11-8-2023

Minimum Wages and Racial Discrimination in Hiring: Evidence from a Field Experiment

Alec Brandon
Johns Hopkins University, alec.brandon@gmail.com

Justin E. Holz
University of Michigan, holzj@umich.edu

Andrew Simon
The University of Chicago, Australian National University, Research School of Economics, andrewsimon@uchicago.edu

Haruka Uchida
The University of Chicago, uchida@uchicago.edu

Upjohn Institute working paper ; 23-389

Citation

Brandon, Alec, Justin E. Holz, Andrew Simon, and Haruka Uchida. 2023. "Minimum Wages and Racial Discrimination in Hiring: Evidence from a Field Experiment." Upjohn Institute Working Paper 23-389. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. https://doi.org/10.17848/wp23-389

W.E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH

Researching the causes and consequences of unemployment

Minimum Wages and Racial Discrimination in Hiring: Evidence from a Field Experiment

Upjohn Institute Working Paper 23-389

Alec Brandon
Johns Hopkins University
Email: **alec.brandon@gmail.com**

Justin E. Holz
University of Michigan
Email: **holzj@umich.edu**

Andrew Simon
University of Chicago, ANU, RSE
Email: **andrewsimon@uchicago.edu**

Haruka Uchida
University of Chicago
Email: **uchida@uchicago.edu**

November 2023

ABSTRACT

When minimum wages increase, employers may respond to the regulatory burdens by substituting away from disadvantaged workers. We test this hypothesis using a correspondence study with 35,000 applications around ex-ante uncertain minimum wage increases in three U.S. states. Before the increases, applicants with distinctively Black names were 19 percent less likely to receive a callback than equivalent applicants with distinctively white names. Announcements of minimum wage hikes substantially reduce callbacks for all applicants but shrink the racial callback gap by 80 percent. Racial inequality decreases because firms disproportionately reduce callbacks to lower-quality white applicants who benefited from discrimination under lower minimum wages.

JEL Classification Codes: J23, C93, J71, J15

Key Words: minimum wage, correspondence study, racial discrimination

Acknowledgments: This research was conducted with approval from the University of Chicago Institutional Review Board, IRB18-0873 and preregistered under AEARCTR-0003333. We thank Dan Alexander, Bocar Ba, Dan Black, Chris Blattman, John Bound, Leonardo Bursztyn, Jeffery Clemens, Steven Durlauf, Yana Gallen, Sara Heller, Nicole Holz, Rafael Jiménez Durán, Attila Lindner, John List, Jonathan Meer, Anant Nyshadham, David Philips, Evan Rose, Mattie Toma, Gonzalo Vazquez-Bare, Udayan Vaidya, Derek Wu, Jaren Ye, and audiences at Advances in Field Experiments, the Institute for Research on Poverty's Summer Research Workshop at the University of Wisconsin IRP SRW), the Online Seminar on Discrimination and Disparities, the University of Michigan, Harris School of Public Policy, the University of Chicago Experimental Workshop, Queens University, the University of Missouri, the University of Texas at Dallas, Bentley University, the University of Nebraska-Lincoln, and the University of Boston, for helpful feedback. Jacob Benge, Noelani Bernal, Emilia Van Buskirk, Zack Fattore, Elliot Grays, Lawrence Guindine, David Liu, Jim Liu, Kelly Liu, Beatrice Del Negro, Naina Prasad, Diego Ruiz, Steven Shi, Alejandro Turriago, Josh Xu, and Daniel Zhu provided excellent research assistance. This research was funded in part by grants from the University of Chicago Experiments Initiative, the University of Michigan, Department of Economics and Rackham Graduate School, the Institute of Education Sciences grant R305B150012, and the W.E. Upjohn Institute for Employment Research.

1. Introduction

Twenty-seven million workers earned under $12 an hour in 2019 (Shrider et al., 2021). Minimum wages are popular policies meant to redistribute income to these workers and provide economic justice, especially racial economic justice (Smythe and Hsu, 2023), but critics argue that disemployment effects undermine this goal. A large literature evaluating these claims finds that hikes increase the average earnings of low-wage workers, with little effect on the number of low-wage jobs (Cengiz et al., 2019), which belies larger changes in hiring (Dube et al., 2016). Hiring changes raise an additional redistributional concern that minimum wage hikes disproportionately harm disadvantaged groups, especially those who face discrimination (Friedman, 1966; Stiglitz, 1973; Clemens et al., 2021).

We conduct a large-scale natural field experiment to evaluate how increases in the minimum wage affect racial discrimination in the hiring process. We sent nearly 35,000 fictitious job applications to low-wage job postings in three states before and after minimum wage increases were announced and in a contiguous state that did not change its policy. In each application, we randomized whether the applicant had a distinctively Black or white name, along with other characteristics such as educational attainment and unemployment duration.[1] In response, we observe whether an application receives a callback offering a job interview, a necessary step in the hiring process for many jobs. The data generated from our experiment allow us to estimate the causal effect of minimum wage hikes on racial discrimination in callbacks.

We estimate the effect of minimum wage hikes on racial discrimination against young Black men. While minimum wage hikes could affect discrimination in hiring against other races, age groups, and genders, unemployment data suggest that young Black men are the most likely to be on the margins of employment. When we launched our experiment in the summer of 2018, the unemployment rate in the U.S. for Black men 18 to 19 years old was 24.7

[1] See Jowell and Prescott-Clarke (1970), Bertrand and Mullainathan (2004), Kroft et al. (2013), and the papers reviewed in Bertrand and Duflo (2017) for other experiments using this approach.

percent, the highest for any demographic group (BLS, 2018). Relative to other demographic groups in the United States, this high unemployment rate suggests that young Black men are likely to bear a disproportionate share of the effects of minimum wage increases.

We find that minimum wage hikes eliminate the vast majority of the racial differences in callback rates. Before minimum wage hikes were announced, applicants with distinctly Black names were 3.2 percentage points (19 percent) less likely to receive a callback than equally qualified applicants with distinctly white names. Black applicants received eight callbacks for every ten that a white applicant received. After minimum wages are increased, this gap shrinks by 2.6 percentage points (80 percent). Black applicants received a little over nine callbacks for every ten that white applicants received. This effect occurs once a hike in the minimum wage is announced and persists for at least a year after it is enacted. These results suggest that, contrary to the concerns of Friedman (1966) and others, minimum wage increases do not exacerbate racial discrimination.

To understand why minimum wage increases shrink the racial callback gap, we consider the role of different types of racial discrimination. We start by developing an economic model of the relationship between minimum wages, hiring, and two forms of racial discrimination: taste-based (Becker, 1957) and statistical (Phelps, 1972; Arrow, 1973; Aigner and Cain, 1977). In the model, firms infer the quality of the worker using information from the resume characteristics and the applicant's race. Differences in the callback rates between Black and white applicants with identical resumes are either driven by differences in the firm's beliefs about productivity by race or by a distaste for working with Black employees. We show how these mechanisms jointly determine which types of workers are on the margin of receiving a callback and thereby affected by the policy change.

We estimate the model using data generated from the correspondence study to partially decompose taste-based and statistical discrimination, following Neumark (2012) and Neumark et al. (2019). We estimate the ratio of the variances of perceived applicant productivity by race, one form of statistical discrimination, as well as a composite term that reflects both taste-based discrimination and statistical discrimination in mean productivity. Since white applicants have less dispersed perceived productivity distribution and receive favoritism in hiring, a larger share of the initially marginal applicants are white, and they experience a larger decrease in callbacks after a minimum wage hike. The nature and extent of discrimination, along with the frequency of callbacks, jointly determine the change in the racial callback gap from minimum wage increases. More broadly, we provide a framework to understand how labor market policies affect racial disparities.

We additionally test whether the decrease in the racial callback gap can be attributed to explanations outside the scope of our model. For example, minimum wage hikes may cause firms to pay more attention to nonrace aspects of the resume, may change the composition of firms hiring workers, or may change the composition of the labor pool. Though these other explanations would not change the policy implication that minimum wages reduce racial disparities in hiring, the richness of our data allows us to address these potential issues and more cleanly isolate the mechanisms.

First, we use exogenous variation in other productivity signals to rule out changes in attention to nonrace resume characteristics. We find no evidence of differentially increasing returns to applicants' productivity by race, suggesting that firms are not paying more attention to these characteristics when the minimum wage increases. By showing similar effects of the minimum wage across establishments within the same parent company and within establishments, we can rule out changes in the composition of hiring firms driving these results. Finally, our research design randomizes the saturation of the portion of Black and long-term-unemployed applications, allowing us to identify spillovers across resumes. We find no evidence that resumes are rivalrous within an establishment, making changes in the labor supply of nonexperimental applicants an unlikely explanation of the results.

Our central contribution is providing causal evidence that minimum wage hikes reduce racial disparities in hiring by combining resume-level randomization with state-level policy variation. Our experiment offers several advantages over other approaches using existing data sets.[2] We measure and randomize the set of observable applicant characteristics, allowing us to rule out many confounds arising when examining labor-labor substitution in observational data. We can also isolate demand-side responses and examine responses both within and across firms. Moreover, we observe outcomes at a highly disaggregated level, allowing us to measure the minimum wage's impacts without relying on infrequent measurements of all workers in a geographic area or coarse demographic groupings that might be less suited to studying labor-labor substitution (Clemens and Wither, 2019; Clemens et al., 2021).

The results complement Derenoncourt and Montialoux (2021) and Bailey et al. (2021) who both find that the 1966 Fair Labor Standards Act increased wages for Black workers, but who arrive at different conclusions about the disemployment effects.[3] More generally, our finding expands our limited understanding of which policies reduce racial disparities. Previous work has considered the role of antidiscrimination policies like the 1964 Civil Rights Act (e.g. Brown, 1984), the Voting Rights Act of 1965 (Donohue and Heckman, 1991), and Ban the Box laws (Agan and Starr, 2018), along with educational desegregation, attainment, and quality (e.g. Lillard et al., 1986; Card and Krueger, 1992; Johnson, 2019). We show that minimum wage increases also reduce racial disparities.[4]

Moreover, we explicitly map the magnitudes of taste-based and statistical discrimination to policy consequences. We do this in two steps. First, we estimate functions of the taste-based and statistical discrimination parameters, without the policy variation. Our large-scale correspondence study confirms that firms hold these beliefs based on whom they call back, implying that white applicants are more likely to be marginal and hence negatively affected by the policy change. We then use the policy variation to estimate the model and illustrate

[2]See Turner and Demiralp (2001), Cengiz et al. (2019), and Wursten and Reich (2021).

[3]A large national minimum wage increase for select industries in 1967 may not have the same effects as more recent state-level changes. Derenoncourt and Montialoux (2021) argue that there was a small effect on disemployment because labor demand was inelastic and that certain important industries had monopsony power. They also argue that there was near zero racial labor-labor substitution because of segregation. The effects of the 1967 increase should also be considered with the Civil Rights Act, since both shaped labor markets and racial disparities at the time. While the impacts of the minimum wage could be similar across periods because racial discrimination and earnings differences persist to the present day, changes in institutions, labor market conditions, and segregation could lead to different effects of the policy.

[4]Similarly, David et al. (2016) show that minimum wages reduce earnings inequality.

how changes in the minimum wage shrink the racial callback gap. Overall, we illustrate how statistical discrimination impacts racial differences.[5] We then explicitly link the mechanism to the policy response and introduce a framework policymakers can use to predict the effects of the minimum wage before the policy is enacted.

Second, by measuring callbacks, we contribute to the minimum wage literature focusing on changes in job flows, which has been found to reflect minimum wage effects better than the stock of jobs (Meer and West, 2016; Gopalan et al., 2021; Jardim et al., 2022). Additionally, most previous research on the minimum wage focuses on employment, wages, and earnings overall[6] or the earnings of teenagers (e.g. Kreiner et al., 2020), or those with low levels of education (Clemens and Wither, 2019). Our experiment allows us to more directly measure the impacts on those at risk of being affected by these policies.[7]

Finally, to make these contributions, our design addresses several methodological challenges. First, by choosing a sample of states where minimum wage hikes were uncertain, we show the importance of accounting for anticipation effects. Agan and Starr (2018) published the first study to use changes in the racial callback gap to learn about the effects of policy interventions. Yet their study only considers changes after policies were announced. In our setting, market wages increase, average callbacks fall, and the racial callback gap shrinks immediately after the announcement, even before the minimum wage changes. Had we ignored anticipation effects, we would have erroneously concluded that the minimum wage does not affect callback gaps. Second, additional resume variation allows us to avoid the confound of unobserved productivity variance raised by Heckman and Siegelman (1993) and Heckman (1998). Neumark and Rich (2019) note that nearly all of the correspondence studies fail to account for differential variances in the perceived productivity of the applicants. After accounting for this possibility in the papers where it is possible to, the majority of discrimi-

[5] E.g. Altonji and Pierret (2001); List (2004); Autor and Scarborough (2008); Charles and Guryan (2008); Gneezy et al. (2012); Doleac and Stein (2013); Fryer et al. (2013); Guryan and Charles (2013); Benson and Lepage (2022); Benson et al. (2022).

[6] Neumark et al. (2007) provide a recent overview of the literature.

[7] By studying the distributional implications of the minimum wage, our work also relates to the literature on optimal minimum wages (e.g. Lee and Saez, 2012; Simon and Wilson, 2021).

nation estimates become statistically insignificant or change sign. However, we find that our results are robust to differences in perceptions of unobserved productivity variances. Finally, our two-stage randomization procedure allows us to separately identify direct resume effects and potential spillovers between applications (Phillips, 2019; Abel, 2017).

The remainder of this article proceeds as follows: In Section 2, we discuss the field experiment and policy variation in more detail. In Section 3, we introduce our parameter of interest and discuss how our research design identifies that parameter. Section 4 discusses the effects of an applicant's race and policy variation. In Section 5, we illustrate the role of taste-based and statistical discrimination. Section 6 describes our strategies for ruling out alternative interpretations and presents our findings. Section 7 concludes.

2. Setting and Experimental Design

2.1 Setting

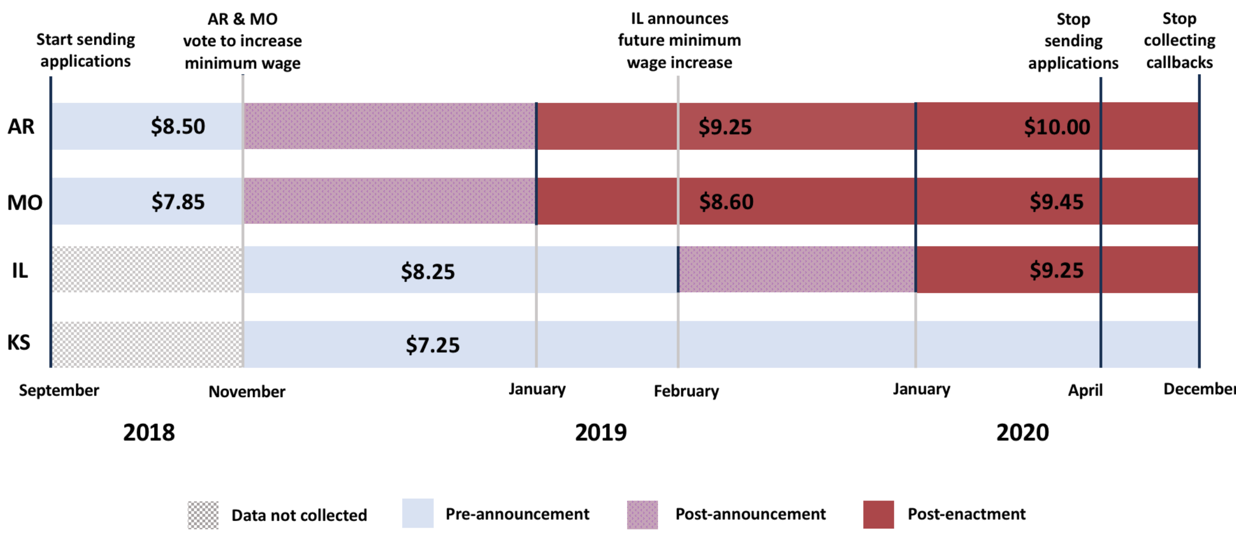
We sent fictitious applications to low-wage, entry-level ads on the largest online job-search website, which is affiliated with 47 percent of hires in the United States (Indeed.com, 2020). We applied to postings in Arkansas and Missouri starting in September 2018, and in Kansas and Illinois starting in November 2018. We sent applications through April 30, 2020. During this period, Arkansas and Missouri voted in November 2018 to increase their minimum wages in January 2019, and Illinois's legislature passed a resolution in February 2019 to increase the minimum wage beginning in January 2020.[8]

We limited our sample to jobs near Little Rock and Fayetteville in Arkansas, Kansas City (Kan.) in Kansas, East St. Louis in Illinois, and Kansas City (Mo.), Springfield, and St. Louis in Missouri. We did not anticipate the increase in Illinois, so we chose cities in Kansas and Illinois to focus on those states' borders with Missouri. We chose to use contiguous cities because contiguous counties have been shown to have more similar covariates and

[8]Polling indicated that 58–75 percent of people supported these bills. The Missouri ballot was approved with 62 percent of the vote, while the Arkansas ballot was approved with 68 percent of the vote (Ballotpedia, 2018b,a).

trends (Allegretto et al., 2011; Dube et al., 2011; Allegretto et al., 2017). Moreover, Jha et al. (2022) show the importance of using multistate commuting zones, as we do (i.e., St. Louis and Kansas City). Figure 1 outlines the timeline for each state. We categorize periods as 1) before the announcement of a minimum wage increase, 2) after the announcement but before the increase, and 3) after the new minimum wage is enacted.

Figure 1: Experimental Timeline



Notes: This figure presents the experiment timeline and minimum wage variation. Blue bars denote that the minimum wage is the same as at the beginning of the experiment, with no announced increase. Purple bars denote that the minimum wage increase has been announced. Red bars indicate an increased minimum wage.

In Table 1, we show that policy changes affected the labor market wages, using information from the job ads in our sample. First, we note that not every job post on Indeed.com provides information about the expected wages to applicants. The top panel shows that firms are slightly less likely to post any wage or an hourly wage after the minimum wage is increased. Firms may choose to post a wage when it is higher than the minimum to attract higher-quality applicants. The change in posting could reflect that after the policy change, more firms bunch at the new minimum wage and no longer benefit from posting wages.

The bottom part of Table 1 shows hourly wage estimates. Ads on Indeed have heterogeneous information about the expected wage. Some jobs post hourly wages, while others post

Table 1: Posted Wages in States Increasing the Minimum Wage

	Pre	Announced	Enacted
All vacancies			
Posts Wage (%)	39	36	36
Posts Hourly Wage (%)	37	32	28
All vacancies that posted hourly wage			
Estimated Hourly Wage	9.25	10.37	11.84
Estimated Hourly Wage $\leq$ State 2018 MW (%)	10	6	3
Estimated Hourly Wage $\leq$ State 2018 MW + 2 (%)	79	65	43
Estimated Hourly Wage $\leq$ State 2020 MW (%)	62	52	35
Estimated Hourly Wage $\leq$ State 2020 MW + 2 (%)	100	84	66

Notes: The top panel of the table shows the percentage of job ads to which we apply that post in any way and as an hourly wage, as a function of the minimum wage policy for Arkansas, Missouri, and Illinois. Ads without an hourly wage may post a daily, weekly, or monthly wage. The bottom panel shows the average hourly wage as well as the percentage of ads that are close to the 2018 or 2020 minimum wages.

daily, weekly, monthly, or annual salaries. We estimate the hourly wage using the salary and number of hours required by the employer. The average estimated hourly wages and the bottom of the wage distribution both increase. For example, at the start of our experiment, 62 percent of jobs with a posted hourly wage paid less than the 2020 minimum wage. However, after the minimum wage increases were enacted, the incidence of wages posted in this range fell by 56 percent.[9] This evidence suggests that the minimum wages were binding.[10]

[9]The number of wages at or below the 2020 minimum wage is relatively high (35 percent) because the "Enacted" column includes Arkansas and Missouri in 2019.

[10]Estimated hourly wages can be below the minimum wage for two reasons. First, some jobs, such as tip jobs, are not subject to the minimum wage, but may still increase because employers who are subject to the minimum wage compete with these firms for workers. Second, there may be measurement error because some posted wages are not hourly wages.

2.2 Design

We implemented a factorial design, randomizing the race through the name,[11] 1- or 12-month unemployment duration,[12] and whether the applicant was a high school graduate or had a GED. All three characteristics were reflected directly in the job application.[13] Fictitious applicants were 19- to 20-year-old males with one to two years of job experience.[14] Research assistants blind to the experiment constructed the applications directly on Indeed.com, consistent with the website format. The age range was chosen to make resumes as comparable as possible and so that our results relate to the majority of prior work on minimum wage effects, which focuses on younger applicants.

Each application also contained a home address, high school completion status, and a short description of previous employment. Prior fictitious work experience was constructed by mimicking sample resumes of real minimum wage job applicants on Indeed.com in the cities of our experimental sample. For each applicant, we randomly assigned job titles for prior employment to be team member, janitor, food preparation, cashier, or store associate.

All applications were sent to job postings listed on Indeed.com that fit our salary and location criteria. We applied to 17,737 vacancies across 14,488 establishments (8,376 firms) for low-wage jobs.[15] We generally sent two applications to an establishment for each job-

[11]See Appendix A for more information about the names used in the experiment. Individuals with the first names we use cover 5.9 percent of the U.S. population, and as much of 12 percent of the male population in the U.S. (Tzioumis, 2018). Appendix B discusses the possibility that implications of the names are also signaling SES.

[12]We assign one or twelve months for two months because employed applicants receive fewer callbacks than those who have been unemployed for one month (Kroft et al., 2013) and after twelve months, additional time unemployed does significantly reduce callbacks (Kroft et al., 2013). We also choose contemporary unemployment instead of historical, given that the former is associated with significant effects on callback rates but not the latter (Eriksson and Rooth, 2014).

[13]For more information about the construction of the resumes, see Appendix A.1.

[14]We only included male names for three reasons. First, many studies find mixed or no impacts of sex on employer callbacks making detecting decreases in baseline discrimination against women infeasible (Nunley et al., 2015; Zschirnt and Ruedin, 2016; Bertrand and Duflo, 2017; Baert, 2018; Kline et al., 2022). Second, further dividing the sample would have resulted in substantial reductions in power. Third, Gaddis (2017) finds that predicted race matches the intended signal for Black female names by less than Black male names. Thus, using female names would reduce the quality of the signal differentially by race. However, using only male names on the resumes limits the external validity of our experiment.

[15]Since prior work has found that employers disfavor applicants with long commutes (Phillips, 2020), we restricted search criteria to job postings by establishments located within a 10-mile radius of the cities in

posting period. If a vacancy remained up for multiple periods, we applied to it again in the subsequent period. To the best of our ability, we did not send establishments the same application type (race, unemployment duration, and high school diploma status) more than once, to minimize the risk that hiring managers would suspect that the applications were fictitious.[16]

We assigned applications to each vacancy through three main steps. First, we randomly assigned the first applicant's race, unemployment duration, and human capital, with 50-50 probability for each trait.[17] Then, we randomly stratified the second application: half on race, and half on unemployment duration. Stratification ensures that each vacancy receives zero, one, or two applications with distinctively Black names. This stratification design addresses a critique raised by Angrist (2014), Phillips (2019), and Vazquez-Bare (2022) that some correspondence studies confound discrimination with applicant pool composition.[18] All other resume characteristics were randomly assigned.[19] Establishments could call the applicant back using either the phone number or email address on the resume.

We classified callbacks as contacts requesting an interview. This method follows previous papers such as Bertrand and Mullainathan (2004) and Kroft et al. (2013). If the establishment contacted the applicant to request additional information, we did not classify that as a callback. Because there is a lag between our treatment assignment and the observation of the outcome, we had to classify callbacks with respect to the timing of the minimum wage hike. We consider callbacks as an outcome of a particular application, given the timing of the application. For example, suppose we send an application to an establishment before the minimum wage hike is announced, and the establishment calls the applicant back for an interview *after* the minimum wage hike is enacted. We classify that callback as being part of the pre-announcement period. If anything, this classification strategy biases us toward

our sample. More information about the characteristics of our sample is presented in Appendix C.

[16]Appendix Section D presents robustness checks that rule out firms learning that the resumes are ficticious.

[17]Appendix C provides evidence that observable firm characteristics are balanced across treatments.

[18]For other papers that use two stages of exogenous variation to identify spillover effects, see Duflo and Saez (2003), Crépon et al. (2015) and Holz et al. (2019).

[19]For more information on the job sampling and application process, see Appendices A.2 and A.3.

not finding an effect of the minimum wage on callbacks, because we would be misclassifying observations before the minimum wage hike as observations after the minimum wage hike.[20]

3. Minimum Wages and Discrimination

We consider a hiring model in which firms solicit and receive applications for a job that pays a wage m. The firm faces applicants i differentiated by their race $B_i \in \{0, 1\}$ (either white or Black) and other characteristics X_i that are represented in the job application (e.g., educational attainment). For each applicant, the firm must decide whether to invite an applicant to an interview, $Y_i \in \{0, 1\}$.

We aim to understand the causal effect of applicant race on callbacks, or what we call the "racial callback gap" (RCG), and how that parameter varies with the minimum wage. In our context, firms discriminate against Black applicants by calling back Black applicants less frequently than white applicants with otherwise identical characteristics, X_i:

$$RCG = \mathbb{E}\Big[\mathbb{P}[Y_i = 1 | B_i = 1, X_i] - \mathbb{P}[Y_i = 1 | B_i = 0, X_i]\Big] \tag{1}$$

where $RCG < 0$ corresponds to discrimination against Black applicants.

Generally, the two primary threats to identifying the RCG are that X_i 1) is often unobserved to the researcher despite being observed by the hiring manager, and 2) may be correlated with the applicant's race. Our experiment is designed to address both of these challenges. First, since our only correspondence with the firm is through the application, we observe all relevant information held by the firm at the time of the firm's decision. Moreover, by randomizing names, we ensure that resume characteristics are independent of the applicant's signaled race amongst the applications we send out. These two features of our experiment allow us to observe how firms respond to Black and white applicants who have otherwise identical characteristics, and thus identify the racial callback gap.

[20] We show robustness of our results to using alternative classification schemes in Appendix E.

To understand how the RCG changes with the minimum wage, we formalize the firm's decision to call back an applicant. Each applicant has a latent true quality, Q_i, which the firm does not observe during the interview and hiring process. In our context, Q_i may represent the applicant's marginal revenue product of labor.

Since the firm does not observe latent quality, Q_i, at the hiring stage, it must form an inference on i's quality with the available information. Let $\mathcal{I}$ denote the firm's information set at the time of callback choice, so that $\mathcal{I}$ includes all information (B_i, X_i) presented in i's job application, as well as the firm's beliefs about the conditional distribution of Q_i given (B_i, X_i). Let $q_i(B_i, X_i) \equiv \mathbb{E}[Q_i|\mathcal{I}]$, so that it represents the firm's expectation of i's quality at the time of callback choice. $q_i(B_i, X_i)$ in our context may represent the firm's perceived expectation of the applicant's marginal revenue product of labor, which may be different from the true expected latent quality, $\mathbb{E}[Q_i|B_i, X_i]$, because of the firm's beliefs.[21]

The perceived benefit of hiring worker i may differ from her true quality, Q_i, as a function of her race for two reasons. First, under statistical discrimination, firms may perceive Black and white candidates with otherwise identical resumes as differing in average quality, or $q_i(1, X_i) < q_i(0, X_i)$. Second, even if firms believe that the expected quality of two workers is the same, firms may engage in taste-based discrimination, in such a way that the perceived benefit of hiring worker i is lower when i is Black rather than white.

Let $v(m)$ denote the firm's expected marginal cost associated with hiring an additional worker, which is a function of the minimum wage m. The firm calls back applicant i when the expected benefit from hiring i exceeds the expected cost of doing so:

$$Y_i = \mathbb{1}\{\underbrace{q_i(B_i, X_i) + B_i d}_{\text{expected benefit}} > \underbrace{v(m)}_{\text{expected cost}}\} \tag{2}$$

where the parameter $d < 0$ captures the taste-based-driven discrimination penalty that Black

[21]Without loss of generality, we can decompose $q_i(B_i, X_i)$ into the true conditional expectation in latent quality, $E[Q_i|B_i, X_i]$, and an idiosyncratic term, i.e. $q_i(B_i, X_i) = E[Q_i|B_i, X_i] + \epsilon_i(B_i, X_i)$. The idiosyncratic component can drive heterogeneity in responses, and can reflect expectational errors made by firms, as well as both true and perceived firm-specific productivity differences (Neumark, 2012).

applicants face.[22] This decision rule is consistent with an Extended Roy Model (Heckman and Vytlacil, 2007) wherein the firm chooses whether to call back i by minimizing the total expected cost, given applicant information (B_i, X_i).

We assume that the conditional distribution of $q_i(B_i, X_i)$ has a strictly decreasing pdf around the cutoff. Taking together the RCG definition (Eq. 1) and the firm decision rule (Eq. 2), the change in the RCG from a minimum wage increase is therefore given by:

$$\frac{\partial RCG}{\partial m} = v'(m) \times \big[f(v(m)|B_i = 0, X_i) - f(v(m) - d|B_i = 1, X_i)\big] \tag{3}$$

where f is the pdf of $q_i(B_i, X_i)$. From McCall (1970), we expect that firms reduce the number of (relatively low-quality) applicants they call back when they face a higher labor cost from the minimum wage hike. In other words, we assume that $v(m)$ is an increasing function of the minimum wage policy: $v'(m) > 0$.

The second term on the right-hand side compares the share of white marginal applicants to the share of Black marginal applicants. Without knowing more about the productivity distributions and the magnitude of the taste-based discrimination, we cannot sign the second term. Therefore, the minimum wage could increase or decrease the RCG, depending on the extent and type of discrimination. The primary threat to identifying $\partial RCG/\partial m$ is changes in other aspects of the firm's decision that are contemporaneous with the firm's decision. Hence, the identification of this parameter relies on an assumption that the RCG would have remained constant but for the minimum wage hikes. We overcome these challenges by using comparable states as a control group and unannounced changes in the minimum wage.

As a special case of the model above, fix $X_i = x$, suppose that $q_i(B_i, x)$ is normally distributed so that $q_i(0, x) \sim N(\mu_0, \sigma_0^2)$ and $q_i(1, x) \sim N(\mu_1, \sigma_1^2)$. Racial differences in firms' perceptions of mean productivity, and the variance in productivity, reflect statistical discrimination. In this special case, we allow the means and variances of productivity to vary

[22] Alternatively, we could interpret d as a race-dependent variation in perceived cost of hiring i. This would not affect our qualitative conclusions.

by race. Under these assumptions, the change in the RCG from an increase in the minimum wage is

$$\frac{\partial RCG}{\partial m} = v'(m) \times \left[\phi((v(m) - \mu_0)/\sigma_0)\sigma_0^{-1} - \phi((v(m) + d - \mu_1)/\sigma_1)\sigma_1^{-1}\right]$$

where ϕ is the standard normal pdf. Both statistical and taste-based discrimination determine how the RCG changes with the minimum wage. For example, minimum wages can attenuate the RCG if σ_1 is sufficiently larger than σ_0—that is, if firms believe that the variance of the quality distribution for Black applicants is larger than for white applicants.[23] Appendix G discusses the implications of alternative assumptions on $q_i(B_i, \cdot)$. For example, if $q_i(B_i, \cdot)$ is uniformly distributed, the $\partial RCG/\partial m$ only depends on the perceived variance of applicant productivity, and not at all on statistical discrimination in means, or taste-based discrimination.

4. Results

4.1 Which Workers Receive Callbacks?

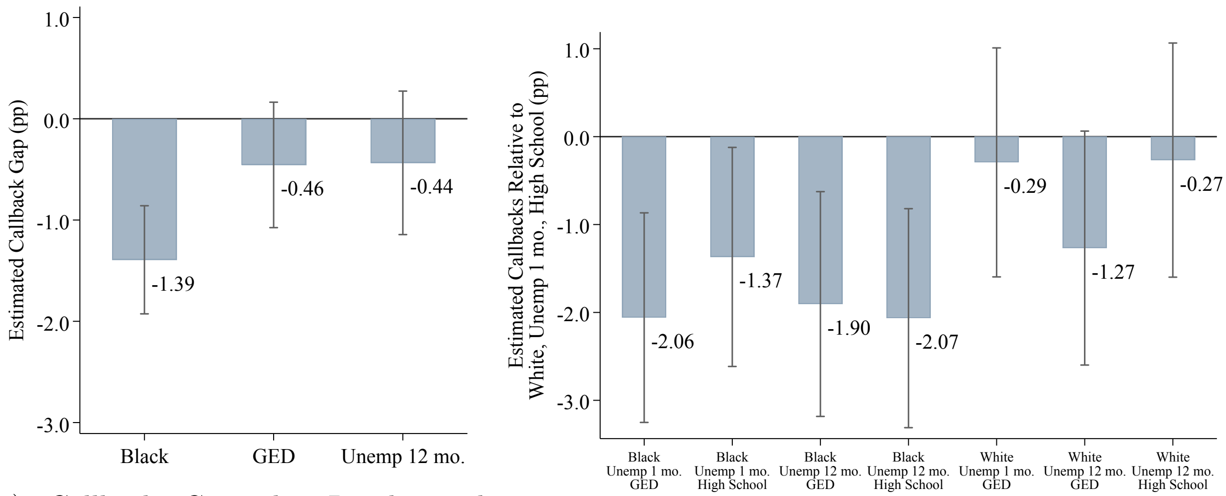
We begin by estimating the causal effect of perceived race, human capital, and duration unemployed on the likelihood of receiving a callback over our full sample:

$$Y_{ict} = \alpha_1 \text{Black}_i + \alpha_2 \text{GED}_i + \alpha_3 \text{Unemp12}_i + X_i'\gamma + \eta_c + \epsilon_{ict} \tag{4}$$

where Y_{ict} is an indicator representing whether application i received a callback from a job posting by firm j in city c at time t. $Black$, GED, and $Unemp12$ are indicators for whether an applicant was randomized to have a distinctively Black name, hold a GED rather than a high school degree, and be 12 months unemployed rather than 1 month, respectively. X is a vector of other randomized applicant characteristics. In all specifications, we include

[23] Under normality, this result requires that the baseline callback rate is less than 50 percent for both groups, as is the case in our and many other audit and correspondence studies.

Figure 2: Callback Gaps in the Full Sample



(a) Callback Gaps by Randomized Characteristic

(b) Callbacks by Applicant Type

Notes: Panel (a) presents estimates and 95 percent confidence intervals for the callback gaps based on the randomized characteristics: (1) Black-white, (2) GED–high school graduate, and (3) 12 months–1 month unemployed. Panel (b) similarly presents estimates for the likelihood an applicant receives a callback relative to White applicants with a high school diploma who have been unemployed for one month. Both specifications include city fixed effects, control for applicant age, and include the full sample of 34,986 observations. Standard errors are clustered by establishment.

city fixed effects, η_c. The α coefficients estimate the causal effect of the characteristic on callbacks.

Overall, the callback rate for white applicants is 11 percent. Similar to previous correspondence studies, we find that applicants with distinctively Black names are about 1.4 percentage points (12 percent) less likely to receive a callback than applicants with distinctively white names (Figure 2a). The callback gaps based on prior education and duration unemployed are about one-third of the size of the RCG and not statistically significant.

We present the effect of each of our types of applicants in Figure 2b, relative to white applicants with a high school diploma who have been unemployed for the previous one month. We find little heterogeneity by education and duration unemployed in the callback rates for Black applicants. This is consistent with "attention discrimination," where hiring managers pay less attention to applications from minorities and so do not see their qualifications

(Bartoš et al., 2016). White applicants with a high school degree who are one month unemployed are 1.3 percentage points more likely to receive a callback than white applicants with a GED who are 12 months unemployed ($p = 0.062$).

4.2 The Association between Minimum Wages and the Overall Callbacks

Next, we present descriptive evidence on how the overall callback rate changes with the minimum wage. For the sake of simplicity, we focus on two time periods, 1) before the announcements of the minimum wage changes and 2) after the announcements. We estimate the following equation:

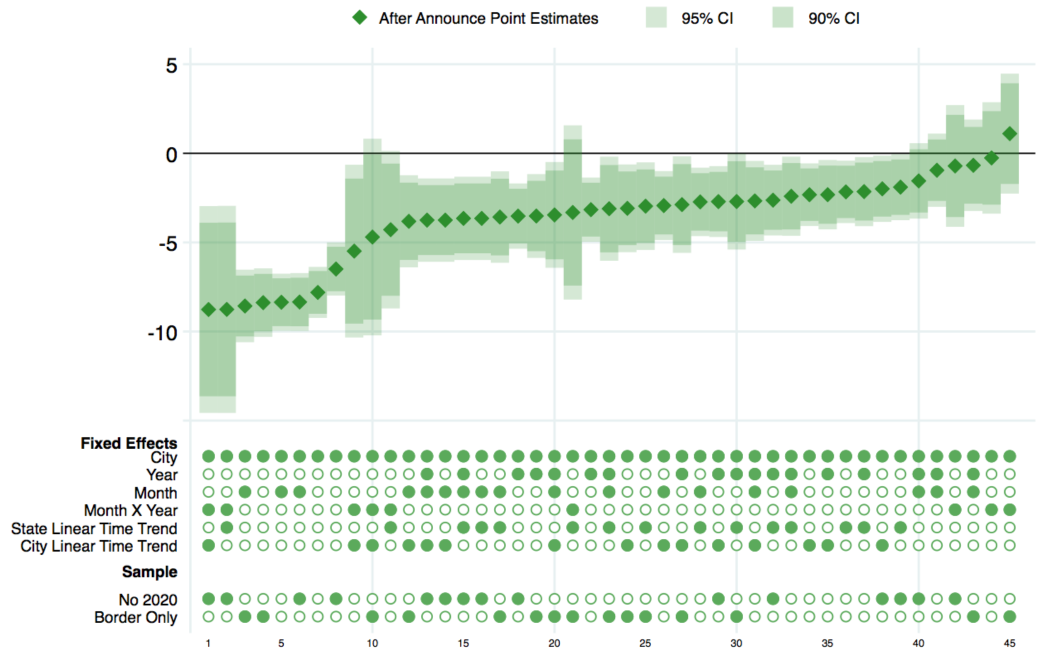
$$Y_{ict} = \psi \text{After Announced}_{ct} + X_i' \gamma + \eta_{ct} + \epsilon_{ict}. \tag{5}$$

Here, ψ captures the change in callbacks after an announcement of a minimum wage increase. Both the variation in the timing of the announcements across states and whether the states raise their minimum wages at all allow us to estimate ψ. We include city fixed effects and city linear time trends, defined by the month of the date of application, η_{ct}. We additionally consider more flexible controls, as well as sample restrictions, and show how our results change across specification in Figure 3.

The likelihood of a callback decreases after the minimum wage hikes by about 2.3 percentage points ($p = 0.010$), or 14.8 percent, based on our preferred specification. Since the average minimum wage increase in our sample is about 13.5 percent,[24] we estimate a callback elasticity of –1.09. This elasticity likely reflects a combination of the firm's decisions to only hire higher-quality applicants at the higher price, hiring fewer workers for each posted ad, and an increase in the likelihood of an applicant accepting an offer. This effect is large in magnitude compared to the Congressional Budget Office's median employment elasticity of –0.25, but is still within the CBO's overall reported range of 0.4 to –1.7 (Alsalam, 2019).

[24]We estimate the average minimum wage percent increase in our sample by regressing the log of the minimum wage on indicators for announced and enacted, controlling for city fixed effects. The coefficient on enacted gives the relevant percent increase for our sample.

Figure 3: The Minimum Wage and Callbacks



Notes: This figure presents a specification curve for the change in the likelihood of receiving a callback from a minimum wage increase. The filled-in circles indicate which controls are included. The "Border-Only" sample includes only observations from Kansas City (Kan. and Mo.) and St. Louis (Ill. and Mo.). In all specifications, standard errors are clustered by establishment.

The callback elasticity is also considerably larger than previous work's estimates of separation elasticities of around –0.3 (e.g. Dube et al., 2016). Together with the previous work on employment flows, our results suggest that the search costs of finding employment increase after the minimum wage hikes.

Our empirical strategy to estimate the callback elasticity relies on geographic-specific linear time trends. Specifications that also include month-of-year or year fixed effects, that restrict the sample, or that use state instead of city linear time trends, mostly have similar estimates. Among those 24 specifications, our point estimates for ψ range from –1.9 to –3.8. We find much more extreme estimates when we either include both geographic linear time trends and month-by-year fixed effects, or exclude the geographic time trends.

4.3 The Causal Effect of Minimum Wage Increases on the Racial Callback Gap

We now examine how the minimum wage affected the RCG, by estimating

$$\begin{aligned} Y_{ict} = &\beta_1 \text{Announced}_{ct} + \delta_1 \text{Black}_i \times \text{Announced}_{ct} + \beta_2 \text{Enacted}_{ct} \\ &+ \delta_2 \text{Black}_i \times \text{Enacted}_{ct} + \alpha \text{Black}_i + X_i'\gamma + \eta_c + \epsilon_{ict} \end{aligned} \tag{6}$$

where Y_{ict} is an indicator representing whether application i received a callback from a job posting in city c at time t, *Black* is an indicator representing whether the applicant is Black, *Announced* is an indicator variable representing whether the minimum wage increase has been announced but not yet enacted, *Enacted* is an indicator representing whether the minimum wage has increased,[25] and X is a vector of randomized applicant characteristics, including the applicant's length of unemployment, human capital, and age. *Announced* and *Enacted* are mutually exclusive. We also include city fixed effects, η_c.

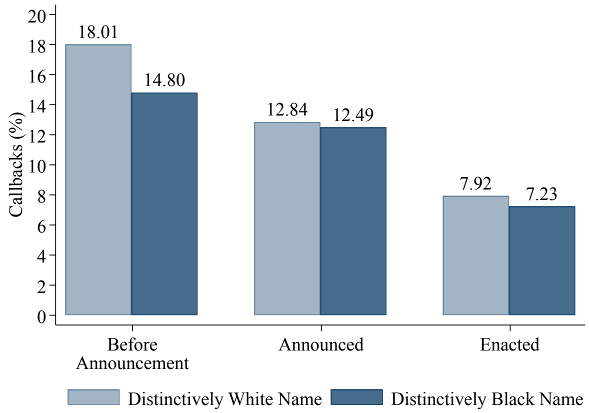
Our research design identifies the effect of minimum wage increases on the RCG under an assumption of parallel trends in the racial callback gap using a difference-in-differences design. The first difference comes from randomization, which allows us to measure the callback gap between Black and white applicants, holding all other characteristics fixed that might also affect labor market disparities (Card and Krueger, 1992; Smith and Welch, 1989). The second difference comes from measuring callbacks before and after the policy change. In Appendix F, we present evidence that there were not differential trends prior to the announcement.[26]

Our results in Figure 4 show that before the announcement of the minimum wage increase,

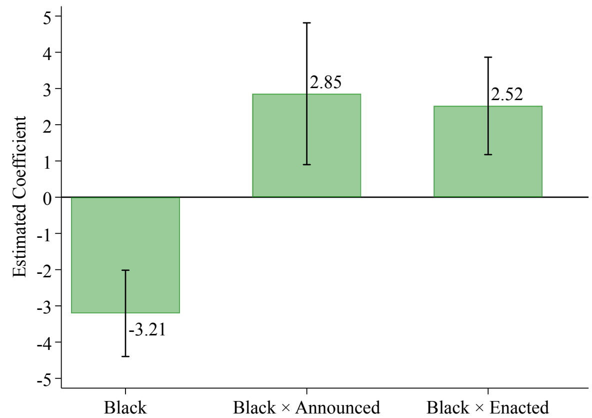
[25] Appendix E shows that assigning treatment based on the earliest date of callback for vacancies where at least one applicant receives a callback does not affect our results.

[26] In Appendix I we also consider specifications with firm fixed effects to compare the behavior of different establishments that belong to the same parent company, but where one experiences a minimum wage increase and the other does not because they are located in different states. Note that since an observation is at the applicant × job-posting × establishment level, this strategy, whether or not we include firm or establishment fixed effects, does not rely on two-way fixed effects.

Figure 4: The Minimum Wage and the Racial Callback Gap



(a) Average Callback Rates by Race and Time Period



(b) Racial Differences

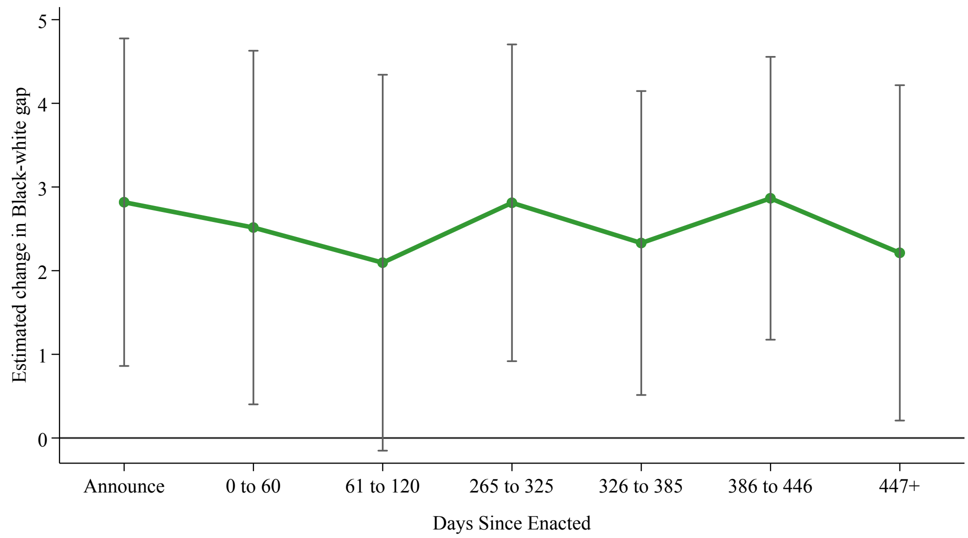
Notes: Panel (a) presents estimates of the average callback rates by race and time period, based on the estimates from Equation 6. Panel (b) displays coefficient estimates from Equation 6 and 95 percent confidence intervals. The coefficient estimate on *Black* represents the baseline difference in callback rates between applications with distinctively Black names and those with distinctively white names. The coefficient estimate on *Black* $\times$ *Announced* can be interpreted as the change in the racial callback gap after the minimum wage hike is announced. Similarly, the coefficient estimate on *Black* $\times$ *Enacted* can be interpreted as the change in the racial callback gap after the minimum wage hike is enacted. Figures include the full sample of 34,986 observations. In Appendix F, we present evidence that there were not differential trends prior to the announcement.

applicants with distinctively Black names were about 3 percentage points less likely to receive a callback. When the minimum wage increase is announced, all applicants are less likely to receive a callback, but the callback rate decreases by less for Black applicants; after the policy is enacted, the racial callback gap shrinks to about 0.7 percentage points but is still different from 0 ($p = 0.029$). We therefore reject the claim from Friedman (1966) that minimum wage increases exacerbate racial differences in hiring. We also show in Appendix H that minimum wage increases do not affect the callback gaps by education or duration unemployed.

Our results presented in Figure 4 suggest that employers respond immediately to the news that the minimum wage will increase soon. This may be because there is a relatively short time between the announcement and the enactment, so all applicants considered during the announced period would not begin work until after the minimum wage increase is in effect.

To further understand the dynamics, we estimate a specification in which we divide the enacted period into intervals of the same size as the pre-period, and present the estimates in Figure 5. The effect remains constant for more than a year after the enactment. Based on these results, we pool the announced and enacted periods for the remainder of the analysis.

Figure 5: The Minimum Wage and the Racial Callback Gap Dynamics

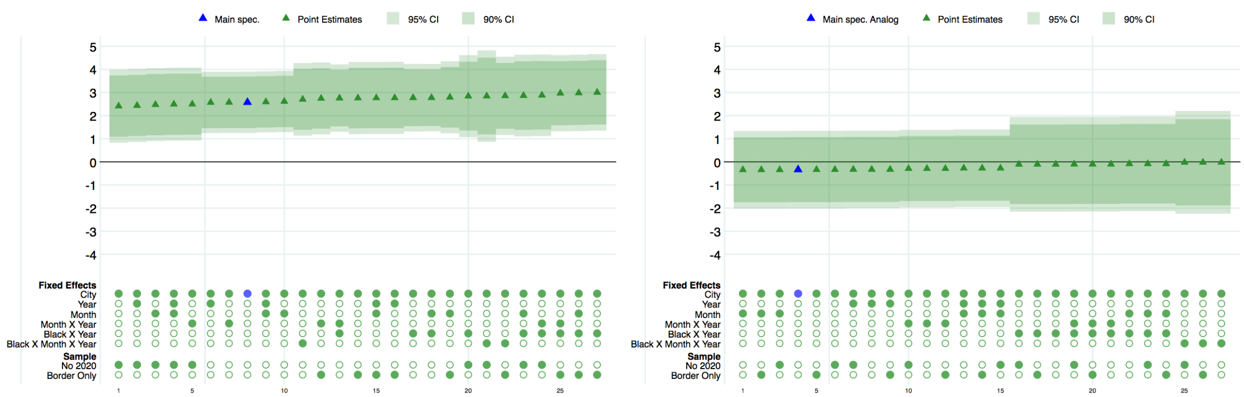


Notes: The figure presents estimates and confidence intervals for the change in the racial callback gap by number of days since the minimum wage was enacted. We group days into 60-day bins to reflect the length of the announced period in Arkansas and Missouri. No applicants were sent applications between days 120 and 265. The specification includes city fixed effects and controls for GED, duration unemployed, and age. Standard errors are clustered by establishment. The specification includes the full sample of 34,986 observations.

4.4 Robustness and the Importance of Announcement Effects

Our research design allows us to consider several alternative specifications that vary the sample or include additional controls. We present the results from these specifications in Figure 6a, which estimates variations of Equation 6 and plots the change in the gap after the minimum wage increase in green, with our preferred specification from Figure 4 highlighted in blue. These specifications vary temporal fixed effects to account for potential concerns with differential seasonality by city. We also include variation of these fixed effects interacted with the applicant's race to allow for differences in seasonality by race. And, we also present specifications where we exclude 2020 because of the COVID-19 pandemic. Finally, we include estimates from the contiguous county design recommended by Allegretto et al. (2011) Dube et al. (2011), Allegretto et al. (2017), and Jha et al. (2022). The estimates across all specifications are nearly identical, and all point to the same qualitative conclusion that the minimum wage increases reduced the racial callback gap.[27]

Figure 6: Specification Curves with and without Pre-announcement Period



(a) Including Pre-announce Period

(b) Excluding Pre-announce Period

Notes: These figures present specification curves for the change in the racial callback gap from a minimum wage increase. Panel (a) displays estimates of the main specification, pooling the after-announcement and after-enactment effects. Panel (b) displays the effect of enactment, using the after-announcement period as the pre-period. The preferred specification and its analog are highlighted in blue. The filled-in circles indicate which controls are included. The "Border Only" sample includes only observations from Kansas City (Kan. and Mo.) and St. Louis (Ill. and Mo.). In all specifications, standard errors are clustered by establishment.

In all specifications, we cluster the standard errors by establishment. Alternative levels of clustering have a small effect on the precision of our estimate for the coefficient on *Black* × *After Announce*. We consider three alternatives. First, Agan and Starr (2018) suggest clustering by firm, because establishments within a firm may be susceptible to serially correlated shocks. When doing this, the standard errors and p-values minimally change. While randomization is at the job vacancy level, the policy variation is by state, and so Neumark et al. (2019) suggest clustering at that level. Since the sample only has four states, we estimate p-values using the wild bootstrap (Cameron et al., 2008; Canay et al., 2021). In this case, the

[27]See Appendix I for analysis with other specifications that include firm or establishment fixed effects.

p-value for the coefficient on *Black* $\times$ *After Announce* is still less than 0.001, although the confidence interval is about 1.5 times as large as when we cluster by establishment. Finally, for completeness, when we cluster by city using the wild bootstrap, the p-value is 0.016.

Figure 6b presents these same specifications, using the postannouncement period as the baseline. The estimates represent what the results from the study would have been, had we begun measuring the policy effect after the policy change was certain, as in Agan and Starr (2018). Similar to Figure 4a, the figure shows that the conclusions from the study hinge on the ability to measure announcement effects. The coefficient estimates in this figure are never positive or statistically distinguishable from zero. Had we not measured the change in callbacks caused by the policy announcement, we would have falsely concluded that the minimum wage hikes had no effect, or, if anything, a small negative effect on callbacks.

5. Illustrating the Role of Taste-Based and Statistical Discrimination

The results in Section 4.3 show that increases in the minimum wage reduce the racial callback gap. In line with our model, this may imply that firms believe that the variance of applicant quality is larger for Black applicants. However, the total impact of the minimum wage change depends on both taste-based and statistical discrimination. We follow Neumark (2012) and Neumark et al. (2019) to partially separate these two channels and illustrate the role of taste-based and statistical discrimination for our findings.

As in Section 3, suppose $q_i(0,x) \sim N(\mu_0, \sigma_0^2)$ and $q_i(1,x) \sim N(\mu_1, \sigma_1^2)$, and that firms require Black applicants to be of higher quality to receive a callback due to taste-based discrimination, d. Then we can partially recover these parameters by estimating a heteroskedastic probit:

$$Y_{ict} = \mathbb{1}\big[\underbrace{\alpha}_{(\mu_1-\mu_0)-d}\text{Black}_i + X_i'\eta + (\text{Black}_i\sigma_1 + (1-\text{Black}_i)\sigma_0)\epsilon_{ict} > 0\big] \tag{7}$$

where $\epsilon_{ict} \sim N(0,1)$ and X_i is a vector of observable randomized characteristics and city fixed effects. α captures both taste-based discrimination and statistical discrimination from differences in the mean of the perceived quality distribution. If observable randomized characteristics and city fixed effects affect a firm's beliefs about the applicant's productivity similarly by race, η is the same for Black and white applicants, and we can identify σ_1^2/σ_0^2. While the coefficient is assumed to be the same, that does not imply that the effect on callbacks is identical. Both η and σ determine the likelihood of receiving a callback.

We first pool the data over the whole sample and estimate $(\mu_1 - \mu_0) - d$ and σ_1/σ_2 (column 1 of Table 2). To estimate the relative variance, σ_1^2/σ_0^2, shown in Table 2, we use 10 characteristics that affect the likelihood of receiving a callback. From our randomized characteristics, we include indicators for high school degree and unemployed 1 month, high school degree and unemployed 12 months, and GED and unemployed 1 month, as well as applicant age.[28] We also include city fixed effects. Our approach allows us to first estimate the discrimination parameters using only the randomized characteristics that affect the likelihood that an applicant receives a callback. We then additionally include the policy variation from the minimum wage hike as an eleventh variable that also affects the likelihood of receiving a callback.

Using multiple characteristics allows us to test the assumption that η does not vary by race. When the assumption holds, the marginal effects of a characteristic on callbacks only differ by race because of the difference in the variance of the unobservables. As Neumark (2012) notes, the ratios of the estimated probit coefficients for Black and white applicants, for each characteristic, should therefore be the same. We estimate a probit including the characteristics and their interactions with an indicator for whether the resume has a distinctively Black name, and we fail to reject that the ratios are the same (p-value $= 0.73$).[29]

We find that $(\mu_1 - \mu_0) - d < 0$, meaning that firms discriminate against Black applicants

[28] Alternatively, only including indicators for GED and being unemployed 12 months instead of also including the interaction, and/or including state fixed effects instead of city fixed effects, has a very minimal effect on our estimates of the discrimination parameters.

[29] The p-value for the specification in Equation 7 that also uses the policy variation, presented in column (2) of Table 2, is 0.88.

through some combination of taste-based and beliefs about mean productivity. We also find that $\ln\left(\sigma_1^2/\sigma_0^2\right) > 0$, implying that firms believe the variance of the quality distribution is larger for Black applicants. These estimates are similar to those in Neumark (2012), who finds suggestive evidence that the perceived productivity variance of Black applicants is larger using data from Bertrand and Mullainathan (2004). However, he does not have the power to reject that the two variances are the same.[30]

Table 2: Unobserved Productivity Differences and the Minimum Wage

	Callback	
	(1)	(2)
Discrimination Parameters		
$(\mu_1 - \mu_0) - d$	-0.40**	-0.38**
	(0.16)	(0.12)
$-v'(m)$		-0.45***
		(0.04)
$\ln\left(\sigma_1^2/\sigma_0^2\right)$	0.22*	0.21**
	(0.10)	(0.08)
Characteristics (X)		
HS × Unemp 1 mo.	0.06^{+}	0.06*
	(0.03)	(0.03)
HS × Unemp 12 mo.	0.03	0.03
	(0.03)	(0.03)
GED × Unemp 1 mo.	0.03	0.03
	(0.03)	(0.03)
Age	0.03	0.04
	(0.03)	(0.03)
N	34,990	34,990

Notes: Standard errors are clustered by establishment. Specifications include city fixed effects.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $^{+}$ $p < 0.10$

We then measure the extent to which the callback threshold changes after the minimum

[30]The results and framework are also consistent with interviews with hiring managers who are more likely to perceive racial differences in applicants than in their own employees (Pager and Karafin, 2009).

wage increases, using the following equation,

$$Y_{ict} = \mathbb{1}\big[\underbrace{\alpha}_{(\mu_1-\mu_0)-d}\text{Black}_i + \underbrace{\beta_1}_{-v'(m)}\text{After Announced}_{ct} + X_i'\eta + (\text{Black}_i\sigma_1 + (1-\text{Black}_i)\sigma_0)\epsilon_{ict} > 0\big]. \quad (8)$$

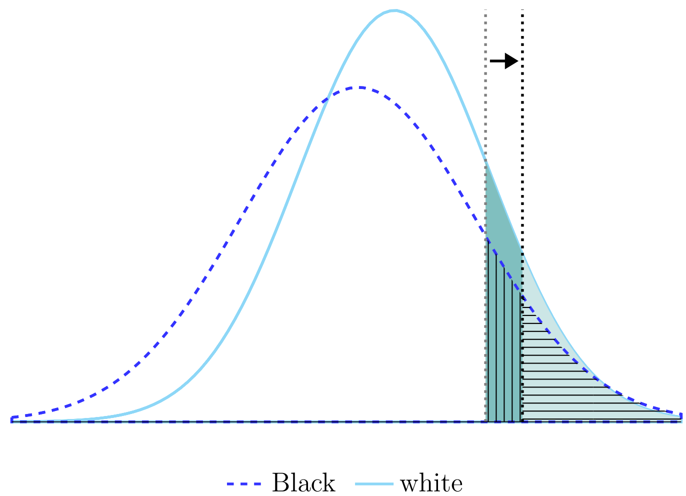
In column (2) of Table 2, we provide estimates from Equation 8. Consistent with the intuition from McCall (1970), we find that the threshold moves to the right ($v'(m) > 0$), and so fewer applicants receive an interview request after the policy change. Our estimates of the taste-based and statistical discrimination parameters remain nearly the same, which suggests that other researchers and policymakers can use this approach, without the policy variation, to study the impacts of potential minimum wage hikes on racial discrimination in hiring.

We now illustrate in Figure 7, using our estimates from column (2) of Table 2, that differences in the perceived variance of applicant quality for Black and white applicants, or statistical discrimination, can help explain why the racial callback gap shrinks with minimum wage increases. The perceived quality distribution of white applicants is given by the solid curve, and the distribution for Black applicants is shown by the dashed curve. The dotted vertical lines represent the callback minimum-quality threshold.[31]

Based on the discriminatory penalty and perceived variance of the productivity distributions between Black and white applicants, a larger number of white applicants receive callbacks before the minimum wage increase; there is relatively more mass under the solid curve to the right of the vertical dotted line than for the dashed curve. After the new minimum wage is announced, the callback threshold shifts to the right, and establishments only call back higher-quality applicants. Initially marginal applicants, those between the two dotted lines, no longer receive a callback. We find that, in our setting, because white

[31] We cannot separately identify differences in perceived mean quality and the callback threshold by race (f). For exposition, following the model, we assume that Black and white applicants face the same threshold. We interpret d and animus discrimination as directly affecting the value of hiring Black applicants, like μ, rather than a difference in the hiring standard. This assumption does not affect the interpretation of the results since our discussion is based on the relative mass of applicants around the threshold. Alternative assumptions would shift the distribution and threshold by the same constant.

Figure 7: Minimum Wage Increases and Callbacks by Applicant Productivity



Notes: This figure plots the share of applicants who receive a callback by race, based on our estimates from Table 2. The solid normal line shows the distribution of the perceived benefit of hiring a white applicant, while the dashed line is for Black applicants. The vertical dotted lines represent the callback thresholds before and after the minimum wage increase. An applicant receives a callback if he is to the right of the threshold, corresponding to the shaded region.

applicants have a smaller perceived variance and are favored in hiring, a larger share of the initially marginal applicants are white, and therefore they experience a larger decrease in callbacks.

Our framework estimates the relative perceived variances of the productivity distributions by race, which are not necessarily the true relative variances. By randomization in the correspondence study, the productivity distributions should be identical. However, firms may have inaccurate beliefs (Bohren et al., 2019) or limited previous experience in hiring different types of applicants (Lepage, 2021).

These results highlight that minimum wage increases are not guaranteed to reduce the racial callback gap. Instead, the effects of the minimum wage critically depend on the nature and extent of discrimination in the low-wage labor market. However, the unambiguous sign of $v'(m)$ allows researchers or policymakers to determine how a minimum wage will affect the racial callback gap even *without* variation in the minimum wage. In any low-skill labor market, one can conduct a simplified correspondence study that varies the race of the applicant and the perceived quality of the applicant to identify $(\mu_1 - \mu_0) - d$ and $\ln\left(\sigma_1^2/\sigma_0^2\right)$ using Equation 7. Then, given these parameters, the model predicts the sign of $\partial RCG/\partial m$.

6. Assessing Alternative Explanations for Why the Racial Callback Gap Shrinks

Our experiment allows us to rule out competing mechanisms driving the reduction in the racial callback gap. First, increases in labor costs may affect the extent to which hiring managers rely on race as a signal of worker quality. Second, increases in the cost of employing applicants may affect the composition of establishments posting job ads. Finally, we investigate the possibility that nonexperimental job applications respond to the increase in the minimum wage, changing the composition of applicants and, thereby, the probability of a callback. We find only limited evidence of any of these channels, suggesting that statistical discrimination and taste-based discrimination mainly drive our findings.

6.1 Observable Signals of Worker Productivity

Bartoš et al. (2016) show that managers spend less time reviewing a Black applicant's resume. Minimum wage hikes increase the costs of hiring a low-quality applicant, perhaps leading managers to rely more on additional signals of quality in workers, such as their human capital or previous relevant experience. If increases in the minimum wage induce managers to spend more time reviewing each application, then employers will be more likely to know that Black applicants are of high quality when the minimum wage is high. Thus, attention discrimination in this market implies that the returns to quality increase for Black applicants relative to white applicants following the minimum wage increase.

Using education and previous relevant experience as quality measures, we test whether attention discrimination drives our results by estimating the extent to which the racial callback gap changes with the minimum wage, based on these characteristics:

$$Y_{ict} = \sum_r \alpha_r B_i \mathbb{1}_{ir} + \sum_r \beta_r \mathbb{1}_{ir} \text{After Announced}_{ct} \tag{9}$$
$$+ \sum_r \delta_r B_i \mathbb{1}_{ir} \text{After Announced}_{ct} + X_i' \gamma + \eta_c + \epsilon_{ict}$$

where $\mathbb{1}_{ir}$ indicates whether applicant i has randomized characteristic r. We first estimate the effect of the minimum wage increase on the racial callback gap, allowing for heterogeneity by each combination of an applicant's human capital and duration of unemployment. Figure 8a presents the estimates of δ_r. While we find that the callback gap shrinks the most among applicants with a high school degree who are 12 months unemployed, we cannot reject the notion that any pair of coefficients is different. The results are consistent with a closing of the racial callback gap for all types, since applicants with high school degrees and a 12-month unemployment duration have the largest baseline racial gap. Overall, the results do not suggest that the highest-quality Black applicants see the largest relative increases in callbacks, as attention discrimination would suggest.

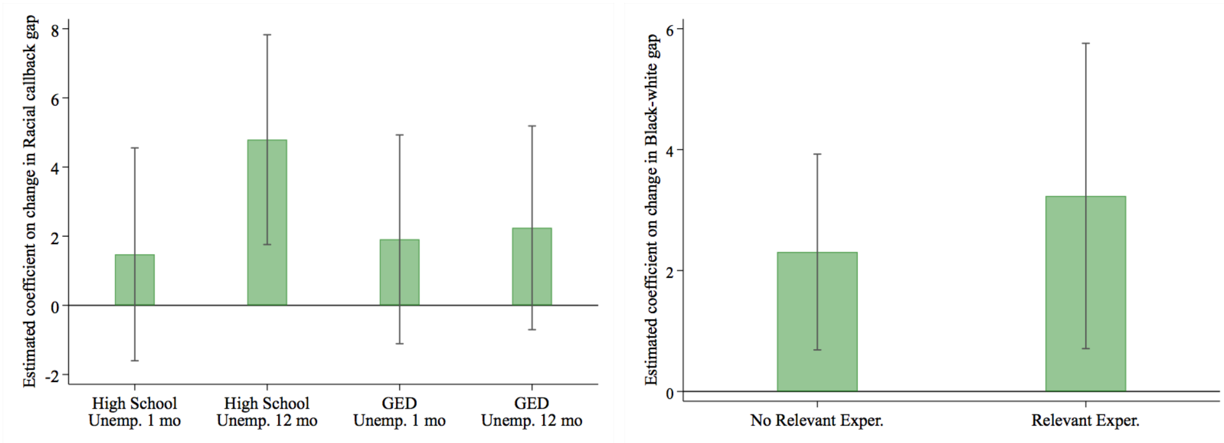
As a second test, we estimate the effect of the minimum wage increases on the racial callback gap for applicants with and without relevant experience.[32] For each applicant, we randomized the job title associated with his or her previous job. We similarly categorize job ads based on the job titles and descriptions for each position. We estimate Equation 9 with r denoting whether or not an applicant had previous relevant experience, and we present the results in Figure 8b. While the point estimates suggest that the racial callback gap shrinks more for applicants with previous experience, we cannot reject that they are different.

6.2 Changes in the Composition of Job Vacancies

Kline et al. (2022) show substantial heterogeneity in discrimination against minority applicants across firms, implying that in the composition of firms, as well as perhaps in es-

[32]These results should be interpreted with caution as the relevant work experience depends on the job ads posted by employers after the minimum wage increase, which may be a function of the minimum wage.

Figure 8: Change in Racial Callback Gap by Previous Relevant Experience



(a) Other Characteristics (b) Applicant Quality

Notes: Panel (a) presents estimates and confidence intervals for the change in the racial callback gap, allowing for heterogeneity by combinations of education and months unemployed. Panel (b) presents estimates and confidence intervals for the change in the racial callback gap, allowing for heterogeneity by whether the applicant was randomized to have previous relevant experience. Both specifications control for city fixed effects and other characteristics. Standard errors are clustered by establishment.

tablishments, hiring will affect the estimates of discrimination. This composition is potentially a function of the minimum wage. Luca and Luca (2019) show that minimum wage increases lead low-productivity firms to close, and firms that discriminate are likely less efficient (Becker, 1957; Pager, 2016). When facing higher labor costs, establishments may also respond by demanding more skilled applicants or altering the types of tasks performed on the job (Clemens et al., 2021). In this section, we show that our results are unlikely to be driven by changes in firm composition.

To show that changes in firm composition do not drive our results, we first add firm fixed effects to Equation 6 and compare the behavior of different establishments belonging to the same parent company. Some establishments experience minimum wage increases, while others do not because they are located in different states.[33] This specification limits the role of compositional changes by focusing on the within-firm variation. In this specification,

[33]Since an observation is at the applicant × job-posting × establishment level, this strategy does not rely on two-way fixed effects.

shown in column (2) of Table 3, the estimates on *Black* and *Black* × *After Announce* are slightly smaller in magnitude, but not different from column (1).

Next, we estimate a version of Equation 6 that includes establishment fixed effects to limit the role of establishment compositional changes further. The coefficients are smaller in magnitude, but the results still indicate that minimum wage increases shrink the racial callback gap by 56 percent (Table 3, column [3]). Overall, this pattern of results suggests that establishment composition likely plays a role, but it does not fully explain why the minimum wage shrinks the racial callback gap. Appendix I presents a specification curve showing that these results are robust to alternative specifications, as in Figure I2.[34]

Table 3: The Minimum Wage and the Racial Callback Gap

	(1)	(2) Callback	(3)
Black	-3.21***	-2.94***	-2.33***
	(0.61)	(0.65)	(0.58)
After Announce	-9.09***	-6.61***	-8.50***
	(0.82)	(1.00)	(1.31)
Black × After Announce	2.58***	2.11**	1.30*
	(0.68)	(0.73)	(0.65)
Firm Fixed Effects	No	Yes	No
Establishment Fixed Effects	No	No	Yes
Observations	34,990	34,990	33,398
R-Squared	0.02	0.36	0.62

Notes: All specifications include city fixed effects and control for applicant duration unemployed, human capital, and age. Two observations are excluded from column (3) with establishment fixed effects, because they are the only applications to those establishments. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05, $^{+}$ p-value < 0.1

As a complementary exercise to measure the role of establishment compositional changes, we divide our establishments into groups based on when the establishment was hiring and when we applied.[35] If establishment composition solely contributes to the shrinking of the

[34]The appendix also presents the number of postings by unique firms and establishments over time.

[35]Whether and when an establishment chooses to hire is endogenous, but splitting our sample in this way helps illustrate both changes in composition and policy responses.

racial gap, then we would expect no differences in the callback behavior of establishments hiring both before and after the minimum wage. We would expect there only to be a difference between establishments that *only* hired either before or after the minimum wage hike. Instead, if establishments are only changing their behavior because of the policy, then we should see similar patterns between establishments hiring in both periods and those hiring only before or after the hikes. Appendix J provides additional evidence that establishment composition does not drive out the effect of minimum wages on the RCG.

6.3 Robustness to Changes in Nonexperimental Applications

Next, we consider relaxing the view that the firms' beliefs about the distributions of worker productivity are primitive and the minimum wage only changes the hiring threshold. One concern is that the minimum wage may change the search efforts of nonexperimental applicants, increasing the competition for callbacks. However, several conditions must hold for these changes to confound our results. The minimum wage must change the search effort of other applicants, applications to the same vacancy must be rivals in callbacks, and the effects of additional resumes must affect callbacks differentially by race. If any of these conditions are economically insignificant, then changes in labor supply will not affect our interpretation of the results. We address the validity of each of these factors in turn.

Labor market search-and-matching models predict that increases in the minimum wage lead to an increase in the supply of applicants and search intensity, which in turn leads to better job matches (Flinn, 2006; Ahn et al., 2011). However, empirically, Adams et al. (2022) find evidence of increases in search efforts of those previously looking for minimum wage jobs following minimum wage hikes, and no changes to the composition of those searching for minimum wage jobs. Moreover, the intensive margin changes quickly dissipated one month after the hike. Thus, if changes in search effort were driving our results, we would expect to see the callback gap rebound shortly after the minimum wage hike. However, as we showed in Figure 5, the reduction in the RCG persists throughout the sample period, suggesting that changes in search effort are not a substantial concern in our experiment.

Next, we investigate whether applications for the same vacancy are rivals in callbacks. In our model, firms call back all workers above the threshold. However, an alternative model is that the firm only calls back a limited number of the best candidates from its pool of applicants. While we cannot measure nonexperimental resumes, our randomization strategy allows us to learn whether experimental applications for the same application are rivalrous—that is, whether the receipt of one application affects the callback probability of another application.[36] We model resume rivalry as a spillover effect. Since we sent two resumes to each firm in our study, the spillover effect is the effect of one application on the probability of a callback for the other. We identify this spillover effect using the following regressions:

$$Y_{j,i} = \alpha + \tau T_{j,i} + \theta_0 T_{j,-i}(1 - T_{j,i}) + \theta_1 T_{j,-i} T_{j,i} + X'_{j,i}\gamma + \xi_{j,i} \tag{10}$$

where each vacancy is indexed by $j = 1, ..., J$ with 2 applications per vacancy, so that each application i to vacancy j has one other application. The variable $T_{j,-i} \in \{\text{Black}_{j,-i}$, Unemployed 12 months$_{j,-i}$, GED$_{j,-i}\}$ is the treatment status of the other application sent to vacancy j. Finally, $X_{j,i}$ is a vector of application characteristics.

Recall that vacancies were first randomized to be of type $R_j \in \{$Race pair, Unemployment duration pair$\}$. These regressions can be considered a partial population experiment. Potential outcomes in this experiment are now given by $Y_{j,i}(T_{j,i}, R_j)$. Vazquez-Bare (2022) shows that the parameters from this regression map into the direct and spillover effects of resume characteristics.

Here, τ is the direct effect of having a distinctively Black name, a long unemployment duration, or a GED. In the case of a distinctively Black name, it is the racial callback gap when no other fictitious resumes are sent to the firm. The coefficient θ_0 is the average

[36] Previous papers have examined this possibility and found mixed evidence. Abel (2017) finds that a firm's hiring decision in South Africa depends on the applicant pool's composition. Phillips (2019) finds evidence that applicants who compete against higher quality applicant pools receive more callbacks. However, in larger samples, Kline et al. (2022) do not find any evidence of spillovers across resumes.

spillover effect of having a treatment resume for control resumes (white, short unemployment duration, or high school). Similarly, θ_1 is the average spillover effect from the firm receiving other treated applications on treated applicants. If our applications within the establishment are not rivalrous, we would find $\tau_d \neq 0$ while $\theta_0 = 0$ and $\theta_1 = 0$. Estimates of Equation 10 appear in Table 4. Column (1) of Table 4 replicates Equation 4 with the spillover sample.[37] These estimates closely match those of the full sample. Columns (2) through (4) of Table 4 display estimates of Equation 10 with different definitions of treatment.

We cannot reject the null hypothesis of no spillovers onto treatment or control applications in any specification. Moreover, column (2) of Table 4 shows that the spillover coefficients are both an order of magnitude smaller than the direct effects. Including these terms reduces the precision of the estimate of the direct effect but does not affect its size. Column (3) shows that the spillovers of treatment resumes on other treatment and control resumes are larger in magnitude for the long unemployment duration. In contrast, column (4) shows that the spillovers from the GED treatments are also economically small.[38] Moreover, introducing these terms into the regression never meaningfully changes our estimate of the direct effect of race. Together, these results suggest very limited evidence of rivalry between applications.[39]

Finally, we investigate whether the effects of additional resumes affected callbacks differentially by race. Columns (5) and (6) in Table 4 show spillover effects from the unemployment duration and human capital treatments separately for Black and white applicants. We find that the spillover effects from these characteristics are an order of magnitude smaller than the direct effect of race. Importantly, we also find that the spillover effects are not significantly different for Black and white applicants, implying that the effect of different-quality resumes does not affect callbacks differentially by race.

[37] This sample excludes 2,595 establishments that only receive one application and four establishments that erroneously receive too many applications.

[38] We do not estimate all of the spillover effects in the same regression because spillovers from unemployment duration and race are colinear, given our randomization strategy.

[39] Appendix Table D2 provides an additional test that resumes are rivalrous with each other. In this table, we use the randomized order of the resumes to test whether there are order effects of receiving a resume and whether any potential order effects influence callbacks differentially by race. We find that resume order is not important, either overall, or differentially by race.

Table 4: Direct and Indirect Effects of Resume Characteristics

	(1) Callback	(2) Callback	(3) Callback	(4) Callback	(5) Callback	(6) Callback
Black_i, j	-1.389***	-1.266*	-1.406***	-1.389***		
	(0.280)	(0.671)	(0.429)	(0.280)		
Unemp 12 mo_i, j	-0.510	-0.510	-0.120	-0.510	0.028	-0.427
	(0.377)	(0.377)	(0.494)	(0.377)	(0.629)	(0.747)
GED_i, j	-0.530	-0.530	-0.530	-0.505	-0.555	-0.477
	(0.330)	(0.330)	(0.330)	(0.536)	(0.673)	(0.715)
Black_i, j × Black_$-i, j$		0.106				
		(0.680)				
White_ij × Black_$-i, j$		-0.171				
		(0.628)				
Unemp 1 mo_i, j × Unemp 12 mo_$-i, j$			0.357		0.230	0.369
			(0.557)		(0.762)	(0.810)
Unemp 12 mo_i, j × Unemp 12 mo_$-i, j$			-0.426		-0.587	-0.207
			(0.538)		(0.735)	(0.799)
HS_i, j × GED_$-i, j$				0.019	-0.043	0.074
				(0.546)	(0.692)	(0.717)
GED_i, j × GED_$-i, j$				-0.031	0.026	-0.083
				(0.534)	(0.662)	(0.711)
Constant	11.640***	11.560***	11.471***	11.631***	10.051***	11.506***
	(0.375)	(0.674)	(0.570)	(0.494)	(0.611)	(0.854)
Applicants	All	All	All	All	Black	White
R-squared	0.011	0.011	0.011	0.011	0.009	0.013
Observations	32,388	32,388	32,388	32,388	16,268	16,120

Notes: This table reports results from tests of the assumptions that applications at each job have no spillovers onto other applications at the same job. Column (1) displays estimates of Equation 4 using the sample with firms that receive one other resume. This excludes 2,595 establishments that received only one resume and four establishments that received more than two resumes. Columns (2) through (4) estimate versions of Equation 10 with the same sample to investigate spillovers. All specifications control for the applicant's age and the firm's city. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05.

All together, the results suggest that changes in the labor composition are not driving the results. We find limited evidence that resumes were rivalrous. Even if the resumes were rivalrous with each other, the direct effects of treatment are stable when controlling for potential spillovers. While this is not direct evidence that labor supply does not change, it provides some evidence that these potential changes do not impact our conclusions.

7. Concluding Remarks

We use a correspondence study to test whether increases in the minimum wage exacerbate racial labor market disparities. Before states announce that they will increase the minimum wage, applicants with distinctively Black names are 19 percent less likely to receive a callback than applicants with distinctively white names. After the announcements, the racial callback gap shrinks by 80 percent. We provide evidence that the gap decreases because white applicants are more likely to be on the margin of a callback, partly due to a less-dispersed distribution of perceived productivity. Both taste-based and statistical discrimination models predict that employers will call back a larger portion of relatively low-quality white applicants. When it becomes more costly to employ workers, these applicants are no longer called back.

Our framework provides a method to predict how hikes will change the racial callback gap. The data generated by a correspondence study without policy variation can capture animus and statistical discrimination in the distributions of perceived applicant productivity. Those parameters can provide sufficient information to learn which types of applicants are most at risk from potential future minimum wage increases.

Readers should interpret our results with several caveats in mind. First, we can only observe whether firms call back applicants for an interview. The evidence for whether callbacks are a valid surrogate for hiring is mixed (Quillian et al., 2020); nevertheless, it is illegal to discriminate in the callback stage. Second, we only collected data for a year after the minimum wage changed. These effects could differ from those in the long run (Sorkin, 2015; Aaronson et al., 2018). Third, our discussion of the mechanisms and prediction exercises depends on assumptions about the firm's beliefs about the distribution of productivity by race. These results are not generic to all distributions. Finally, racial differences in callbacks or hiring outcomes cannot speak to the full welfare consequences of minimum wage policies.

Nevertheless, our results suggest that policymakers should consider the minimum wage as one method of reducing discrimination against Black candidates. For the circumstances we estimate in Section 5, the results also suggest that other policies that increase the callback threshold will lead to less discrimination. In contrast, policies that reduce the callback threshold will have the unintended consequence of increasing discrimination.

However, given that Kline et al. (2022) shows that discrimination is concentrated among a small number of firms, policymakers should not always expect the same levels of discrimination parameters, or policy impacts, across settings. We show how one can estimate these moments without policy variation and use prior knowledge about the effect of the policy on callback thresholds to predict how the policy will affect labor market discrimination.

Finally, our results suggest that future research combining policy and experimental variation should consider announcement effects in their settings, as, in our setting, the conclusions of our study would have been reversed had we ignored these effects.

References

Aaronson, D., E. French, I. Sorkin, and T. To (2018). Industry dynamics and the minimum wage: a putty-clay approach. International Economic Review 59(1), 51–84.

Abel, M. (2017). Labor market discrimination and sorting: evidence from south africa. World Bank Policy Research Working Paper (8180).

Adams, C., J. Meer, and C. Sloan (2022). The minimum wage and search effort. Economics Letters 212, 110288.

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

Ahn, T., P. Arcidiacono, and W. Wessels (2011). The distributional impacts of minimum wage increases when both labor supply and labor demand are endogenous. Journal of Business & Economic Statistics 29(1), 12–23.

Aigner, D. J. and G. G. Cain (1977). Statistical theories of discrimination in labor markets. Ilr Review 30(2), 175–187.

Allegretto, S., A. Dube, M. Reich, and B. Zipperer (2017). Credible research designs for minimum wage studies: A response to neumark, salas, and wascher. ILR Review 70(3), 559–592.

Allegretto, S. A., A. Dube, and M. Reich (2011). Do minimum wages really reduce teen employment? accounting for heterogeneity and selectivity in state panel data. Industrial Relations: A Journal of Economy and Society 50(2), 205–240.

Alsalam, N. (2019). The effects on employment and family income of increasing the federal minimum wage. Congressional Budget Office.

Altonji, J. G. and C. R. Pierret (2001). Employer learning and statistical discrimination. The Quarterly Journal of Economics 116(1), 313–350.

Angrist, J. D. (2014). The perils of peer effects. Labour Economics 30, 98–108.

Arrow, K. J. (1973). The Theory of Discrimination, pp. 3–33. Princeton University Press.

Autor, D. H. and D. Scarborough (2008). Does job testing harm minority workers? evidence from retail establishments. The Quarterly Journal of Economics 123(1), 219–277.

Baert, S. (2018). Hiring discrimination: An overview of (almost) all correspondence experiments since 2005. Audit studies: Behind the scenes with theory, method, and nuance, 63–77.

Bailey, M. J., J. DiNardo, and B. A. Stuart (2021). The economic impact of a high national minimum wage: Evidence from the 1966 fair labor standards act. Journal of Labor Economics 39(S2), S329–S367.

Ballotpedia (2018a). Arkansas issue 5, minimum wage increase initiative (2018). https://ballotpedia.org/Arkansas_Issue_5,_Minimum_Wage_Increase_Initiative_(2018). Accessed: February 24, 2023.

Ballotpedia (2018b). Missouri proposition b, $12 minimum wage initiative (2018). https://ballotpedia.org/Missouri_Proposition_B,_$12_Minimum_Wage_Initiative_(2018). Accessed: February 24, 2023.

Bartoš, V., M. Bauer, J. Chytilová, and F. Matějka (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. American Economic Review 106(6), 1437–75.

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

Benson, A., S. Board, and M. Meyer-ter Vehn (2022). Discrimination in hiring: Evidence from retail sales. Available at SSRN 4179847.

Benson, A. and L. P. Lepage (2022). Learning to discriminate on the job. Available at SSRN 4155065.

Bertrand, M. and E. Duflo (2017). Field experiments on discrimination. Handbook of economic field experiments 1, 309–393.

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

BLS (2018). Labor force statistics from the current population survey: Unemployment rates by age, sex, race, and hispanic or latino ethnicity.

Bohren, J. A., K. Haggag, A. Imas, and D. G. Pope (2019). Inaccurate statistical discrimination: An identification problem. Technical report, National Bureau of Economic Research.

Brown, C. (1984). Black-white earnings ratios since the civil rights act of 1964: The importance of labor market dropouts. The Quarterly Journal of Economics 99(1), 31–44.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics 90(3), 414–427.

Canay, I. A., A. Santos, and A. M. Shaikh (2021). The wild bootstrap with a "small" number of "large" clusters. Review of Economics and Statistics 103(2), 346–363.

Card, D. and A. B. Krueger (1992). School quality and black-white relative earnings: A direct assessment. The quarterly journal of Economics 107(1), 151–200.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer (2019). The effect of minimum wages on low-wage jobs. The Quarterly Journal of Economics 134(3), 1405–1454.

Charles, K. and J. Guryan (2008). Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination. Journal of Political Economy 116(5), 773–809.

Clemens, J., L. B. Kahn, and J. Meer (2021). Dropouts need not apply? the minimum wage and skill upgrading. Journal of Labor Economics 39(S1), S107–S149.

Clemens, J. and M. Wither (2019). The minimum wage and the great recession: Evidence of effects on the employment and income trajectories of low-skilled workers. Journal of Public Economics 170, 53–67.

Crépon, B., F. Devoto, E. Duflo, and W. Parienté (2015). Estimating the impact of microcredit on those who take it up: Evidence from a randomized experiment in morocco. American Economic Journal: Applied Economics 7(1), 123–50.

David, H., A. Manning, and C. L. Smith (2016). The contribution of the minimum wage to us wage inequality over three decades: a reassessment. American Economic Journal: Applied Economics 8(1), 58–99.

Derenoncourt, E. and C. Montialoux (2021). Minimum wages and racial inequality. The Quarterly Journal of Economics 136(1), 169–228.

Doleac, J. L. and L. C. Stein (2013). The visible hand: Race and online market outcomes. The Economic Journal 123(572), F469–F492.

Donohue, J. J. and J. Heckman (1991). Continuous versus episodic change: The impact of civil rights policy on the economic status of blacks. Journal of Economic Literature 29(4), 1603–1643.

Dube, A., T. W. Lester, and M. Reich (2011). Do frictions matter in the labor market? accessions, separations and minimum wage effects.

Dube, A., T. W. Lester, and M. Reich (2016). Minimum wage shocks, employment flows, and labor market frictions. Journal of Labor Economics 34(3), 663–704.

Duflo, E. and E. Saez (2003). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. The Quarterly journal of economics 118(3), 815–842.

Eriksson, S. and D.-O. Rooth (2014). Do Employers Use Unemployment as a Sorting Criterion When Hiring? Evidence from a Field Experiment. American Economic Review 104(3), 1014–1039.

Flinn, C. J. (2006). Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates. Econometrica 74(4), 1013–1062.

Friedman, M. (1966). Minimum-wage rates. Newsweek.

Fryer, R. G., D. Pager, and J. L. Spenkuch (2013). Racial disparities in job finding and offered wages. The Journal of Law and Economics 56(3), 633–689.

Gaddis, S. M. (2017). How black are lakisha and jamal? racial perceptions from names used in correspondence audit studies. Sociological Science 4, 469–489.

Gneezy, U., J. List, and M. K. Price (2012). Toward an understanding of why people discriminate: Evidence from a series of natural field experiments. Technical report, National Bureau of Economic Research.

Gopalan, R., B. H. Hamilton, A. Kalda, and D. Sovich (2021). State minimum wages, employment, and wage spillovers: Evidence from administrative payroll data. Journal of Labor Economics 39(3), 673–707.

Guryan, J. and K. K. Charles (2013). Taste-based or statistical discrimination: the economics of discrimination returns to its roots. The Economic Journal 123(572), F417–F432.

Heckman, J. J. (1998). Detecting discrimination. Journal of economic perspectives 12(2), 101–116.

Heckman, J. J. and P. Siegelman (1993). The urban institute audit studies: Their methods and findings. Clear and convincing evidence: Measurement of discrimination in America, 187–258.

Heckman, J. J. and E. J. Vytlacil (2007). Econometric evaluation of social programs, part i: Causal models, structural models and econometric policy evaluation. Handbook of econometrics 6, 4779–4874.

Holz, J., R. Rivera, and B. A. Ba (2019). Spillover effects in police use of force. U of Penn, Inst for Law & Econ Research Paper (20-03).

Indeed.com (2020). Indeed named the leading source of hire in north america.

Jardim, E., M. C. Long, R. Plotnick, E. Van Inwegen, J. Vigdor, and H. Wething (2022). Minimum-wage increases and low-wage employment: Evidence from seattle. American Economic Journal: Economic Policy 14(2), 263–314.

Jha, P., D. Neumark, and A. Rodriguez-Lopez (2022). What's across the border? re-evaluating the cross-border evidence on minimum wage effects.

Johnson, R. C. (2019). Children of the dream: Why school integration works. Basic Books.

Jowell, R. and P. Prescott-Clarke (1970). Racial discrimination and white-collar workers in britain. Race 11(4), 397–417.

Kline, P., E. K. Rose, and C. R. Walters (2022). Systemic discrimination among large us employers. Technical Report 4.

Kreiner, C. T., D. Reck, and P. E. Skov (2020). Do lower minimum wages for young workers raise their employment? evidence from a danish discontinuity. Review of Economics and Statistics 102(2), 339–354.

Kroft, K., F. Lange, and M. J. Notowidigdo (2013). Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*. The Quarterly Journal of Economics 128(3), 1123–1167.

Lee, D. and E. Saez (2012). Optimal minimum wage policy in competitive labor markets. Journal of Public Economics 96(9-10), 739–749.

Lepage, L. P. (2021). Endogenous learning, persistent employer biases, and discrimination. Persistent Employer Biases, and Discrimination (March 2, 2021).

Lillard, L., J. P. Smith, and F. Welch (1986). What do we really know about wages? the importance of nonreporting and census imputation. Journal of Political Economy 94(3, Part 1), 489–506.

List, J. A. (2004). The nature and extent of discrimination in the marketplace: Evidence from the field. The Quarterly Journal of Economics 119(1), 49–89.

Luca, D. L. and M. Luca (2019). Survival of the fittest: the impact of the minimum wage on firm exit. Technical report, National Bureau of Economic Research.

Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles (2022). Ipums national historical geographic information system: Version 16.0. American Economic Review.

McCall, J. J. (1970). Economics of information and job search. The Quarterly Journal of Economics, 113–126.

Meer, J. and J. West (2016). Effects of the minimum wage on employment dynamics. Journal of Human Resources 51(2), 500–522.

Neumark, D. (2012). Detecting discrimination in audit and correspondence studies. Journal of Human Resources 47(4), 1128–1157.

Neumark, D., I. Burn, and P. Button (2019). Is it harder for older workers to find jobs? new and improved evidence from a field experiment. Journal of Political Economy 127(2), 922–970.

Neumark, D., I. Burn, P. Button, and N. Chehras (2019). Do state laws protecting older workers from discrimination reduce age discrimination in hiring? evidence from a field experiment. The Journal of Law and Economics 62(2), 373–402.

Neumark, D. and J. Rich (2019). Do field experiments on labor and housing markets overstate discrimination? a re-examination of the evidence. ILR Review 72(1), 223–252.

Neumark, D., W. L. Wascher, et al. (2007). Minimum wages and employment. Foundations and Trends® in Microeconomics 3(1–2), 1–182.

Nunley, J. M., A. Pugh, N. Romero, and R. A. Seals (2015). Racial discrimination in the labor market for recent college graduates: Evidence from a field experiment. The BE Journal of Economic Analysis & Policy 15(3), 1093–1125.

Pager, D. (2016). Are firms that discriminate more likely to go out of business? Sociological Science 3, 849–859.

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

Phelps, E. S. (1972). The statistical theory of racism and sexism. The american economic review 62(4), 659–661.

Phillips, D. C. (2019). Do Comparisons of Fictional Applicants Measure Discrimination When Search Externalities are Present? Evidence from Existing Experiments. The Economic Journal 129(621), 2240–2264.

Phillips, D. C. (2020). Do low-wage employers discriminate against applicants with long commutes? evidence from a correspondence experiment. Journal of Human Resources 55(3), 864–901.

Quillian, L., J. J. Lee, and M. Oliver (2020). Evidence from field experiments in hiring shows substantial additional racial discrimination after the callback. Social Forces 99(2), 732–759.

Shrider, E. A., M. Kollar, F. Chen, J. Semega, et al. (2021). Income and poverty in the united states: 2020. US Census Bureau, Current Population Reports (P60-273).

Simon, A. and M. Wilson (2021). Optimal minimum wage setting in a federal system. Journal of Urban Economics 123, 103336.

Smith, J. P. and F. R. Welch (1989). Black economic progress after myrdal. Journal of economic literature 27(2), 519–564.

Smythe, A. and L. Hsu (2023). The minimum wage as a tool for racial economic justice. Journal of Economic Literature 61(3), 977–87.

Sorkin, I. (2015). Are there long-run effects of the minimum wage? Review of economic dynamics 18(2), 306–333.

Stiglitz, J. E. (1973). Approaches to the economics of discrimination. The American Economic Review 63(2), 287–295.

Turner, M. D. and B. Demiralp (2001). Do higher minimum wages harm minority and inner-city teens? The Review of Black Political Economy 28(4), 95–116.

Tzioumis, K. (2018). Demographic aspects of first names. Scientific data 5(1), 1–9.

Vazquez-Bare, G. (2022). Identification and estimation of spillover effects in randomized experiments. Journal of Econometrics.

Wursten, J. and M. Reich (2021). Racial inequality and minimum wages in frictional labor markets.

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

For Online Publication: Appendix

A. Details of Experimental Design

A.1 Resume Characteristics and Creation

Names: The names in the experiment are drawn from combinations of first and last names used in Agan and Starr (2018). Each resume was randomly assigned a first and last name from the designated race with replacement. The list of first and last names used in the experiment appears in Table A1 below.

Table A1: First and Last Names Assigned by Race

First Name		Last Name	
Distinctively Black	Distinctively White	Distinctively Black	Distinctively White
Daquan	Cody	Alston	Brennan
Darnell	Douglas	Banks	Fox
Darryl	Dylan	Bryant	Hansen
Denzel	Jacob	Byrd	Hoffman
Dwayne	John	Charles	Kane
Elijah	Kyle	Fields	Meyer
Isaiah	Matthew	Franklin	O'Niell
Jamal	Nicholas	Hawkins	Romano
Jaquan	Ryan	Ingram	Russo
Jermaine	Scott	Jackson	Ryan
Malcolm	Sean	Jenkins	Schmidt
Marquis	Shane	Darryl	Snyder
Maurice	Stephen	Joseph	Sullivan
Reginald	Thomas	Pierre	Wagner
Terrance	Tyler	Robinson	Weber
Terrell		Simons	
Tyree		Washington	
Tyrone		Williams	

Notes: This table lists the first and last names assigned by race. The names were taken from Agan and Starr (2018).

To assess whether these names signal the intended race, we use the procedure from Kaplan (2021) to predict the posterior probability that an individual with a given first and last name is of the intended race. The probability for each first name is averaged over each time the name is used in conjunction with the last name in Table A2. As this table shows, the empirical probability that individuals with the names we chose are of the intended race is extremely high. These values are slightly higher for resumes with distinctly Black names than those with distinctly white names. However, we estimate ITT effects from our experiment, meaning that any interpretation by firms that resumes with distinctly white names are nonwhite would bias our estimates toward zero.

Table A2: Empirical Likelihood That First Names Signal the Intended Race

Distinctively Black		Distinctively White	
First Name	$\mathbb{E}[Black \mid \text{first name, surname}]$	First Name	$\mathbb{E}[White \mid \text{first name, surname}]$
Daquan	99.3	Cody	87.5
Darnell	99.3	Douglas	91.3
Denzel	98.7	Jacob	79.3
Dwayne	96.7	John	82.9
Elijah	97.9	Kyle	83.4
Isaiah	97.2	Matthew	85.8
Jamal	99.2	Nicholas	76.7
Jaquan	99.6	Ryan	82.7
Jermaine	98.5	Scott	94.6
Malcolm	91.4	Sean	77.6
Marquis	99.5	Shane	88.7
Reginald	97.9	Thomas	94.1
Terrance	98.6	Tyler	89.4
Terrell	97.3		
Tyree	95.7		
Tyrone	98.6		
Total	98.6		85.6

Notes: This table reports the first names used in the experiment and the probability that a person with a given name is of the intended race. We calculate the probabilities by recovering the posterior probability that a person with a given first and last name is of the intended race using Kaplan (2021). Then, we average this probability over all iterations of the first names used in the experiment.

Locations. After assigning names to races, we randomly assigned each application to be from one of four cities: St. Louis, Springdale/Fayetteville, Little Rock, or Springfield. In wave two, we repeated this process for the new cities we added to the experiment: Kansas City (Kansas and Missouri) and East St. Louis, Illinois. Table A3 shows the average posterior probability that a resume sent in a given state throughout the experiment is of the intended race. Similar to the overall sample, the probability that a resume is of the intended race is exceedingly high for both Black and white resumes but higher for Black resumes. Moreover, these probabilities are similar across states for both Black and white resumes.

Table A3: Empirical Likelihood That Names Signal the Intended Race by State

	Black Resumes				White Resumes			
	AR	MO	KS	IL	AR	MO	KS	IL
$\mathbb{E}[Black \mid \text{first name, surname}]$	97.9	97.4	98.4	97.4	4.3	2.9	4.1	5.2
	(2.7)	(2.3)	(1.8)	(3.5)	(4.7)	(3.6)	(4.4)	(5.9)
$\mathbb{E}[White \mid \text{first name, surname}]$	1.3	1.6	0.8	1.8	84.3	86.0	84.1	84.2
	(2.2)	(1.7)	(1.2)	(2.8)	(7.9)	(5.9)	(8.6)	(7.4)
Observations	4,527	8,050	2,825	1,955	4,471	8,159	2,844	1,878

Notes: This table reports, by state, the estimated probability that a person with a given name is of a given race. Standard deviations are in parentheses. We calculate the probabilities by recovering the posterior probability using Kaplan (2021) with the subject's first name and surname. Then, we average this probability over all applications sent out in the experiment.

Unemployment Duration. Within each city and race, we assign half of the resumes to have an unemployment duration of 1 month and half of the cities to have an unemployment duration of 12 months. We chose these two unemployment durations because Kroft et al. (2013) found the largest difference in callbacks between these two lengths. We operationalize these durations by making the employment end date of their previous job to be either 1 or 12 months before the month that our experiment started. Regularly after the beginning of the experiment, we would increase the end date of the previous employment history to maintain the 1-month and 12-month unemployment duration.

Ages and Years of Experience. We assigned ages and years of experience to applicants based on their unemployment duration. If an application was given an unemployment duration of 12 months, we gave the applicant a birth year 20 years before the beginning of the experiment. These individuals were also given an employment start date one year before their employment end date, to provide them with one year of work experience. If an application was given an unemployment duration of one month, we randomly assigned half of the applications to be from people 19 years old and half of the applications to be from people 20 years old. If the applicant was 19 years old, we assigned that person one year of work experience; if the applicant was 20 years old, we assigned the person two years of work experience by having a start date two years before his or her employment end date.

Contact Information. Each application provided firms with three different ways to contact the applicant. First, since each application was sent through Indeed, firms could contact the applicants directly through their Indeed accounts. Second, we manually created e-mail addresses for all of our applicants. Each e-mail account was associated with a resume and was created using combinations of the applicant's first name, last name, and arbitrary integers. These e-mail addresses were also used as the login information for the Indeed accounts.

Finally, we provided provisional phone numbers from Tresta. We chose phone numbers with area codes local to the cities we operated in, so that each application had a local phone number. In waves one through three, we assigned phone numbers to resumes based on their type, defined by the applicant's race, unemployment duration, and human capital attainment. Our randomization strategy ensured that employers would not see two applicants with the same phone number. In waves four and five, we increased the number of phone numbers so that each resume had a unique phone number. Phone calls to each number were automatically directed to a voicemail with a standard, nonpersonalized message through Tresta.

Addresses. We assigned each application a home address in the cities where we submitted applications. To obtain the addresses, we went on apartment finder websites like Zillow.com, Apartments.com, and Domu.com and found one-bedroom apartments with rental prices that were a third of a full-time minimum wage worker's income for a month. We always chose apartments near the cities we operated in, because Phillips (2020) found evidence that firms are less likely to call back applicants with long commutes. Addresses were randomly assigned to applications without replacement.

Educational History. We randomly assigned each application to have a high school diploma or a GED within each race, unemployment duration, and city. If the individual had a high school diploma, we recorded the graduation date as 18 years after the birth date, as determined by age. We assigned each high school graduate to one of five local high schools within each city and race unemployment-duration pair. If the applicant had a GED, we did not list a high school on the application.

Employment History. Each applicant was assigned to one previous employer. To obtain a list of potential employers, we downloaded the available resumes from other applicants on Indeed.com seeking employment at low-wage jobs in the cities we applied to and the surrounding areas. For each previous job, we assigned the job title of a cashier, food preparation, team member, store associate, or janitor.

We only chose job titles that were relevant to the previous employer. For example, an applicant whose previous employment was at Walgreens could be a team member but could not work in food preparation. However, someone who had previously worked at Arby's could have worked in food preparation. For each of the four job types, we generated a bank of generic work descriptions based on those from the other applications we found on Indeed.com and randomly assigned these descriptions to each resume. For example, a store associate's work description might have been "Cleaned and stocked the store. Provided customer service," and someone who worked in food preparation would have a description similar to "Took food orders, handled payments, and prepared food."

Miscellaneous Resume Characteristics. Resumes also included other small characteristics like a mission statement, which we randomly added to each resume based on a bank of statements we created in line with those posted on Indeed.com. It was common for applications to require idiosyncratic questions, which we could not prepare answers to beforehand. Following Agan and Starr (2018), we instructed RAs, who were blind to the purpose of the study, to answer these questions positively and in line with their judgment. Before they began, we instructed these research assistants about what employers generally look for in an application and how to answer these questions in a way that would increase the likelihood of receiving a callback. We did not include any references on the job applications because it was expensive to create additional phone numbers for fictitious references. Moreover, Agan and Starr (2018) found that no employers ever called the phone numbers for their references, suggesting that employers did not pay attention to them.

A.2 Job Sampling

We developed code that scraped vacancies from Indeed regularly throughout the experimental period. During waves one through three, the scraping was done every day. During waves four and five, we changed this process to once weekly. The sampling procedure followed the following process.

1. Go to Indeed.com advanced job search.

2. Limit jobs to those with salary estimates of $16,000 to $21,000. This range corresponds to $7.70–$10.10 per hour or jobs with the following keywords: "retail," "food+prep," "fast+food," "janitor," "maid," "cashier," "retail+salespeople," "cooks," and "building+cleaners."

3. Limit the location to within 10 miles of the following areas: Little Rock, Ark.; St. Louis, Mo.; Springfield, Mo.; Fayetteville, Ark.; Springdale, Ark.; Kansas City, Mo.; Kansas City, Kan.; East St. Louis, Ill.; Granite City, Ill.; Overland Park, Kan.; Shawnee, Kan.; Olathe, Kan.; Cahokia, Ill.; Washington Park, Ill.; Alton, Ill.; and Belleville, Ill.

4. Use a fuzzy merge with a manual check to ensure that we remove jobs from the same establishments during the same wave.

5. Exclude jobs that are unusual or seem not to be minimum wage jobs (e.g., jobs with the National Guard).

6. When more jobs are available than we have the ability to apply for in that period, we prioritize the jobs that were posted most recently.

A.3 Application Submission

After we have the set of jobs we intend to submit applications to, we randomly assign resumes to the applications using the following approach.

1. For each job, independently and randomly assign the applicant's race, unemployment status, and educational status with a probability of 1/2.

2. Randomly draw profiles with replacement from the set of resumes created using the procedure in Appendix A.1.

3. Generate an alternate resume of the same type as in the previous step if the randomly chosen applicant has a work history at the vacancy firm. When applying to firms, research assistants manually checked whether the application firm and the firm on the resume's work history match. If they matched, the research assistant applied with the alternate resume.

4. Randomize, with 50 percent probability, whether the other resume sent to the firm is of the opposite race or the opposite unemployment duration. Other characteristics are randomized.

5. Randomly draw profiles with replacements from the set of resumes that share the same type determined in the previous step.

6. Generate an alternate resume for the second resume using a similar procedure as the alternate for the first resume.

After each randomization, we gave our research assistants a list of jobs, a link to the job ad, and the Indeed login information they needed to apply for the job. With some time lag, a second research assistant sent the second resume to the same firm using the second profile. Each research assistant was given a spreadsheet with only the information they needed to complete their tasks, so they were not aware of the identity of the other fictitious applicants. The randomization procedure ensured that the characteristics of the resume were independent of the application order.[40]

[40] Table D2 shows that there is no evidence application order plays a role in a firm's callback decision.

It was not always possible to send a complete set of two resumes to the same firm within a wave. It was also not always possible to apply to the same firm across different waves. This was mostly due to firms removing their job ad from Indeed. There were also new jobs posted continuously throughout the experiment. As a result, the sample sizes change heterogeneously across waves.

A.4 Measuring Outcomes

We assigned each e-mail and Indeed message to a firm using the information provided by the employer. In the few instances when the employer did not provide enough information to assign the application, we responded to the e-mail and asked the employer for the information. Voicemails were reviewed by the research team and assigned to a firm using the information provided by the employer. When the employer did not leave enough information to match the application, we called back the employer from a different number, told them we received a call from that number, and asked for the firm's identity. We recorded all callbacks that occurred until August 2020. Regardless of the callback date, we assigned the callback based on the application date. We investigate the implications of this decision in Appendix E.

Following Bertrand and Mullainathan (2004) and Kroft et al. (2013), we instructed research assistants to record any contact from the employers from these three sources. For contact to be considered a "callback," we required that the firm make an explicit request for the individual to come in for an interview. We do not classify other types of communication, such as clarification about a question on the application, as a callback. This choice was made because it is difficult to connect this type of communication to discrimination. For example, an employer who reaches out for additional information from a minority application more frequently may do this either because she is more interested in hiring the minority applicant or because she screens the minority applicant more carefully before hiring. Conversely, interview requests are a stronger signal that the employer is interested in hiring the applicant. We investigate robustness to this decision in Appendix E.

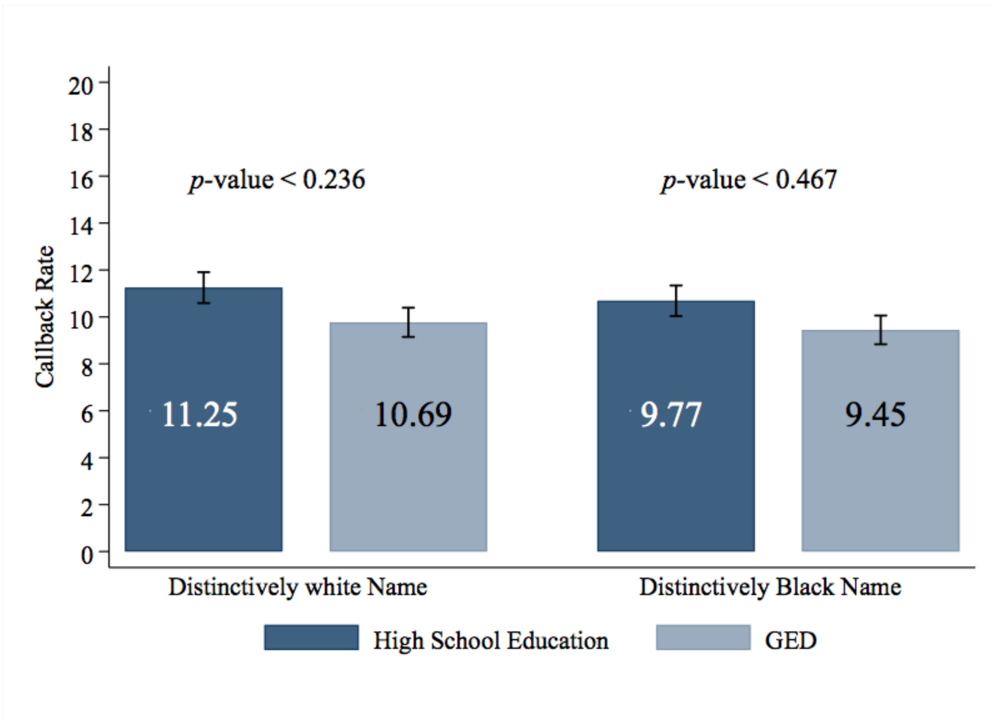
B. Applicant Names Potentially Signaling Socio-Economic Status

The primary treatment in this paper is the manipulation of the name on the resume. While we have interpreted differences in outcomes resulting from the name as racial effects, the name may also signal other characteristics (Cook et al., 2014, 2016). Indeed, Fryer and Levitt (2004) find that children with distinctively Black names tend to have lower socioeconomic status (SES) than other children. Moreover, Gaddis (2015) finds that applicants with distinctively Black names that suggest the mother had low education are especially penalized. Similarly, Kreisman and Smith (2022) finds that test scores, college enrollment, and college completion are negatively correlated with Black names across households but not within households, suggesting that the applicants with distinctively Black and white names may be different along these dimensions even after conditioning on observables.

Despite these findings, previous correspondence studies have found that using distinctively Black and white names successfully signals the applicant's race without unintentionally signalling SES (Bertrand and Mullainathan, 2004; Kline et al., 2022). This may be because employers do not know the correlation between names and productivity or residual differences, conditional on resume characteristics, so the Black name index does not predict productivity. We believe this confounding factor is even less of a concern in our setting, as we apply to low-wage, entry-level jobs that likely primarily attract applicants of low socioeconomic status. Our resumes also provide the applicant's address, work history, educational attainment, and high school. These perhaps more concrete signals of SES likely reduce the employer's potential reliance on the applicant's name as a signal of SES. Nevertheless, we investigate the possibility that the experiment measures disparities by SES rather than racial disparities.

If distinctively Black names primarily signal SES, we would expect that applications with other attributes that signal high SES would benefit Black applicants more than white applicants. Perhaps the strongest signal of social status in our experiment is whether the applicant has a GED instead of a high school degree. Figure B1 displays callback rates by race and whether the applicant has a GED. We find that the returns to a high school education are similar for white and Black applicants and, if anything, a little larger for Black applicants. The lack of treatment-effect heterogeneity on this dimension is inconsistent with the name on the resume signaling SES.

Figure B1: High School Education Effect by Race



Notes: This figure displays the callback rates by the race of the applicant separately by whether the applicant has a high school education or GED. Error bars represent 95 percent confidence intervals. p-values refer to tests showing that the difference between education status within applicant race is significantly different from zero.

Next, we link each applicant's address to a census tract and obtain information about individuals living in the tract containing the fictitious applicant's address from IPUMS (Manson et al. (2022)). We also obtain information about the applicant's high school Great Schools Score and the school's expenditure per student from greatschools.org for the applications where the applicant had a high school degree. The experiment used 34 high schools and 146 unique census tracts. For each of the census tracts, we obtain information on variables associated with socioeconomic status. Table B1 lists these variables and their summary statistics.

Table B1: Socioeconomic Status Variables

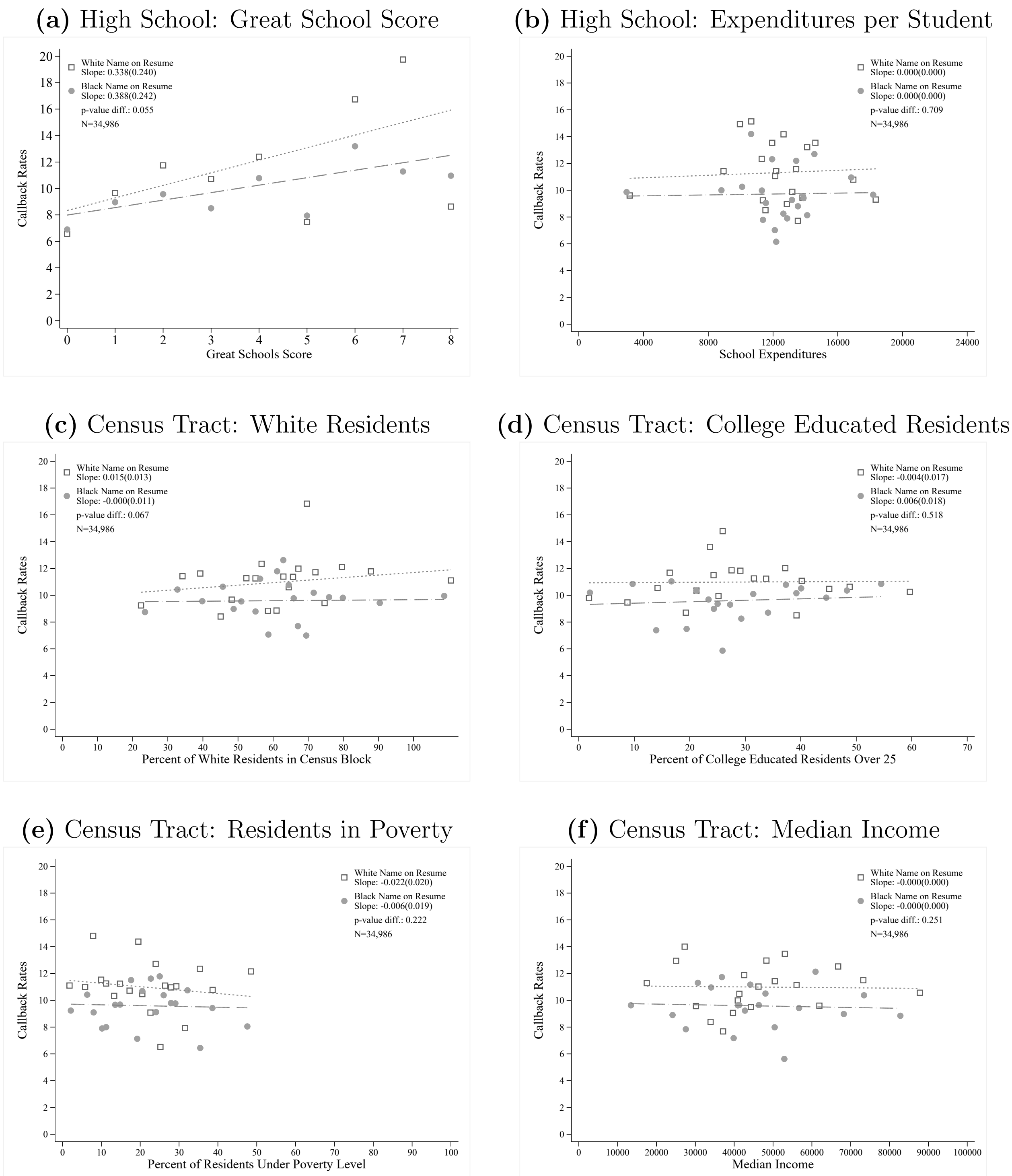
	Observations	Mean	St. Dev.	10^{th} ptile	90^{th} ptile
Great School Summary Rating	17,581	3.1	2.3	1.0	6.0
Expenditures Per Student ($)	17,581	12,046	4,398	9949	16552
White Residents (%)	34,986	61.0	26.1	19.5	88.1
College Educated Residents (%)	34,986	28.3	15.1	10.1	47.7
Under Poverty Line (%)	34,986	21.4	12.5	6.50	39.4
Median Household Income ($)	34,986	45,822	19,090	23,078	73,872
Median Rent ($)	34,986	857	166	666	1,082
Missing High School Characteristics (%)	34,986	53.4	49.9	0.0	100
Missing any Census Variable (%)	34,986	0.93	9.59	0.0	0.0

Notes: This table displays the summary statistics for socioeconomic status variables. High school variables are obtained from greatschools.org. The value of these variables is missing if the subject has a GED rather than a high school diploma, or if the high school does not appear on greatschools.org. Census-tract-level variables are obtained from Manson et al. (2022) and linked to the resume through the address that appears with the application.

Table B1 shows substantial variation in socioeconomic status signals across resumes as measured by the resume's census tract and high school characteristics. We can use this variation to understand whether distinctively Black names primarily signal race or socioeconomic status. A necessary condition for SES signals to confound our results is that employers value hiring applicants of higher SES. While plausible in other settings, we find limited evidence of this in the labor markets we examine.

Figure B2 displays bin scatters of our SES variables and callback rates separately for resumes with distinctively Black and white names. We only find a positive association between the Great School Score and the portion of white residents and callback rates. In contrast, we can detect no significant correlations between high school expenditure per student, the whiteness of the census block, how college-educated the census block is, the portion of the census block living in poverty, or the census block's median income and callback rates. While these relationships are not identified, they suggest modest returns to high SES in this labor market.

Figure B2: Relationship between SES Variables and Callbacks by Race



(a) High School: Great School Score
(b) High School: Expenditures per Student
(c) Census Tract: White Residents
(d) Census Tract: College Educated Residents
(e) Census Tract: Residents in Poverty
(f) Census Tract: Median Income

Notes: This figure displays the relationships between socioeconomic-status variables and callbacks separately by the applicant's race. High school variables are obtained from greatschools.org. The value of these variables is missing if the subject has a GED rather than a high school diploma, or if the high school does not appear on greatschools.org. Census-tract-level variables are obtained from Manson et al. (2022) and linked to the resume through the address that appears with the application.

If distinctively Black names primarily signal socioeconomic status, we would also expect that applications with distinctively Black names are helped more by living in high SES areas or by coming from a higher-quality school. To examine this supposition, we test whether the correlations between our SES variables and callback rates are stronger for applicants with distinctively Black names. In all instances, we either find no evidence of a correlation between SES variables and callback rates for both races or that the correlation is more positive for applications with distinctively white names. These results suggest that Black names are not helped more by living in areas with high SES, thus implying that Black names do not primarily signal social status.

Finally, we investigate whether controlling for our SES measures affects our main results' conclusions. Table B2 shows estimates of our main effects. Columns (1) through (3) display the overall effect of the applicant's race on callbacks, adding either our SES variables or high school and census-tract fixed effects. Columns (4) through (6) display the effects of the minimum wage, adding these same variables. If distinctively Black names primarily signal socioeconomic status, we would expect that adding these variables would lessen the relationship between the race coefficients and callbacks. But we find no evidence that this is the case, providing more evidence that names in our study signal race rather than socioeconomic status.

Table B2: Callback Rate and Socioeconomic Status

	(1) Callback	(2) Callback	(3) Callback	(4) Callback	(5) Callback	(6) Callback
Black	-1.39***	-1.44***	-1.09**	-3.21***	-3.23***	-3.16***
	(0.27)	(0.28)	(0.36)	(0.61)	(0.62)	(0.78)
After Announcement				-9.09***	-8.99***	-8.57***
				(0.82)	(0.83)	(0.89)
Black × After Announce				2.58***	2.52***	2.74**
				(0.68)	(0.68)	(0.86)
Unemp 12 mo.	-0.44	-0.47	-0.15	-0.46	-0.51	-0.21
	(0.36)	(0.37)	(0.47)	(0.36)	(0.37)	(0.47)
GED	-0.46	1.78*		-0.49	1.16	
	(0.32)	(0.82)		(0.31)	(0.81)	
Constant	11.43***	15.70***	10.91***	17.91***	22.89***	17.09***
	(0.36)	(2.00)	(0.36)	(0.74)	(2.12)	(0.79)
SES Controls	NO	YES	NO	NO	YES	NO
SES Fixed Effects	NO	NO	YES	NO	NO	YES
R-squared	0.011	0.011	0.020	0.018	0.018	0.026
Observations	34,986	34,986	34,986	34,986	34,986	34,986

Notes: This table displays estimates of the effects of resume characteristics and minimum wage changes on the callback rate. *SES Controls* refers to whether our set of SES controls, presented in Table B1 , are included in the regression. *SES Fixed Effects* refers to whether resumes include either a GED or one of 35 different randomized high schools, along with fixed effects for the census tract of the address randomly assigned to the applicant. There are 146 different census tracts in the sample. All specifications include city fixed effects and control for applicant age. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05.

C. Sample Characteristics and Treatment Balance

In this section, we describe the sample characteristics and present statistics about treatment balance. Column (1) of Table C1 shows the average characteristics for the firms in our sample across all applications. The table is broken down into four panels. Panel (a) displays information about whether job hours are part time, full time, or unstated, along with whether the job posts an hourly wage and the estimated hourly wage conditional on posting it. Because some jobs state that both full- and part-time jobs are available, these categories are not mutually exclusive. Panel (b) displays the job skills requested in the job ad, calculated using the procedure from Spitz-Oener (2006). *Job tasks* refers to the types of tasks requested in the application text, using the procedure from Atalay et al. (2020). Finally, the job titles from the ad are also shown in Panel (d). Figure C1 illustrates the most frequent job titles displayed in job ads of our sample.

Overall, about 36 percent of jobs post a wage. We believe that the majority of these jobs pay minimum wage and choose not to post it because that is not a selling point of the job. Of those that post it, the mean wage is $13 an hour. The median wage in our sample is lower, $10. This value should be interpreted with caution because we have to estimate the hourly wage for many jobs based on the stated weekly, monthly, or annual salary and the expected wage.

Columns (2) and (3) of Table C1 show the average characteristics across the race of the applicant. All characteristics shown are determined pretreatment and should not be affected by the treatment assignment. There is variation in the means because only some firms receive one resume of each type (see Section 2.2). Column (3) reports p-values for the null hypothesis that the average characteristics are equal across the two treatment categories. As is consistent with successful random assignment, the observable characteristics are balanced across treatment groups.

In Table C2, we present an alternative version of the randomization balance test, breaking down the sample by the resume type. This type is defined by race, unemployment duration, and educational attainment. Similarly to Table C1, this shows that overall, the observable characteristics of advertisements are balanced across treatments. In the two instances where we can detect differences in means across resume types, the differences are not economically meaningful.

Figure C1: Job Titles

Notes: This figure displays the most frequent phrases that appear in job vacancy titles. Sizes of the text correspond to frequency counts.

Table C1: Race Treatment Balance across the Whole Sample Period

		Treatment Arm		
	All	Black Applicant	White Applicant	p-value test
	(1)	(2)	(3)	(4)
a. Job Hours and Pay:				
Part Time Job (%)	34.831	34.703	34.960	0.478
	(0.255)	(0.359)	(0.361)	
Full Time Job (%)	40.916	40.914	40.919	0.990
	(0.263)	(0.371)	(0.372)	
Job Hours Unstated (%)	38.120	38.230	38.011	0.560
	(0.260)	(0.367)	(0.368)	
Has Posted Wage (%)	36.669	36.566	36.772	0.580
	(0.258)	(0.364)	(0.365)	
Estimated Hourly Wage (Dollars)	13.010	13.050	12.971	0.433
	(0.071)	(0.101)	(0.100)	
b. Job Skills:				
Social Skills Demanded (%)	57.795	57.946	57.642	0.422
	(0.264)	(0.373)	(0.374)	
Customer Skills Demanded (%)	78.443	78.170	78.718	0.087
	(0.220)	(0.312)	(0.310)	
Character Skills Demanded (%)	44.724	44.800	44.646	0.689
	(0.266)	(0.375)	(0.376)	
Other Demanded Skill (%)	12.476	12.474	12.479	0.983
	(0.177)	(0.249)	(0.250)	
c. Job Tasks:				
Non-Routine Manual (%)	73.724	73.543	73.906	0.283
	(0.235)	(0.333)	(0.333)	
Routine Manual (%)	51.146	51.211	51.081	0.736
	(0.267)	(0.377)	(0.379)	
Other Task (%)	18.035	18.041	18.031	0.972
	(0.206)	(0.290)	(0.291)	
d. Job Title on Ad:				
Cashier (%)	3.587	3.550	3.624	0.603
	(0.099)	(0.140)	(0.142)	
Food Service (%)	6.917	6.798	7.037	0.220
	(0.136)	(0.190)	(0.194)	
Janitor (%)	8.223	8.114	8.333	0.283
	(0.147)	(0.206)	(0.209)	
Team Member (%)	5.948	5.983	5.913	0.694
	(0.126)	(0.179)	(0.179)	
Other (%)	76.820	77.007	76.636	0.244
	(0.226)	(0.318)	(0.320)	
Observations	34,987	17,549	17,437	

Notes: This table lists the averages of firm characteristics across the full sample period. Standard errors are reported in parentheses. The statistics in Panel (a) are based on the job text obtained from the advertisement. Posted hourly wages are estimated using the information about the salary and the number of hours worked. Estimated hourly wages are based on 12,829 observations rather than the full sample because not every job ad includes a wage. The statistics in Panel (b) are predicted skills from the advertisement text using the procedure from Spitz-Oener (2006). These categories are not mutually exclusive. The statistics in Panel (c) are predicted tasks from the advertisement text using the procedure from Atalay et al. (2020). These categories are not mutually exclusive. The statistics in Panel (d) are predicted from the advertisement text using key words. The categories are not mutually exclusive. Column (1) is based on the entire subject pool. Columns (2) and (3) are based on the firms selected to receive resumes of a given type. Column (3) reports the *p*-value of a test of equal means across the resume types. Standard errors are clustered at the establishment level.

Table C2: Treatment Balance across the Whole Sample Period

		Treatment Arm								
	All	B01GED	B01HS	B12GED	B12HS	W01GS	W01GED	W12GED	W12HS	p-value test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
a. Job Hours and Pay:										
Part Time Job (%)	34.831	34.586	35.130	34.516	34.578	34.908	35.142	34.413	35.371	0.947
	(0.255)	(0.718)	(0.720)	(0.717)	(0.720)	(0.718)	(0.729)	(0.723)	(0.720)	
Full Time Job (%)	40.916	40.123	40.791	41.405	41.338	41.949	40.522	41.073	40.127	0.584
	(0.263)	(0.740)	(0.741)	(0.743)	(0.746)	(0.744)	(0.749)	(0.748)	(0.738)	
Job Hours Unstated (%)	38.120	39.143	38.631	37.381	37.764	37.315	38.216	37.882	38.632	0.471
	(0.260)	(0.737)	(0.734)	(0.730)	(0.734)	(0.729)	(0.742)	(0.738)	(0.733)	
Has Posted Wage (%)	36.669	35.543	36.266	36.880	37.580	36.498	37.611	37.049	35.960	0.306
	(0.258)	(0.723)	(0.725)	(0.728)	(0.733)	(0.726)	(0.739)	(0.735)	(0.722)	
Estimated Hourly Wage (Dollars)	13.010	13.150	12.758	13.108	13.179	12.970	13.162	13.008	12.742	0.508
	(0.071)	(0.234)	(0.191)	(0.205)	(0.180)	(0.190)	(0.201)	(0.215)	(0.194)	
b. Job Skills:										
Social Skills Demanded (%)	57.795	57.599	58.140	58.549	57.493	58.074	57.080	58.326	57.088	0.721
	(0.264)	(0.746)	(0.744)	(0.743)	(0.748)	(0.744)	(0.755)	(0.750)	(0.745)	
Customer Skills Demanded (%)	78.443	79.426	77.558	78.081	77.612	78.696	78.854	77.729	79.574	0.041
	(0.220)	(0.610)	(0.629)	(0.624)	(0.631)	(0.617)	(0.623)	(0.633)	(0.607)	
Character Skills Demanded (%)	44.724	45.409	45.339	44.088	44.363	44.333	44.853	44.889	44.520	0.785
	(0.266)	(0.752)	(0.751)	(0.749)	(0.752)	(0.749)	(0.759)	(0.756)	(0.748)	
Other Demanded Skill (%)	12.476	11.643	12.847	12.437	12.970	12.423	12.599	12.673	12.228	0.595
	(0.177)	(0.484)	(0.505)	(0.498)	(0.509)	(0.497)	(0.506)	(0.506)	(0.493)	
c. Job Tasks:										
Non-Routine Manual (%)	73.724	73.935	72.715	73.761	73.763	74.040	73.614	73.474	74.479	0.650
	(0.235)	(0.663)	(0.672)	(0.663)	(0.666)	(0.661)	(0.673)	(0.671)	(0.656)	
Routine Manual (%)	51.146	51.629	50.864	51.205	51.146	50.943	50.908	50.486	51.970	0.894
	(0.267)	(0.754)	(0.754)	(0.754)	(0.757)	(0.753)	(0.763)	(0.760)	(0.752)	
Other Task (%)	18.035	17.635	18.736	18.031	17.759	17.670	18.444	18.571	17.459	0.638
	(0.206)	(0.575)	(0.588)	(0.580)	(0.579)	(0.575)	(0.592)	(0.591)	(0.571)	
d. Job Title on Ad:										
Cashier (%)	3.587	3.828	3.774	3.251	3.346	3.543	3.377	3.446	4.121	0.342
	(0.099)	(0.290)	(0.287)	(0.267)	(0.272)	(0.279)	(0.276)	(0.277)	(0.299)	
Food Service (%)	6.917	6.881	6.708	6.867	6.737	6.768	7.336	7.100	6.952	0.898
	(0.136)	(0.382)	(0.377)	(0.381)	(0.379)	(0.379)	(0.398)	(0.391)	(0.383)	
Janitor (%)	8.223	8.020	7.844	8.231	8.364	8.290	8.570	8.603	7.880	0.745
	(0.147)	(0.410)	(0.405)	(0.414)	(0.419)	(0.416)	(0.427)	(0.426)	(0.405)	
Team Member (%)	5.948	6.266	6.048	6.139	5.477	5.792	5.659	5.897	6.295	0.630
	(0.126)	(0.366)	(0.359)	(0.362)	(0.344)	(0.352)	(0.353)	(0.358)	(0.366)	
Other (%)	76.820	76.715	77.012	76.830	77.475	77.129	76.339	76.526	76.540	0.868
	(0.226)	(0.638)	(0.635)	(0.636)	(0.632)	(0.633)	(0.649)	(0.645)	(0.638)	
Observations	34,987	4,389	4,398	4,398	4,364	4,403	4,294	4,324	4,416	

Notes: This table lists the averages of firm characteristics across the full sample period. Standard errors are reported in parentheses. The statistics in Panel (a) are based on the job text obtained from the advertisement. Posted hourly wages are estimated using the information about the salary and the number of hours worked. Estimated hourly wages are based on 12,829 observations rather than the full sample because not every job ad includes a wage. The statistics in Panel (b) are predicted skills from the advertisement text using the procedure from Spitz-Oener (2006). These categories are not mutually exclusive. The statistics in Panel (c) are predicted tasks from the advertisement text using the procedure from Atalay et al. (2020). These categories are not mutually exclusive. The statistics in Panel (d) are predicted from the advertisement text using key words. The categories are not mutually exclusive. Column (1) is based on the entire subject pool. Columns (2) through (9) are based on the firms selected to receive resumes of a given type. Column (9) reports the p-value of a test of equal means across the resume types. Standard errors are clustered at the establishment level.

D. Do Firms Learn that Resumes are Fake over Time?

One potential concern with our analysis is that our results are primarily driven by firms learning, over time, that the resumes are likely to be fake. Under this interpretation, callbacks would decrease as more firms learn that the resumes are fake and do not expend effort calling back fake applicants. Additionally, this interpretation would require that the resumes of white applicants are more likely to be discovered as fake, closing the callback gap.

We believe that this interpretation is unlikely for several reasons. First, we only send about two resumes to each firm per period. We consider this to be a small number of applications sent to each firm, minimizing the firm's scope to learn that the resumes are fake. Moreover, we took great steps in the design to minimize this possibility. As mentioned in Section A.1, we avoided sending resumes of the same type, with matching work history, with the same phone number, or with the same name to the employer more than once.

Second, this learning is likely to happen gradually over time. However, Figure 5 shows that the change in callbacks happens immediately after the minimum wage hikes are announced and persists for over a year afterward. For this data pattern to be consistent with the mechanism that firms learn about the resumes being fake over time, we would need the sample to learn that the resumes are fake coincidentally with the minimum wage hikes and for all learning to cease after that.

Third, if learning is driving our results, we would expect that firms that receive more resumes from us are more likely to learn that the applications are fake. In our sample, we have variation in whether we applied to firms only before the minimum wage hike, only after, or both before and after (See Appendix J). Figure J2 shows that firms to which we only sent one set of applications behave similarly to those to which we sent more than one set of applications. This pattern of behavior is inconsistent with firms that we send resumes to more than once learning that our resumes are fake.

Fourth, Table D1, below, displays the main effects for the full sample and those we only applied to during one period. In columns (2) and (4), we also control for the number of applications sent to the firm. As would be inconsistent with the mechanism that firms learn the resumes are fake over time, we find a positive correlation between the number of resumes we send to the firm and the likelihood of a callback. Moreover, neither controlling for the number of applications sent to the firm nor restricting the sample to those to which we only sent applications in a single period meaningfully affects our results.

Table D1: Effects by When We Applied to the Establishment

	(1) Callback	(2) Callback	(3) Callback	(4) Callback
Black	-3.210***	-3.178***	-3.159***	-3.133***
	(0.598)	(0.597)	(0.646)	(0.645)
After Announcement	-9.085***	-8.635***	-7.518***	-8.159***
	(0.611)	(0.612)	(0.731)	(0.733)
Black × After Announce	2.575***	2.521***	2.482***	2.443***
	(0.710)	(0.709)	(0.757)	(0.756)
Number of Resumes Sent to Establishment		1.118***		1.333***
		(0.101)		(0.130)
Constant	5.730	2.271	-1.414	-4.007
	(8.896)	(8.886)	(9.307)	(9.294)
Sample	Full	Full	Applied in Only 1 Period	Applied in Only 1 Period
R-squared	0.018	0.021	0.016	0.020
Observations	34,986	34,986	30,708	30,708

Notes: This table reports estimates of the main effects by whether we applied to the establishment in both periods or only a single period. Even columns also control for the number of resumes sent to each establishment. All specifications control for the applicant's age and the firm's city. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05.

Finally, we can examine whether the order of applications affects the results. If sending applications to a firm taught them that the resumes were fake, we would expect the firms to be more likely to realize that the second resume was fake because they did not hear back from the first application or because the resumes are similar in some way.

Table D2 shows that we find no evidence that application order played a role in a firm's callback decision. There is no effect of being the second application sent to the firm overall or by the applicant's race. These results are robust to using resumes for firms that received an opposite race resume (column 1), an opposite unemployment duration resume (column 2), or the full sample. While we cannot know with certainty that the firms viewed the applications in the order in which they were received, this table provides suggestive evidence against the sequential spillover effects found in Kessler et al. (2022).

Table D2: Effects of Resume Order

	(1) Callback	(2) Callback	(3) Callback
Black	-1.710**	-2.072***	-1.884***
	(0.625)	(0.628)	(0.442)
Second Application	-0.107	-0.364	-0.231
	(0.659)	(0.583)	(0.441)
Black × Second Application	0.870	1.190	1.020
	(1.115)	(0.931)	(0.727)
Constant	11.07***	11.13***	11.10***
	(0.469)	(0.469)	(0.337)
R-squared	0.012	0.010	0.011
Observations	17,569	17,417	34,986

Notes: This table displays the probability of being called back for a job interview by race, the order of the application, and the interaction between race and the order of the application. All specifications control for applicant age and firm city fixed effects. Column (1) shows the results for race pairs—that is, firms that received a second resume from the opposite race. Column (2) shows the results for unemployment duration pairs. Column (3) shows the pooled estimates, controlling for the subject's pair type. Standard errors are clustered at the establishment level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

E. Robustness to Measurement of Treatment Timing and Callbacks

In our preferred specifications, we assign each application to the minimum wage policy at the date the application is submitted. Since most applications do not receive any contact from the potential employer, this date is the only point of reference we observe as researchers. However, significant delays between when the application is received and when it is reviewed may lead us to misclassify the treatment of interest.

We test whether potential misclassification affects our results. We assign treatment, or the minimum wage policy, based on the date of the earliest contact we receive for a given job ad, which is not necessarily a callback. For example, if we sent applications to a job ad before the minimum wage was announced, but then one applicant is called back *after* the minimum wage is enacted and the other is never contacted, we code the time period as "Enacted" for both applications. The date is not defined for job ads where no applicants are contacted, and they are therefore excluded.

Column (1) of Table E1 presents the results with this alternative measurement of treatment timing and shows that the main result still holds—minimum wages shrink the racial callback gap. However, the magnitudes are much larger. This is because we are conditioning on receiving some contact from the employer, which consists primarily of callbacks, and callbacks were more often for white applicants. In the full sample, many employers call back neither of our applicants, and so appear to not discriminate in the data. Since this new definition of treatment timing meaningfully changes the sample, we additionally estimate our preferred specification that defines treatment by application date using this restricted sample. The results are again similar.

We now additionally consider the impact of how we measure callbacks. Following our pre-analysis plan and the previous literature, we define callbacks to be explicit interview requests. However, many employers will contact applicants for more information before offering an interview. These requests for additional information *may* signal that the employer wants

Table E1: Alternative Measurement of Treatment Timing

	Callback	
	(1)	(2)
Treatment Assignment By:	Callback Date	Application Date
Black	-8.33***	-7.82***
	(1.57)	(1.61)
Announced	-7.75**	-8.69***
	(2.85)	(2.59)
Enacted	-9.96***	-13.48***
	(2.11)	(2.05)
Black × Announced	7.30*	4.40
	(2.92)	(2.94)
Black × Enacted	4.29*	4.05+
	(2.07)	(2.09)
N	8,700	8,700
R^2	0.02	0.02

Notes: Both specifications include city fixed effects and control for applicant age. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05, + p-value < 0.10.

to hire the applicant, but others appear to be automatically generated after submission of the application and therefore do not allow us to measure potential discrimination. Over our entire sample period, about 18 percent of applicants receive any response, while only 10 percent receive an interview request. We estimate Equations 4 and 6 using both definitions for callbacks and present the results in Table E2. The results on average discrimination, in columns (1) and (2), and the results on how they change with the minimum wage, in columns (3) and (4), are very similar across definitions.

Table E2: Alternative Measurement of Callbacks

	(1) Interview Request	(2) Any Response	(3) Interview Request	(4) Any Response
Black	-1.39***	-1.30***	-3.21***	-3.44***
	(0.27)	(0.34)	(0.61)	(0.71)
Unemp 12 mo.	-0.44	-0.34	-0.47	-0.36
	(0.36)	(0.45)	(0.36)	(0.45)
GED	-0.46	-1.22**	-0.50	-1.26**
	(0.32)	(0.40)	(0.31)	(0.40)
Announced			-5.17***	-4.24***
			(1.02)	(1.22)
Enacted			-10.1***	-8.43***
			(0.84)	(0.98)
Black × Announced			2.85**	2.97*
			(1.00)	(1.20)
Black × Enacted			2.52***	3.04***
			(0.69)	(0.82)
Observations	34986	34986	34986	34986
R^2	0.01	0.01	0.02	0.01

Notes: All specifications include city fixed effects and control for applicant age. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05, $^{+}$ p-value < 0.10.

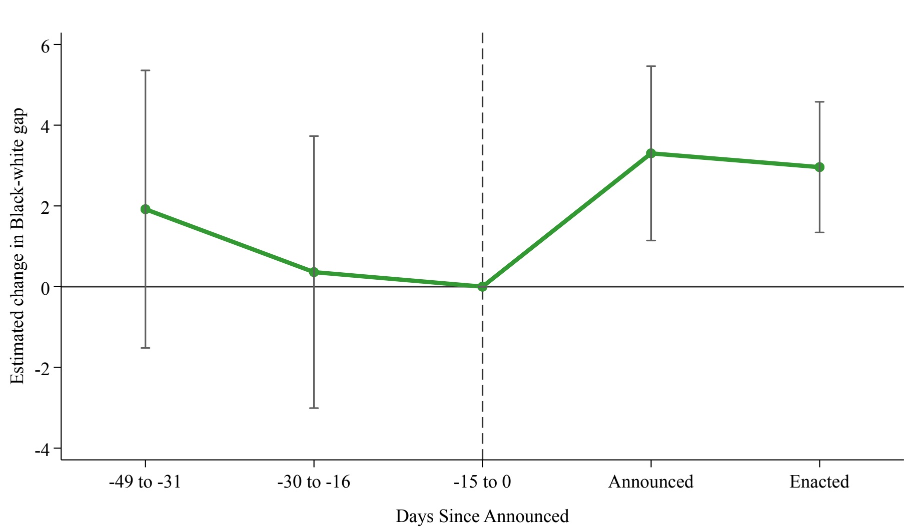
F. Evidence of Common Trends

In this section, we test the common trends assumption by testing for pretrends in the racial callback gap. We first start with binning the time periods before a minimum wage increase was announced. Figure F1 displays the treatment-effect estimates relative to fewer than 15 days prior to the announcement in event time:

$$Y_{ict} = \sum_{E} \left(\beta_E E_{ct} + \delta_E \text{Black}_i \times E_{ct} \right) + \alpha \text{Black}_i + X_i' \gamma + \eta_c + \epsilon_{ict} \tag{11}$$

where E_{ct} is an indicator for whether calendar time t in city c corresponds to event time bin, E.

Figure F1: The Minimum Wage and the Racial Callback Gap Dynamics



Notes: The figure presents estimates and confidence intervals for the change in the racial callback gap by days since the beginning of the experiment—the δ coefficients from Eq 11. The specification includes city fixed effects and controls for GED, duration unemployed, and age. Standard errors are clustered by establishment. The specification includes the full sample of 34,986 observations, but we only present estimates for time bins where we have data from all three states with minimum wage increases. This excludes our estimate for -50 or more days since announced, which we only observe for Illinois.

This figure is analogous to our main results from Equation 6, but with the pre-period split into multiple periods. We find that there is no evidence of differential time trends prior to the minimum wage hike announcements. If anything, discrimination against minority applicants is increasing slightly in the seven weeks prior to the minimum wage hike announcements (since a positive change in the racial callback gap corresponds to a reduction). This suggests that the effect of the minimum wage was likely not driven by a preexisting time trend.

Second, given our relatively short pre-period for the minimum wage increase announcements, we additionally consider another version of Equation 6 that allows the callback rate to linearly change, separately for white and Black applicants, in the seven weeks leading up to the announcement. Specifically:

$$\begin{aligned} Y_{ict} =& \beta_0 \text{Pre x Event Time}_{ct} + \delta_0 \text{Black}_i \times \text{Pre x Event Time}_{ct} \\ &+ \beta_1 \text{Announced}_{ct} + \delta_1 \text{Black}_i \times \text{Announced}_{ct} + \beta_2 \text{Enacted}_{ct} \\ &+ \delta_2 \text{Black}_i \times \text{Enacted}_{ct} + \alpha \text{Black}_i + X_i' \gamma + \eta_c + \epsilon_{ict} \end{aligned} \tag{12}$$

where Event Time_{ct} takes on negative values in pre-announce days. If there are pretrends, we would expect δ_0 to be different from 0, indicating that the callback gap is growing or shrinking as we move closer to the announcement date and further from the start of our data collection. This strategy uses the fact that, because of the nature of our experiment, we have 57 different dates during the pre-announce period in which we apply to at least one job, but with relatively few observations at each specific date. It therefore complements our previous test for pretrends where we, instead, bin pre-announcement dates into three bins.

We present the results in Table F1 and find that the callback rate remains stable for both white and Black applicants during the pre-period. If anything, as in Figure F1, the racial callback gap is increasing, since δ_0 is negative, although it is not different from 0.

Table F1: Testing for Pretrends

	(1)
Announced	-6.17***
	(1.62)
Enacted	-11.1***
	(1.46)
Black	-3.64***
	(0.74)
Pre-Announce × Event Time	0.040
	(0.039)
Black × Pre-Announce × Event Time	-0.032
	(0.027)
Black × Announced	3.29**
	(1.09)
Black × Enacted	2.95***
	(0.80)
Observations	34990
R^2	0.020

Notes: All specifications include city fixed effects and controls for GED, duration unemployed, and age. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05, + p-value < 0.10

G. Alternative Assumptions on the Productivity Distributions by Race

In the model in Section 3 and our empirical application in Section 5, we focus on the case where employers perceive that applicant productivity follows a normal distribution. The normal distribution assumption is a useful case because it allows us to consider the role of three types of discrimination that are often discussed in the literature—taste-based, statistical in means, and statistical in variance. However, it is a particular functional-form assumption that may not hold. The equation for the change in the RCG with respect to the minimum wage also implies that other distributions with similar densities, like the logistic distribution, will lead to similar qualitative conclusions.

We now assume that employers' perception of applicant productivity is uniformly distributed. Specifically, fixing $X_i = x$, $q_i(0, x) \sim U(a_W, b_W)$ and $q_i(1, x) \sim U(a_B, b_B)$. Therefore, the likelihood of receiving a callback may differ for white and Black applicants because employers perceive the supports of the two distributions to differ in addition to a discriminatory penalty, d, for Black applicants. Under this assumption,

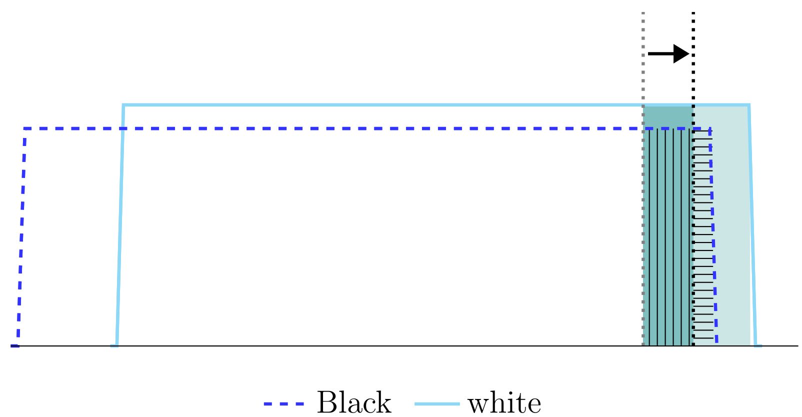
$$RCG = \frac{b_W - v(m)}{b_W - a_W} - \frac{b_B - (v(m) - d)}{b_B - a_B} \tag{13}$$

In this case, the RCG may be positive because employers believe that there are more high-productivity white applicants—for example, because b_W is larger than b_B, or because of the discriminatory penalty. Based on this expression, the change in the RCG from a minimum wage increase is

$$\begin{aligned} \frac{\partial RCG}{\partial m} &= v'(m) \times \left[\frac{1}{b_W - a_W} - \frac{1}{b_B - a_B} \right] & (14) \\ &= v'(m) \times \sqrt{12} \left[\frac{1}{\sigma_W} - \frac{1}{\sigma_B} \right] & (15) \end{aligned}$$

where σ is the standard deviation of $q|B$. While taste-based and statistical discrimination both determine the RCG, only statistical discrimination determines how the gap changes with a minimum wage increase. Using our estimates from the correspondence study, we plot the implied perceived productivity distributions for Black and white applicants in Figure G1 to again show how differences in the perceived variance, or the length and height of the pdf, lead white applicants to be more negatively affected by minimum wage increases. Since the length of the pdf for white applicants is shorter, and therefore the height is larger, a relatively larger share of the marginal applicants are white.

Figure G1: Minimum Wages and Callbacks when Productivity Is Uniformly Distributed



H. Effects of Minimum Wages on Education and Duration Unemployed Gaps

We extend Equation 6 to consider whether minimum wage increases affect the callback gaps between GED and high school graduates as well as 12-month and 1-month unemployed applicants. We find no evidence that the minimum wage increases meaningfully affect the callback gaps by education and duration unemployed. Including these additional terms does not affect our estimates for the change in the racial callback gap.

Table H1: The Minimum Wage and Callback Gaps

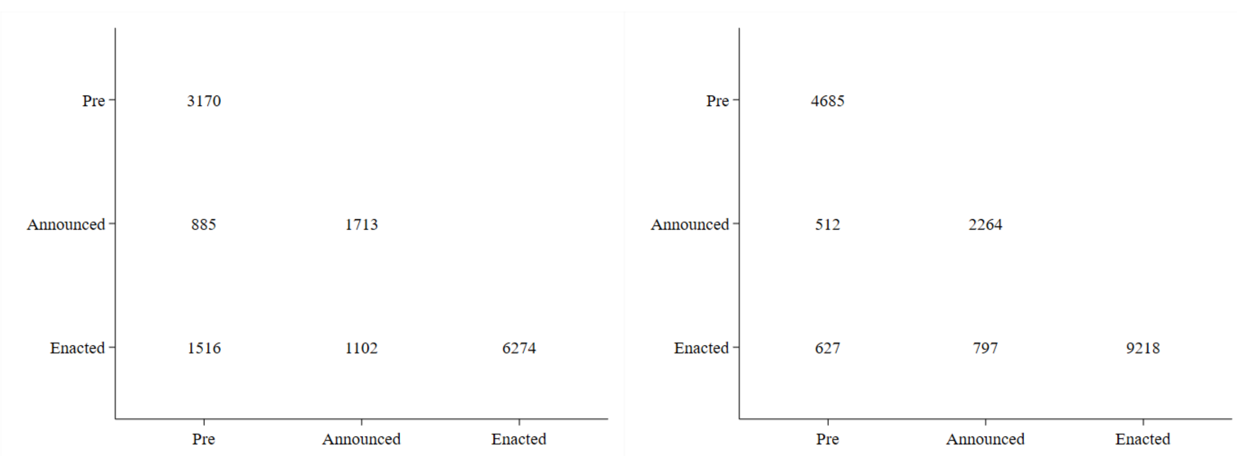
	Callback	
Black	-3.22***	-2.95***
	(0.61)	(0.65)
GED	-0.74	-0.87
	(0.71)	(0.79)
Unemployed 12 months	-1.04	-1.29^{+}
	(0.67)	(0.74)
After Announcement	-9.68***	-7.14***
	(0.98)	(1.17)
Black × After Announce	2.58***	2.11**
	(0.68)	(0.73)
GED × After Announce	0.36	0.46
	(0.79)	(0.89)
Unemployed 12 months × After Announce	0.82	0.59
	(0.69)	(0.76)
Firm Fixed Effects	No	Yes
N	34,990	34,990
R^2	0.02	0.36

Notes: Both specifications include city fixed effects and control for applicant age. Standard errors are clustered by establishment. *** p-value < 0.001, ** p-value < 0.01, * p-value < 0.05, $^{+}$ p-value < 0.10.

I. Robustness of Firm Composition Results across Alternative Specifications

We consider alternative specifications to estimate the change in the racial callback gap after a minimum wage increase using within-firm and within-establishment variation. We first present the number of firms and establishments posting across periods in Panels (a) and (b) of Figure I1, respectively. The diagonal elements give the number of unique firms or establishments in our sample in a period. For example, we applied to 1,713 unique firms between the announcement and enactment of the increase. Off-diagonal elements give the number of unique firms or establishments that we applied to across two periods. For example, there are 1,516 unique firms to which we applied to a posting both before the minimum wage was announced and after it was enacted. These 1,516 are a subset of the 6,274 unique firms in the Enacted period and the 3,170 in the pre-period. The matrix is symmetric, and so we only present the lower triangle. Using the variation shown in this figure, we estimate specifications similar to Equation 6 that include firm and establishment fixed effects and present the results in Figure I3.

Figure I1: Firm and Establishment Postings Over Time

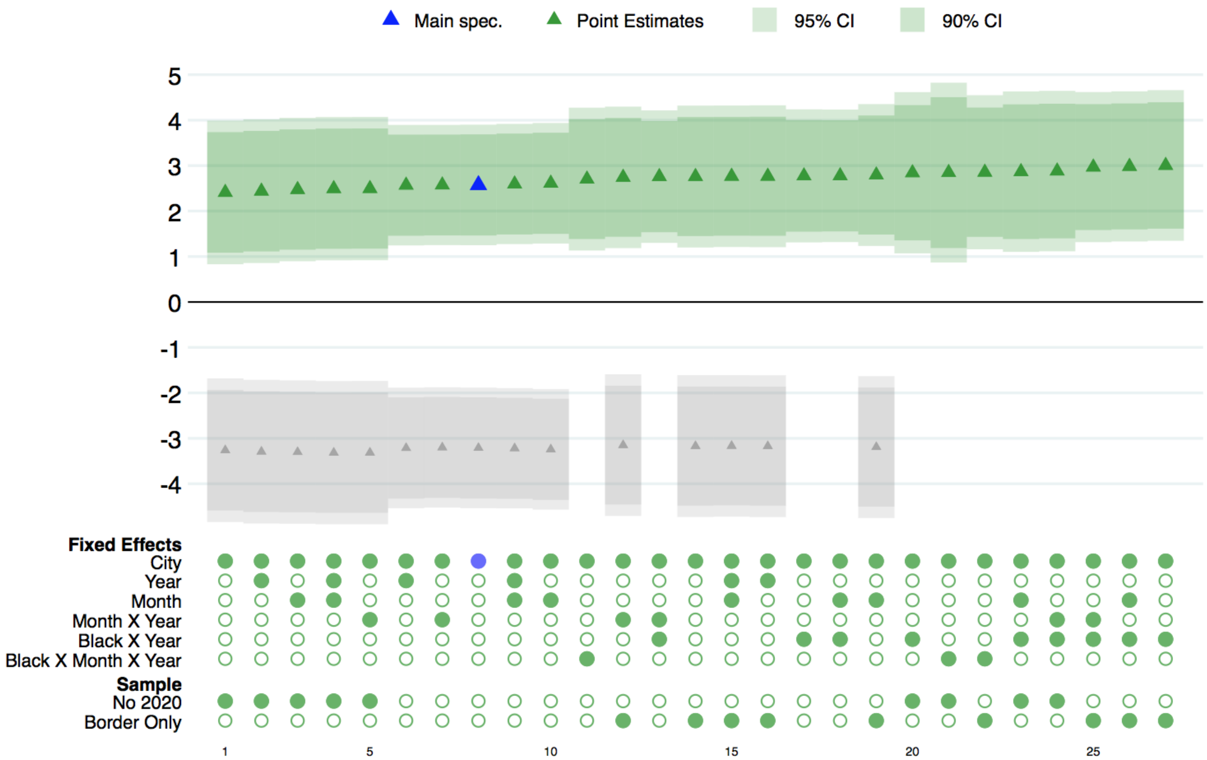


(a) Firms

(b) Establishments

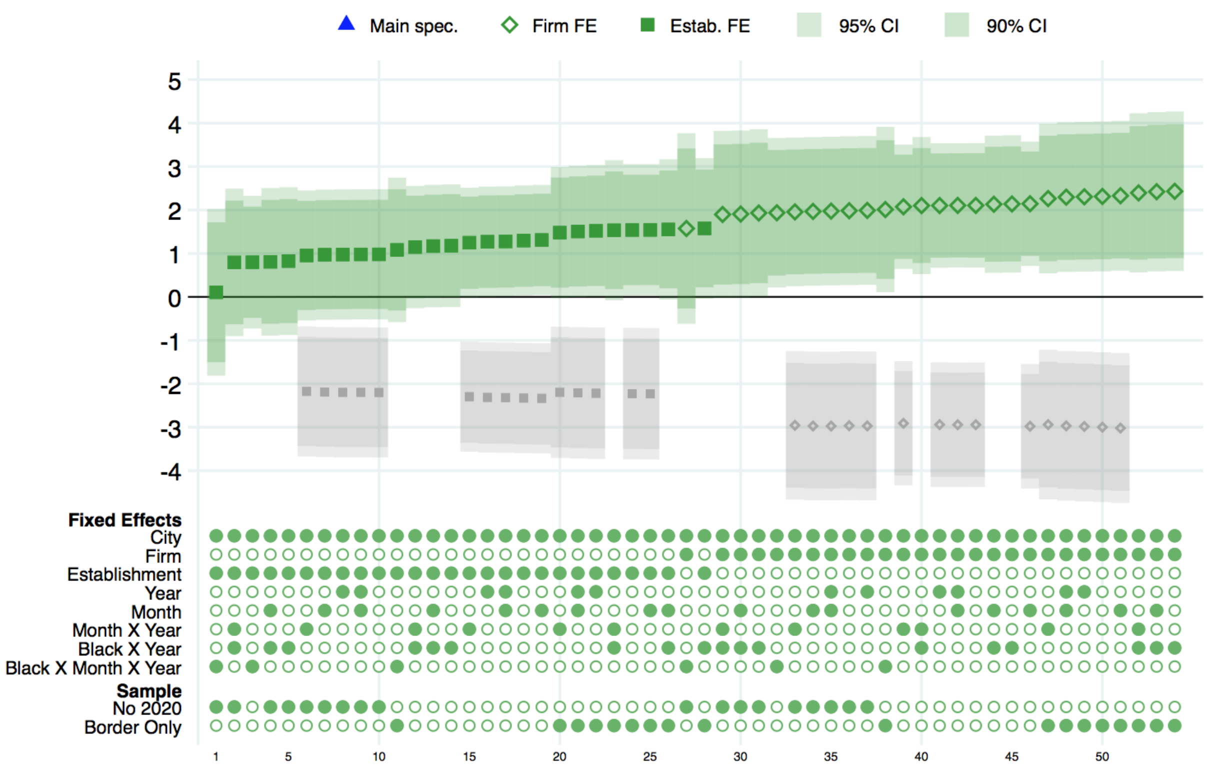
Notes: Elements of Panel (a) show variation across time in postings by firms in our sample. Diagonal elements give the number of unique firms in a given period. Off-diagonal elements give the number of unique firms that appear in both periods. Panel (b) similarly presents the number of unique establishments posting across periods. The matrix is symmetric, and so we present the lower triangle.

Figure I2: Alternative Specifications



Notes: This figure presents a specification curve for the baseline racial callback gap in gray and the change in the racial callback gap from a minimum wage increase. The preferred specification is highlighted in blue. The filled-in circles indicate which controls are included. The baseline gap is not identified in specifications that include Black × time-period fixed effects. In all specifications, standard errors are clustered by establishment.

Figure I3: Specification Curve for the Role of Firm Composition



Notes: Each specification considers an alternative set of fixed effects or limits the sample to test the robustness of the effect of minimum wage increases on the racial callback gap. All specifications include either firm or establishment fixed effects as well as applicant age and city fixed effects. Standard errors are clustered by establishment.

J. Additional Results on Firm Compositional Changes

We provide an additional test of the extent to which changes in firm or establishment composition drive our main result: that minimum wage increases shrink the racial callback gap. To do this, we divide our establishments into four mutually exclusive groups by when the establishment was hiring and when we applied: 1) those who only hire before the minimum wage hike is announced, 2) those who only hire after the minimum wage hike is announced, 3) the pre-period for those who hire both before and after the announcement, and 4) the post-announcement-period for those who hire both before and after the announcement. Whether and when an establishment chooses to hire is endogenous, but splitting our sample in this way helps illustrate both changes in composition and policy responses.

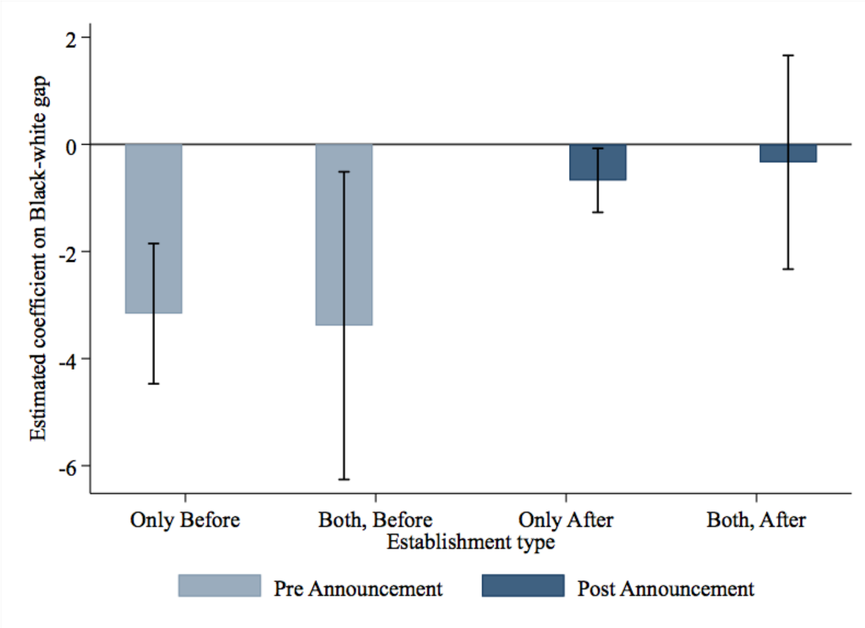
If establishment composition solely contributes to the shrinking of the racial gap, then we would expect no differences in the behavior of establishments hiring in both periods. We would expect there only to be a difference between establishments that only hired either before or after the minimum wage hike. Instead, if establishments are only changing their behavior because of the policy, then we should see similar patterns between establishments hiring in both periods and those hiring only before or after the hikes. We estimate that

$$Y_{ict} = \sum_{e} \omega_e \mathbb{1}_e + \sum_{e} \delta_e \mathbb{1}_e B_i + X_i' \gamma + \eta_c + \epsilon_{ict} \tag{16}$$

where $\mathbb{1}_e$ is an indicator representing whether the establishment is of type e. The main parameters of interest are the four δs, which correspond to the racial callback gaps by establishment type and period. We plot the δs and their confidence intervals in Figure J2.

The effects of minimum wage increases by establishment type suggest that establishment composition matters relatively little. The callback gap shrinks by about 3.0 percentage points among establishments that hire both before and after the announcement ($p < 0.095$). We additionally find a 2.5 percentage point difference between establishments that only hire after and those that only hire before ($p = 0.001$), which is not statistically different from the

Figure J2: Racial Callback Gap by When Establishment Hires



Notes: The figure presents estimates and confidence intervals for the change in the racial callback gap from the minimum wage hike across our three establishment types: 1) only hires before the new minimum wage is announced, 2) only hires after the new minimum wage is announced, and 3) hires both before and after. We split type 3) into the before- and after- announcement periods. The specification control is for city fixed effects and other characteristics. Standard errors are clustered by establishment.

3.0 percentage point change among establishments that hire in both periods. Instead, our results imply that hiring managers change their behavior in response to the higher minimum wage, consistent with the discussion from Section 5. They use a different threshold for deciding whom to call back, which mostly crowds out white applicants.

The results on establishment composition are based on our sample of job postings on Indeed.com. If the minimum wage hike changed the platform some establishments use to search for applicants, we would have missed those jobs, creating selection into the sample. While we tried to apply to every relevant job on Indeed.com, our resources were somewhat limited, and we could not apply to all postings.

K. Multiple Hypothesis Testing

In this section, we will establish that the results in the paper are not false positives resulting from not adjusting the significance value of our hypothesis tests in a manner that reflects the multiple comparisons (List et al., 2019, 2021). We do this by testing families of null hypotheses from the body of this paper using the procedure from Westfall and Young (1993). This procedure controls for the family-wise error rate and allows for dependence amongst p-values within a hypothesis.[41] We control for the family-wise error rate within families, but not across families.

We follow Rubin (2021) and correct for multiple comparisons when performing disjunction testing, but not when performing conjunction or individual testing.[42] Tables K1 and K2, below, display reproductions of the coefficient estimates and p-values from the main body of the paper, along with the Westfall and Young (1993) adjusted p-values. For each hypothesis test, we include the exhibit reference and our family designation. Hypotheses not included in this table were determined to be either conjunction or independent tests.

Table K1 shows that none of the conclusions from these analyses change when correcting for multiple hypothesis testing. Family 3 conducts multiple hypothesis testing on the hypotheses from Table H1, which are a superset of those presented in Figure 4 and help us rule out the race effects as the result of a multiple hypothesis problem involving testing the change in three different characteristics with the minimum wage. While the main results of the paper are robust to this correction, there are a few instances where we lose enough power from using the family-wise approach that we can no longer detect differences from zero in Family 2 (see Figure 2b). We can no longer detect differences between the highest-quality

[41] This procedure is similar to the Romano and Wolf (2016), List et al. (2019) and List et al. (2021) procedures. However, it has an additional assumption of set pivotality, which is not assumed by the the three proceeding procedures. We use this procedure because it can account for clustered standard errors and fixed effects, which are key to our design.

[42] Rubin (2021) refers to disjunction testing as tests that require *at least one* hypothesis to return a significant result before rejecting a joint null hypothesis. The *conjunction testing* approach requires all statistical tests to be significant as a decision rule. The *individual testing* approach is one where the researchers only make decisions about each constituent null hypothesis separately.

Black resume and the highest-quality white resume (one-month unemployment duration and high school education). We can also no longer detect differences between resumes from a white applicant with a 12-month unemployment duration and a GED and the highest-quality white resume. Despite these differences, the conclusion from this family that the treatment effects found in Figure 2a are driven by race does not change. Table K2 also shows that none of the conclusions from Figures 8a or 8b are changed when using a family-wise error approach.

Table K1: P-value Corrections for Multiple Hypothesis Testing on Disjunctive Hypotheses: Part I

	Exhibit	Family	Outcome	Variable	Coeff.	p-values	
						Unadjusted	Westfall and Young (1993)
(1)	Figure 2a	1	Callback	Black	-1.39	< 0.001	< 0.001
(2)	Figure 2a	1	Callback	GED	-0.44	0.229	0.278
(3)	Figure 2a	1	Callback	Unemp 12 mo.	-0.46	0.149	0.278
(4)	Figure 2b	2	Callback	B01GED	-2.06	< 0.001	0.002
(5)	Figure 2b	2	Callback	B01HS	-1.37	0.031	0.105
(6)	Figure 2b	2	Callback	B12GED	-1.90	0.004	0.012
(7)	Figure 2b	2	Callback	B12HS	-2.07	0.001	0.004
(8)	Figure 2b	2	Callback	W01GED	-0.29	0.659	0.883
(9)	Figure 2b	2	Callback	W12GED	-1.27	0.062	0.170
(10)	Figure 2b	2	Callback	W12HS	-0.27	0.695	0.883
(11)	Figure 4 (Table H1)	3	Callback	Black	-3.22	< 0.001	< 0.001
(12)	Figure 4 (Table H1)	3	Callback	GED	-0.74	0.298	0.462
(13)	Figure 4 (Table H1)	3	Callback	Unemp 12 mo.	-1.04	0.119	0.309
(14)	Figure 4 (Table H1)	3	Callback	Post	-9.68	< 0.001	< 0.001
(15)	Figure 4 (Table H1)	3	Callback	Black × Post	2.58	< 0.001	0.001
(16)	Figure 4 (Table H1)	3	Callback	GED × Post	0.36	0.647	0.645
(17)	Figure 4 (Table H1)	3	Callback	Unemp 12 mo. × Post	0.82	0.238	0.462
(18)	Table 3, Column (2)	4	Callback	Black	-2.94	< 0.001	< 0.001
(19)	Table 3, Column (2)	4	Callback	Post	-8.50	< 0.001	< 0.001
(20)	Table 3, Column (2)	4	Callback	Black × Post	1.30	0.045	0.034
(21)	Table 3, Column (3)	5	Callback	Black	-2.33	< 0.001	< 0.001
(22)	Table 3, Column (3)	5	Callback	Post	-8.50	< 0.001	< 0.001
(23)	Table 3, Column (3)	5	Callback	Black × Post	1.30	0.045	0.034

Notes: This table displays the coefficients, p-values, and Westfall and Young (1993) adjusted p-values for different analyses done in the main body of the paper. "Family" refers to the set of hypotheses for which the family-wise error rate is controlled. The "Variable" column refers to the variable for which we estimate the coefficient and conduct the hypothesis test. "Post" refers to the indicator equal to 1 if the period is after a minimum wage hike announcement. All standard errors are clustered at the establishment level, and all Westfall and Young (1993) adjusted hypothesis tests use 999 bootstrap iterations.

Table K2: P-value Corrections for Multiple Hypothesis Testing on Disjunctive Hypotheses: Part II

	Exhibit	**Family**	**Outcome**	**Variable**	**Coeff.**	**p-values**	
						Unadjusted	Westfall and Young (1993)
(24)	Figure 8a	6	Callback	B01HS $\times$ Post	1.47	0.348	0.159
(25)	Figure 8a	6	Callback	B12HS $\times$ Post	4.79	0.002	< 0.001
(26)	Figure 8a	6	Callback	B01GED $\times$ Post	1.91	0.215	0.141
(27)	Figure 8a	6	Callback	B12GED $\times$ $\times$ Post	2.24	0.136	0.087
(28)	Figure 8b	7	Callback	Black $\times$ No Relevant Exp. *times* Post	2.31	0.005	< 0.001
(29)	Figure 8b	7	Callback	Black $\times$ Relevant Exp. *times* Post	3.23	0.012	0.001

Notes: This table displays the coefficients, p-values, and Westfall and Young (1993) adjusted p-values for different analyses done in the main body of the paper. "Family" refers to the set of hypotheses for which the family-wise error rate is controlled. The "Variable" column refers to the variable for which we estimate the coefficient and conduct the hypothesis test. "Post" refers to the indicator equal to 1 if the period is after a minimum wage hike announcement. All standard errors are clustered at the establishment level, and all Westfall and Young (1993) adjusted hypothesis tests use 999 bootstrap iterations.

L. Prespecified Analysis and Deviations from Pre-analysis Plan

We submitted our pre-analysis plan on September 18, 2018, when we started collecting data for the experiment. The pre-analysis plan, titled "An Audit Study on Minimum Wage Legislation," can be found at the AEA Registry under ID No. AEARCTR-0003333. We amended the pre-analysis plan twice—once to add the Illinois and Kansas sample when both Missouri's and Arkansas's minimum wage referenda passed. And the second time was to prolong the sample period because we had not yet reached our desired sample size. As with most studies in economics, we deviated from our pre-analysis plan in a few ways.[43]

The original pre-analysis plan specified our main results, presented in Figure 2a and a version of Equation 6 that had no enactment effects and allowed the unemployment duration effect to change with the announcement. Although we deviated from the pre-analysis plan, these deviations do not affect our results or conclusions. Figure I2 shows that our results are robust to changing the sample period. Figure 5 shows that the effects are concentrated around the announcement period, so allowing for enactment effects does not change our results. Similarly, Table H1 shows that allowing the effect of the unemployment duration to change with the minimum wage announcement does not affect the results.

The paper's results are identical to those from the pre-analysis plan, but the expanded sample has precise estimates. Finally, the analysis of the mechanisms and robustness checks in the appendix were not preregistered. We include these anlayses in the paper to help establish the credibility of the main results and to help us understand how to interpret the effects.

[43]See Abrams et al. (2020) for a discussion of posting and adhering to pre-analysis plans in economics.

M. Ethics Appendix

In this section, we describe the ethical considerations of the experiment. We first note that we underwent ethical review at the Human Subjects Committee at the University of Chicago (IRB18-0873), which played an important role in ensuring that the experiment upheld high ethical standards despite our decision not to obtain consent from subjects (List, 2009). Next, we follow the framework of Asiedu et al. (2021).

1. **Equipoise.** In our experiment, each of our resumes is similar to the others by design. Therefore, we do not expect that any treatment arm clearly dominates another treatment arm from the perspective of the employer. The resumes that are least likely to elicit a response may be better from the hiring manager's view because she will not spend time calling the fictitious applicant for an interview. However, we believe that this benefit is small. Moreover, past research on correspondence studies suggests meaningful uncertainty about the relative likelihood of callbacks from each treatment arm (Neumark and Rich, 2019). That being said, the subjects in our experiment would be better off in the status quo world of no resumes. However, learning about discrimination due to taste-based or statistical motivations is not otherwise feasible. We believe, and the IRB agreed, that the benefits from the knowledge outweigh the small costs to the employer.

2. **Role of the Researchers with Respect to Implementation.** The researchers had direct decision-making power over whether and how to implement the experiment. We did not disclose the experiment to the participants before they received a resume.

3. **Potential Harms to Research Participants from the Interventions.** The experimental design potentially harms our subjects. Employers' time is scarce, and we are having them spend it reviewing applications that are fictitious without obtaining the involved parties' consent or compensating employers for their time. Moreover,

Bertrand and Duflo (2017) note that when an applicant declines an offer, employers may learn that applicants with similar attributes are unlikely to accept offers. They claim that this may lead to employers' being less likely to offer jobs to candidates who share those attributes in the future. They also note that after receiving rejections from candidates, the employers may believe that the market is tighter than previously expected, which would be beneficial for real candidates, but detrimental for employers.

While our experiment potentially harms our subjects, identifying racial discrimination from taste-based or statistical motives requires randomly assigning race, which is only available in a field experiment with fictitious applicants. Both of these mechanisms ended up being important drivers of the treatment effects in our study.[44] So we believe that the benefits to society outweigh the harms to the subject. Moreover, obtaining consent for the experiment would likely bias our estimates against finding discrimination, as employers might change their behavior if they knew that they were participating in a study (Levitt and List, 2007).

Moreover, we designed our experiment to limit the potential harm to participants. To the best of our ability, we sent establishments no more than two resumes in each wave. This helped us minimize the burden on firms. Kessler et al. (2019) estimate that it takes employers about 30 minutes to review 10 resumes. This finding suggests that managers spent about six minutes on our experiment in each wave. Second, because few resumes are sent to each firm, we cannot determine whether an individual firm is discriminating. Finally, we never record the identities of the hiring managers at the firms.

4. **Potential Harms to Research Participants from Data Collection or Research Protocols.** We do not believe subjects experience any harm from data collection. Calling back a subject is no different from the firm's actions in everyday life. The

[44] Kessler et al. (2019) have a method of detecting implicit discrimination against subjects without using deception. However, this method cannot measure discrimination from statistical or taste-based motives.

firm's responses were anonymized so that no individual could link a particular firm's callback decisions to the hiring manager's identification.

5. **Financial and Reputational Conflicts of Interest.** Neither Brandon, Holz, Simon, Uchida, nor any of the research assistants received any form of financial compensation as part of the study. The research questions pursued in this study are novel and different from prior work conducted by the principal investigators (PIs). We perceive no reputational conflicts of interest.

6. **Intellectual Freedom.** This study was conducted without collaborating with organizations. The study was conceived and designed by the PIs, who maintained intellectual freedom throughout all stages of the project. At no point did an outside partner have undue influence on the analysis or the interpretation of the results.

7. **Feedback to Participants and Communities.** We intend to share our results with policymakers after our work is subjected to peer review.

8. **Foreseeable Misuse of Research Results.** We recognize that the results are relevant for public policy in labor markets. We advise policymakers to acknowledge that only one of several potential relevant outcomes is studied in our setting. While our paper concludes that increases in the minimum wage reduce racial disparities in hiring decisions, we cannot speak about racial disparities in other labor market outcomes that may be relevant for optimal policy. Future research should investigate whether the minimum wage reduces turnover, separations, and workplace amenities, and policymakers should acknowledge the uncertainty around the effects on these outcomes when making policy decisions.

References

Abrams, E., J. Libgober, and J. A. List (2020). Research registries: Facts, myths, and possible improvements. Technical report, National Bureau of Economic Research.

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

Asiedu, E., D. Karlan, M. Lambon-Quayefio, and C. Udry (2021). A call for structured ethics appendices in social science papers. Proceedings of the National Academy of Sciences 118(29), e2024570118.

Atalay, E., P. Phongthiengtham, S. Sotelo, and D. Tannenbaum (2020). The evolution of work in the united states. American Economic Journal: Applied Economics 12(2), 1–34.

Bertrand, M. and E. Duflo (2017). Field experiments on discrimination. Handbook of economic field experiments 1, 309–393.

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

Cook, L. D., T. D. Logan, and J. M. Parman (2014). Distinctively black names in the american past. Explorations in Economic History 53, 64–82.

Cook, L. D., T. D. Logan, and J. M. Parman (2016). The mortality consequences of distinctively black names. Explorations in Economic History 59, 114–125.

Fryer, R. G. and S. D. Levitt (2004). The causes and consequences of distinctively black names. The Quarterly Journal of Economics 119(3), 767–805.

Gaddis, S. M. (2015). Discrimination in the credential society: An audit study of race and college selectivity in the labor market. Social Forces 93(4), 1451–1479.

Kaplan, J. (2021). predictrace: Predict the Race and Gender of a Given Name Using Census and Social Security Administration Data. R package version 2.0.0.

Kessler, J. B., C. Low, and X. Shan (2022). Lowering the playing field: Discrimination through sequential spillover effects.

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

Kline, P., E. K. Rose, and C. R. Walters (2022). Systemic discrimination among large us employers. Technical Report 4.

Kreisman, D. and J. Smith (2022). Distinctively black names and educational outcomes.

Kroft, K., F. Lange, and M. J. Notowidigdo (2013). Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*. The Quarterly Journal of Economics 128(3), 1123–1167.

Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? Journal of Economic perspectives 21(2), 153–174.

List, J. A. (2009). The irb is key in field experiments. Science 323(5915), 713–714.

List, J. A., A. M. Shaikh, A. Vayalinkal, et al. (2021). Multiple testing with covariate adjustment in experimental economics. Technical report, The Field Experiments Website.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple hypothesis testing in experimental economics. Experimental Economics 22(4), 773–793.

Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles (2022). Ipums national historical geographic information system: Version 16.0. American Economic Review.

Neumark, D. and J. Rich (2019). Do field experiments on labor and housing markets overstate discrimination? a re-examination of the evidence. ILR Review 72(1), 223–252.

Phillips, D. C. (2020). Do low-wage employers discriminate against applicants with long commutes? evidence from a correspondence experiment. Journal of Human Resources 55(3), 864–901.

Romano, J. P. and M. Wolf (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. Statistics & Probability Letters 113, 38–40.

Rubin, M. (2021). When to adjust alpha during multiple testing: A consideration of disjunction, conjunction, and individual testing. Synthese 199(3), 10969–11000.

Spitz-Oener, A. (2006). Technical change, job tasks, and rising educational demands: Looking outside the wage structure. Journal of labor economics 24(2), 235–270.

Westfall, P. H. and S. S. Young (1993). Resampling-based multiple testing: Examples and methods for p-value adjustment, Volume 279. John Wiley & Sons.