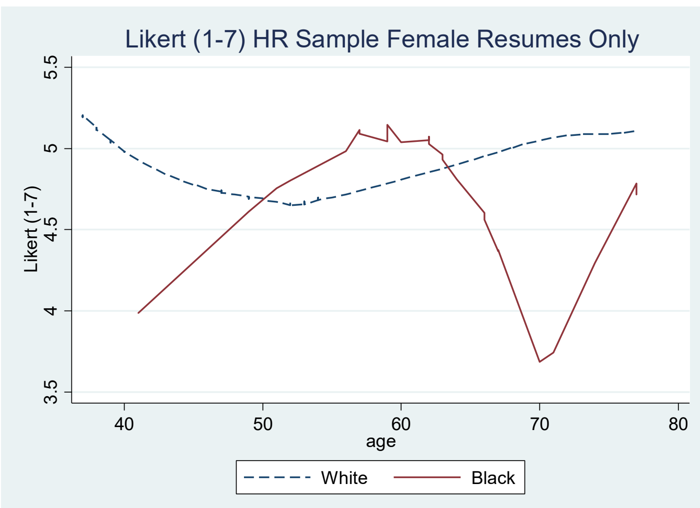
Older White Resumes

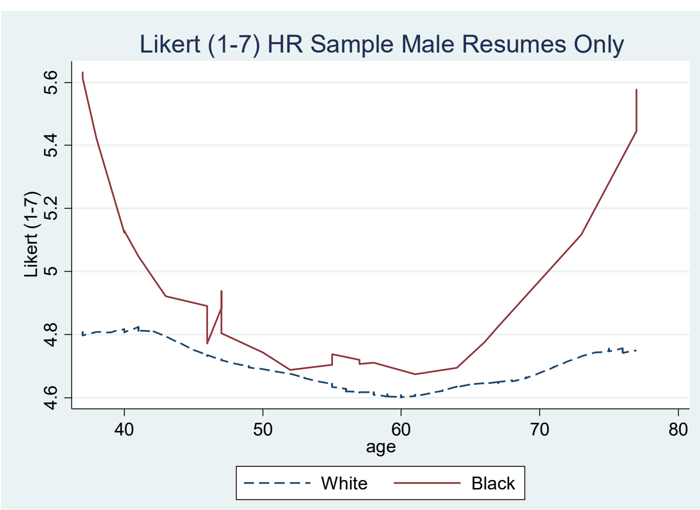


Older Black Resumes

Appendix Figure 4







Appendix Figures 5

References

Agan, Amanda, and Sonja Starr. 2018. "Ban the Box, Criminal Records, and Racial Discrimination: A Field Experiment." *The Quarterly Journal of Economics* 133 (1):191-235.

Aigner, Dennis J., and Glen G. Cain. 1977. "Statistical Theories Of Discrimination In Labor Markets." *Industrial and Labor Relations Review* 30 (2):175-87.

Altonji, Joseph, and Rebecca Blank. 1999. "Chapter 48 Race and gender in the labor market." *Handbook of Labor Economics* 3 (Part C):3143-259.

Altonji, Joseph G. and Charles R. Pierret. 2001. "Employer Learning and Statistical Discrimination." *The Quarterly Journal of Economics* (1):313-50.

Arcidiacono, Peter, Patrick Bayer, and Aurel Hizmo. 2010. "Beyond Signaling and Human Capital: Education and the Revelation of Ability." *American Economic Journal: Applied Economics* (4):76-104.

Aura, Saku, and Gregory D. Hess. 2010. "What's in a Name?" *Economic Inquiry* 48 (1):214-27.

Autor, David H. and David Scarborough. 2008. "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments." *The Quarterly Journal of Economics* (1):219-77.

Barlow, M Rose, and Joanna N Lahey. 2018. "What Race Is Lacey? Intersecting Perceptions of Racial Minority Status and Social Class." *Social Science Quarterly*.

Bartoš, Vojtech, Michal Bauer, Julie Chytilova, and Filip Matejka. 2016. "Attention Discrimination: Theory and Field Experiments with Monitoring Information Acquisition." *American Economic Review* 106 (6):1437-75.

Becker, Gary S. 1971. *The Economics of Discrimination.* ed. by Gary S. Becker Chicago : University of Chicago Press, [1971] 2d ed.

Bertrand, Marianne, and Sendhil Mullainathan. 2004. "Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *The American Economic Review* 94 (4):991-1013.

Bezrukova, Katerina, Chester S. Spell, Jamie L. Perry, and Karen A. Jehn. 2016. "A meta-analytical integration of over 40 years of research on diversity training evaluation." *Psychological Bulletin* 142 (11):1227-74.

Bickenbach, Frank, and Eckhardt Bode. 2003. "Evaluating the Markov Property in Studies of Economic Convergence." *International Regional Science Review* 26 (3):363-92.

Bogardus, Emory Stephen. 1933. "A Social Distance Scale." *Sociology & Social Research*.

Bound, John, and Richard B. Freeman. 1992. "What went wrong? The erosion of relative earnings and employment among young black men in the 1980s." *Quarterly Journal of Economics* 107 (1):201-232.

Card, David, and Alan B Krueger. 1992. "School quality and black-white relative earnings: A direct assessment." *The Quarterly Journal of Economics* 107 (1):151-200.

Chandon, Pierre, J. Wesley Hutchinson, T. Bradlow Eric, and H. Young Scott. 2009. "Does In-Store Marketing Work? Effects of the Number and Position of Shelf Facings on Brand Attention and Evaluation at the Point of Purchase." *Journal of Marketing* 73 (6):1-17.

Charles, Camille Zubrinsky. 2003. "The Dynamics of Racial Residential Segregation." *Annual Review of Sociology* 29:167-207.

Charles, Kerwin Kofi and Jonathan Guryan. 2008. "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination." *Journal of Political Economy* (5):773-809.

Cleveland, Jeanette N., and Andrew H. Berman. 1987. "Age Perceptions of Jobs: Agreement between Samples of Students and Managers." *Psychological Reports 61: 565–66* 61 (Journal Article):565-66.

Crenshaw, Kimberle. 1989. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *Univeristy of Chicago Legal Forum* 1989:139-68.

Dawkins, Casey J, Qing Shen, and Thomas W Sanchez. 2005. "Race, space, and unemployment duration." *Journal of Urban Economics* 58 (1):91-113.

Devine, Patricia G., Patrick S. Forscher, Anthony J. Austin, and William T. L. Cox. 2012. "Long-term reduction in implicit race bias: A prejudice habit-breaking intervention." *Journal of Experimental Social Psychology* 48 (6):1267-78.

Doleac, Jennifer L, and Benjamin Hansen. 2016. "Does "ban the box" help or hurt low-skilled workers? Statistical discrimination and employment outcomes when criminal histories are hidden." National Bureau of Economic Research.

Drydakis, Nick, Peter MacDonald, Vangelis Chiotis, and Laurence Somers. 2018. "Age discrimination in the UK labour market. Does race moderate ageism? An experimental investigation." *Applied Economics Letters* 25 (1):1-4.

Ellen, Ingrid Gould, and Margery Austin Turner. 1997. "Does neighborhood matter? Assessing recent evidence." *Housing Policy Debate* 8 (4):833-66.

Ewens, Michael, Bryan Tomlin, and Liang Choon Wang. 2014. "Statistical discrimination or prejudice? A large sample field experiment." *Review of Economics and Statistics* 96 (1):119-34.

Farber, Henry S, Chris M Herbst, Dan Silverman, and Till Von Wachter. 2018. "Whom Do Employers Want? The Role of Recent Employment and Unemployment Status and Age." NBER Working Paper 24605.

Feng, Gary. 2011. "Eye Tracking: A Brief Guide for Developmental Researchers." *Journal of Cognition and Development* 12 (1):1-11.

Figlio, David N. 2005. "Names, Expectations and the Black-White Test Score Gap." NBER Working Paper 11195.

Firebaugh, Glenn, Francesco Acciai, Aggie J. Noah, Christopher Prather, and Claudia Nau. 2014. "Why Lifespans Are More Variable among Blacks Than among Whites in the United States." *Demography* 51 (6):2025-45.

Fryer Jr., Roland G., and Steven D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *The Quarterly Journal of Economics* (3):767.

Fryer Jr, Roland G., Jacob K. Goeree, Charles A. Holt, and Michael Watts. 2005. "Experience-Based Discrimination: Classroom Games." *Journal of Economic Education* 36 (2):160-70.

Giuliano, Laura, David Levine, and Jonathan Leonard. 2011. "Racial Bias in the Manager-Employee Relationship: An Analysis of Quits, Dismissals, and Promotions at a Large Retail Firm." *Journal of Human Resources* 46 (1):26-52.

Giuliano, Laura, and Michael R. Ransom. 2013. "Manager ethnicity and employment segregation." *Industrial and Labor Relations Review* 66 (2):346-79.

Glover, Dylan, Amanda Pallais, and William Pariente. 2017. "Discrimination as a Self-Fulfilling Prophecy: Evidence from French Grocery Stores." *Quarterly Journal of Economics* 132 (3):1219-60.

Goldin, Claudia, and Lawrence F. Katz, eds. 2018. *Women Working Longer: Increased Employment at Older Ages*. Chicago, IL: University of Chicago Press.

Greenwald, Anthony G., Brian A. Nosek, and Mahzarin R. Banaji. 2003. "Understanding and Using the Implicity Association Test: I. An Improved Scoring Algorithm." *Journal of Personality and Social Psychology* 85 (2):197-216.

Henkens, Kène. 2005. "Stereotyping Older Workers and Retirement: The Managers' Point of View." *Canadian Journal on Aging* 24 (4):353-66.

Holmqvist, Kenneth, Marcus Nystrom, Richard Andersson, Richard Dewhurst, Halszka Jarodzka, and Joost Van de Weijer. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. New York: Oxford University Press.

Holzer, Harry J. 1998. "Employer Skill Demands and Labor Market Outcomes of Blacks and Women." *Industrial & Labor Relations Review* 52 (1):82-98.

Holzer, Harry J. , Steven Raphael, and Michael A. Stoll. 2006. "Perceived Criminality, Criminal Background Checks, and the Racial Hiring Practices of Employers." *The Journal of Law and Economics* 49 (2):451-80.

Immergluck, Daniel. 1998. *Neighborhood Jobs, Race, and Skills: Urban Unemployment and Commuting*. New York: Garland Publishing.

Johnson, Emily, and Joanna Lahey. 2011. "The Resume Characteristics Determining Job Interviews for Middle-Aged Women Seeking Entry-Level Employment." *Journal of Career Development* 38 (4):310-30.

Jones, Kristen P., Isaac E. Sabat, Eden B. King, Afra Ahmad, Tracy C. McCausland, and Tiffani Chen. 2017. "Isms and schisms: A meta-analysis of the prejudice-discrimination relationship across racism, sexism, and ageism." *Journal of Organizational Behavior* 38 (7):1076-110.

Kirby, James B., and Toshiko Kaneda. 2010. "Unhealthy and Uninsured: Exploring Racial Differences in Health and Health Insurance Coverage Using a Life Table Approach." *Demography* 47 (4):1035-51.

Knoepfle, Daniel T., Joseph Tao-yi Wang, and Colin F. Camerer. 2009. "STUDYING LEARNING IN GAMES USING EYE-TRACKING." *Journal of the European Economic Association* 7 (2-3):388-98.

Lahey, Joanna N. 2008. "Age, Women, and Hiring - An experimental Study." *Journal of Human Resources* 43 (1):30-56.

———. 2018. "Understanding Why Black Women Are Not Working Longer." In *Women Working Longer: Increased Employment at Older Ages*, ed. C. Goldin and L. F. Katz. Chicago, IL: University of Chicago Press.

Lahey, Joanna N., and Ryan A. Beasley. 2009. "Computerizing Audit Studies." *Journal of Economic Behavior and Organization* 70 (3):508-14.

Lahey, Joanna N., and Douglas Oxley. 2016. "The Power of Eye Tracking in Economics Experiments." *American Economic Review* 106 (5):309-13.

———. "Patterns of Age Discrimination in Hiring" *mimeo.*

Lang, Kevin, and Jee-Yeon K. Lehmann. 2012. "Racial Discrimination in the Labor Market: Theory and Empirics." *Journal of Economic Literature* (4):959.

Lang, Kevin, and Michael Manove. 2011. "Education and Labor Market Discrimination." *The American Economic Review* (4):1467.

Lang, Kevin, Michael Manove, and William T. Dickens. 2005. "Racial Discrimination in Labor Markets with Posted Wage Offers." *The American Economic Review* (4):1327.

Levitt, Steven D., and Stephen J. Dubner. 2005. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything*. New York: William Morrow and Company.

Lhamon, Catherine E. , and Jocelyn Samuels. 2014. "Dear Colleague Letter on the Nondiscriminatory Administration of School Discipline." Ed. U.S. Department of Education and U.S. Department of Justice.

Lieberson, Stanley, and Eleanor O. Bell. 1992. "Children's First Names: An Empirical Study of Social Taste." *American Journal of Sociology* (3):511.

Lieberson, Stanley, and Kelly S. Mikelson. 1995. "Distinctive African American names: An experimental, historical, and linguistic analysis of innovation." *American Sociological Review* 60 (6):928-46.

Lindner, Nicole M., Alexander Graser, and Brian A. Nosek. 2014. "Age-Based Hiring Discrimination as a Function of Equity Norms and Self-Perceived Objectivity." *PLoS ONE* 9 (1):1-6.

Lohse, Gerald L. 1997. "Consumer Eye Movement Patterns on Yellow Pages Advertising." *Journal of Advertising* 26 (1):61-73.

Madrian, Brigitte C., and Lars John Lefgren. 2000. "An Approach to Longitudinally Matching Current Population Survey (CPS) Respondents." *Journal of Economic and Social Measurement* 26 (1):31-62.

Margo, Robert A. 1986. "Educational Achievement In Segregated School Systems: The Effects of 'Separate-but-Equal'." *The American Economic Association* 76 (4):794-801.

———. 1990. *Race and Schooling in the South, 1880-1950: An Economic History*: University of Chicago Press.

Masters, Ryan K., Robert A. Hummer, Daniel A. Powers, Audrey Beck, Shih-Fan Lin, and Brian Karl Finch. 2014. "Long-Term Trends in Adult Mortality for U.S. Blacks and Whites: An Examination of Period- and Cohort-Based Changes." *Demography* 51 (6):2047-73.

Maughan, Lizzie, Sergei Gutnikov, and Rob Stevens. 2007. "Like more, look more. Look more, like more: The evidence from eye-tracking." *Journal of Brand Managment* 14 (4):335-42.

Mehrabian, A., and M. Piercy. 1993. "Positive or negative connotations of unconventionally and conventionally spelled names." *Journal of Social Psychology* 133 (4):445-51.

Neumark, David, Ian Burn, and Patrick Button. N.d. "Is It Harder for Older Workers to Find Jobs? New and Improved Evidence from a Field Experiment." *Journal of Political Economy*. Forthcoming.

Norton, Michael I., Joseph A. Vandello, and John M. Darley. 2004. "Casuistry and Social Category Bias." *Journal of Personality and Social Psychology* 87 (6):817-31.

Nosek, Brian A., Anthony G. Greenwald, and Mahzarin R. Banaji. 2007. "The Implicit Association Test at Age 7: A Methodological and Conceptual Review." In *Social psychology and the unconscious: The automaticity of higher mental processes.*New York, NY: Psychology Press.

Nunley, John, Adam Pugh, Nicholas Romero, and R. A. Seals. 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):1093-125.

Oyer, Paul, and Scott Schaefer. 2002. "Litigation Costs and Returns to Experience." *The American Economic Review* (3):683.

Pais, Jeremy. 2014. "Cumulative Structural Disadvantage and Racial Health Disparities: The Pathways of Childhood Socioeconomic Influence." *Demography* 51 (5):1729-53.

Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *The American Economic Review* 62 (4):659-61.

Pieters, Rik, Edward Rosbergen, and Michel Wedel. 1999. "Visual Attention to Repeated Print Advertising: A Test of Scanpath Theory." *Journal of Marketing Research (JMR)* 36 (4):424-38.

Pieters, Rik, and Michel Wedel. 2004. "Attention Capture and Transfer in Advertising: Brand, Pictorial, and Text-Size Effects." *Journal of Marketing* 68 (2):36-50.

Pinkston, J. C. 2006. "A test of screening discrimination with employer learning." *Industrial and Labor Relations Review* 59 (2):267-84.

Quillian, Lincoln, Devah Pager, Ole Hexel, and Arnfinn H Midtbøen. 2017. "Meta-analysis of field experiments shows no change in racial discrimination in hiring over time." *Proceedings of the National Academy of Sciences* 114 (41):10870-5.

Reutskaja, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel. 2011. "Search Dynamics in Consumer Choice under Time Pressure: An Eye- Tracking Study." *The American Economic Review* 101 (2):900-26.

Rich, Judith. 2014. "What Do Field Experiments of Discrimination in Markets Tell Us? A Meta Analysis of Studies Conducted Since 2000." IZA Discussion Paper No. 8584.

Ritter, Joseph A. and Lowell J. Taylor. 2011. "Racial Disparity in Unemployment." *The Review of Economics and Statistics* (1):30-42.

Russo, J. Edward, and France Leclerc. 1994. "An Eye-Fixation Analysis of Choice Processes for Consumer Nondurables." *Journal of Consumer Research* 21 (2):274-90.

Sherman, Jeffrey W., Bertram Gawronski, and Yaacov Trope. 2014. *Dual-Process Theories of the Social Mind*. Edited by B. G. Jeffrey W. Sherman, Yaacov Trope. New York: The Guilford Press.

Sigall, Harold, and Richard Page. 1971. "Current stereotypes: A little fading, a little faking." *Journal of Personality and Social Psychology* 18 (2):247-55.

Smith, James P., and Finis R. Welch. 1989. "Black Economic Progress After Myrdal." *Journal of Economic Literature* 27 (2):519-64.

Tews, Michael, Kathryn Stafford, and J. Bruce Tracey. 2011. "What matters most? The perceived importance of ability and personality for hiring decisions." *Cornell Hospitality Quarterly* 52 (2):94-101.

Uhlmann, Eric Luis and Geoffrey L. Cohen. 2005. "Constructed Criteria: Redefining Merit to Justify Discrimination." *Psychological Science* (6):474-80.

Wang, Joseph Tao-yi, Michael Spezio, and Colin F. Camerer. 2010. "Pinocchio's Pupil: Using Eyetracking and Pupil Dilation to Understand Truth Telling and Deception in Sender-Receiver Games." *The American Economic Review* (3):984-1007.

Wicker, Allan W. 1969. "Attitudes versus Actions: The Relationship of Verbal and Overt Behavioral Responses to Attitude Objects." *Journal of Social Issues* 25 (4):41-78.

Yinger, John. 1986. "Measuring Racial Discrimination with Fair Housing Audits: Caught in the Act." *American Economic Review* 76 (5):881.