Do Employers' Neighborhoods Predict Racial Discrimination?

Amanda Agan and Sonja Starr[1]
September 23, 2020

ABSTRACT

Using evidence from a large field experiment, we show that the racial composition of employer neighborhoods predicts racially discriminatory employment decisions, in a direction suggesting in-group bias. Building on prior work on Ban-the-Box laws, we also show that employers in less Black neighborhoods appear much likelier to stereotype Black applicants as potentially criminal when they lack criminal record information. Finally, our data show racial disparities in the geographic distribution of job postings. We show that when jobs are located far from Black neighborhoods, Black applicants are doubly disadvantaged: discrimination patterns disfavor them, and they have fewer nearby potential opportunities.

INTRODUCTION

Numerous studies have provided evidence of persistent employment discrimination favoring white applicants (see Quillian et al. 2017, for a meta-analysis of field experiments). This paper focuses on geographic heterogeneity in that pattern. To identify discrimination, we use data from a large field experiment, in which we sent over 15,000 fictitious job applications in Black-white pairs to businesses throughout New Jersey and New York City. Our paper makes three key contributions. First, we show that neighborhood demographics predict racial discrimination in employer callbacks: employers in neighborhoods with whiter and less Black populations discriminate much more heavily in favor of white applicants. Second, we build on our prior work showing that when employers lose access to

---

[1] Rutgers University and University of Chicago, respectively. Thanks to research assistants Linfeng Li, Monica Mogollon, Humberto Martinez Beltran, Jason Marquez, Keerthana Nunna, Yongbo Sim, and workshop participants at the University of Chicago, the University of Michigan, and Wayne State University. We also thank the many research assistants who helped to carry out the underlying experiment, as well as all those who offered feedback on that work (see Agan and Starr 2018, for detailed acknowledgments). The experiment was generously funded by the Princeton University Industrial Relations Section, the University of Michigan Empirical Legal Studies Center, and the University of Michigan Office of Research.

criminal records, they discriminate more based on race (Agan and Starr 2018); we now show that that effect is driven by employers in neighborhoods with low Black populations, who appear especially prone to stereotype Black applicants as criminal. Third, we provide evidence of racial disparities in where jobs are located, and show how these disparities combine with local variation in discrimination patterns to shape race gaps in employment opportunities.

The evidence that neighborhood racial composition predicts employment discrimination is strong and robust, persisting when we control for other neighborhood and employer characteristics. In our simplest specification, we estimate a 2.6 percentage-point Black advantage in entirely Black neighborhoods and a 3.3 percentage-point white advantage in entirely non-Black neighborhoods—large differences relative to baseline callback rates. Because job openings were mainly in non-Black neighborhoods, white applicants have a clear net advantage.

Although there are some good reasons to expect a pattern like this, it is not predicted by every theory of discrimination, so our results can help inform understanding of discrimination mechanisms. Our findings are best explained by some form of stereotyping or other in-group preference. This could entail managers indulging their own biases or catering to those of customers or staff. In contrast, as we discuss below, strong neighborhood effects should *not* be expected based on theories of statistical discrimination.

No prior field study has provided convincing evidence of the patterns we find. Some prior observational research suggests that hiring patterns favor those who match the racial composition of communities or of management (Holzer and Ihlanfeldt 1998; Stoll, Raphael, and Holzer 2004; Jackson and Schneider 2011; West 2018). Such studies, however, are hard to interpret causally; for example, companies serving Black customers could have stronger pools of Black applicants. In addition, observational studies of in-group bias usually cannot say which direction(s) discrimination runs in.[2] Meanwhile, lab experiments, which have extensively documented in-group biases, have strong causal identification, but do not explore effects on real-world decision-making (see Hewstone, Rubin, and Willis 2002 and Anderson, Fryer, and Holt 2006, for reviews).

Our research design, by contrast, can support strong causal inferences in a real-world setting. Similar field experiments, known as "audit" studies, have long been a key tool for studying discrimination in employment and other areas (Neumark, Bank, and Van Nort 1996; Riach and Rich 2002; Pager 2003; Bertrand and Mullainathan 2004; Lahey 2008; Oreopoulos 2011; Kroft, Lange, and Notowidigdo 2013; Deming et al. 2016). To our knowledge, this is the first audit study to closely examine neighborhood effects on racial discrimination in employment, other than Mobasseri (2019), a smaller study focused on neighborhood crime rates. Our study is particularly well suited to this purpose for several reasons. The targeted positions are overwhelmingly service jobs at employers with customer bases that are localized (mainly restaurants and retail). Those positions are distributed across two large jurisdictions (New Jersey and New York City), with wide variation in racial composition and other neighborhood characteristics. We tailored applications to be competitive for local jobs in all those localities, carefully choosing applicant addresses in nearby neighborhoods.

Our second major finding concerns one specific stereotyping mechanism that appears to contribute to the patterns we found: employers in communities with fewer Black residents appear more likely to hold negative stereotypes about Black criminality. In Agan and Starr (2018), which used the same data, we provided evidence that Ban-the-Box laws (which deprive employers of criminal record

[2] For example, if employers in Black neighborhoods are more likely to hire Black applicants, is that because they are biased in their favor, or because employers in other neighborhoods are biased against them, or both? Or could all the employers be biased in the same direction, but to different degrees? Because characteristics of the applicant pool are usually unobserved, these questions usually cannot be answered.

information) caused a large spike in the Black-white callback gap, suggesting that employers make negative assumptions about Black applicants' records. These assumptions appeared very exaggerated relative to real-world differences in conviction rates, supporting a stereotyping theory. Now, we show that the Ban-the-Box effect is driven entirely by neighborhoods with low Black shares, suggesting that stereotypes about race and crime are more prevalent there. Nonetheless, the interaction between neighborhood racial composition and applicant race remains quite strong even when employers *do* have individual criminal record information, implying that other forms of in-group bias or stereotyping are also at work.

Finally, we explore how these employment discrimination patterns combine with local variation in job *availability* (which our data also illustrate) to produce racial disparities in employment access. In our experimental sample, we artificially made black and white applicants come from identical neighborhoods and apply to identical jobs. To understand impacts on real-world populations, we also must take into account the real geographic distribution of jobs, relative to that of people—the topic of a large literature on "spatial mismatch," to which we also seek to contribute new empirical evidence. We applied to every job we could find within certain constraints in New Jersey and New York City, so our sample composition provides a snapshot of where job opportunities meeting these constraints are located.

Our analysis shows that although discrimination patterns may be reversed in Black neighborhoods, this does not mean their effects "even out." There are more white neighborhoods; job postings are quite disproportionately concentrated there; and overall callback rates are higher there. One might hope that for real-world populations, geographic self-sorting (the tendency to apply to jobs close to home) could mitigate disparities because Black applicants would tend to apply where they face less discrimination. But we show that this theory depends entirely on the geographic distribution of employers who are hiring. When jobs are scarce in and near Black neighborhoods, this sorting will *exacerbate* disparity.

We illustrate this point through simulations that reweight our data such that the Black and white applicant distribution by neighborhood mirrors the real-world population, incorporating commuting-time data to define hypothetical job search parameters. In New York City, where job availability very heavily favors white neighborhoods, all versions of these simulations predicted racial disparities far exceed those observed in our experiment itself (where Black and white applicants applied to the same jobs). The simulations project that white job-seekers in New York City will receive between 68% and 190% more callbacks than equally qualified Black job-seekers. In New Jersey, the job distribution pattern is more nuanced (white neighborhoods have more jobs *in* them, but Black neighborhoods are in denser regions and may have more jobs *near* them), and the simulation results varied based on the assumptions we made about job searches. Overall, the evidence that geographic self-sorting can alleviate racial disparities is weak.

I. Theoretical and Empirical Background

A. *Mechanisms of Discrimination*

Economists have long debated various explanations for the persistence of employment discrimination against Black applicants and other applicants of color (see Gersen 2007, for a useful review). We believe this exploration of mechanisms is important, but not because any form of such racial discrimination is more defensible legally or morally than another.[3] Rather, understanding how discrimination works can potentially shape responses. Likewise, some theories of discrimination suggest that it thrives in the absence of individualized information, which, if true, might shape policy concerning access of employers to certain types

---

[3] Early models of statistical discrimination "tacitly assumed" that it is efficient (Schwab 1986, 228), and this has been explicitly argued more recently (see, for example, Norman 2003), although others have countered that view or argued that it depends on circumstances (see Schwab 1986). Others, such as Epstein (1992), have argued that discrimination that appeals to the tastes of customers or staff is rational and that forbidding it is inefficient. U.S. law in general rejects these distinctions. Title VII of the Civil Rights Act of 1964 prohibits employers from engaging in disparate treatment based on race, whatever the reason.

of information. Here, we briefly discuss several potentially relevant theories of discrimination and related empirical research.

In-Group Preferences.—In-group preferences are preferences for one's own group. Lab experiments have extensively documented such biases (see Hewstone, Rubin, and Willis 2002 and Anderson, Fryer, and Holt 2006, for reviews). Implicit bias studies generally find that non-Black subjects display anti-Black prejudices, while Black subjects do not (Greenwald and Krieger 2006). Surveys have documented explicit racial in-group preferences (Greenwald and Krieger 2006).

Employment discrimination could potentially be shaped by the perceived in-group preferences of customers, hiring managers, existing staff, or even the applicant (if the employer seeks to increase yield from job offers). Observational research has provided suggestive evidence for such theories, although causal inference is complicated by lack of data about the applicant pool. Stoll, Raphael, and Holzer (2004) and Giuliano, Levine, and Leonard (2009) each found a correlation between race of hiring managers and race of those hired. Holzer and Ihlanfeldt (1998) find that customer demographics predict the race of companies' most recent hire, and Combes et al. (2016) find that relative underemployment of African immigrants in France is greater where their population concentration is lower. Such findings may suggest customer discrimination, although hiring managers and existing staff likely tend to reflect the community's demographics, so disentangling these potential mechanisms is difficult.[4]

Any of these forms of in-group bias might drive employment discrimination in the direction of "matching" local communities. But in-group preference in

[4] Becker (1957) argued that discrimination on the basis of the hiring manager's preference should not persist in a perfectly competitive market, while appeals to customers' racial prejudices *could* persist (manifesting primarily as wage and price gaps). Real-world markets, however, are subject to many frictions that alter this expectation. For example, wages and prices may be inelastic to local conditions due to laws and chain employer policies (DellaVigna and Gentzkow 2019), such that race gaps might primarily affect hiring *levels* instead. Moreover, recent research suggests that the concentration of U.S. labor markets gives employers considerable market power (Azar et al. 2020).

employment has been subjected to little experimental investigation. In an audit study, Bertrand and Mullainathan (2004) report one regression that finds a small, statistically insignificant interaction between the ZIP code's Black share and applicant race (in a sample not focused on jobs that serve local neighborhoods). They do not separate racial composition from other neighborhood characteristics. Neumark et al. (1996) found that women got fewer callbacks at restaurants with more male customers; with a small sample, this estimate was imprecise. Carlsson and Eriksson (2019) found that women had an advantage at Swedish firms with female recruiters and those with majority-female employees.

Statistical Discrimination and Stereotyping.—Theories of statistical discrimination suggest that absent reliable individualized information, employers and other decision-makers rely on group generalizations as proxies for unobservable, legitimately decision-relevant considerations (Phelps 1972; Arrow 1973; Stoll 2009; Fang and Moro 2011). To describe discrimination as "statistical" implies that it could serve the employer's economic ends (although it is still illegal and promotes racial disparities). But for this to be so, the generalization must have enough empirical basis to render reliance on it helpful in achieving those ends.

Some studies have sought to assess whether, as this theory predicts, discrimination is altered by the presence of individualized information. Several observational studies find that employer access to criminal records predicts hiring of African-Americans (Bushway 2004; Holzer, Raphael, and Stoll 2006; Stoll 2009; Doleac and Hansen 2020). Researchers have also found increased reliance on race after elimination of drug testing and credit checks (Wozniak 2015; Bartik and Nelson 2016; Ballance, Clifford and Shoag 2020). In Agan and Starr (2018), we found that white applicants had a much larger advantage after Ban-the-Box laws deprived employers of information about criminal records—suggesting that employers made race-based assumptions about criminality. Using a simple model to extrapolate employers' priors, we showed that this spike was far too large to be explained by real-world differences in the distribution of records. We concluded that employers were stereotyping (Agan and Starr 2018).

Stereotyping occupies a middle ground between what economists have traditionally defined as "taste-based" and statistical forms of discrimination. Like the statistical discrimination theory, stereotyping theories posit that employers will use racial generalizations to fill gaps in information (see, for example, Bordalo et al. 2016). But unlike purportedly "rational" statistical discrimination, stereotypes generally entail generalizations that are inaccurate or substantially exaggerated even as to group averages. Bohren et al. (2019) similarly discuss "inaccurate statistical discrimination" as a middle ground theory often ignored by scholars.

It would not be surprising if employers in communities with few Black residents more frequently relied on anti-Black stereotypes. In addition to the above-discussed implicit association tests (which document anti-Black associations that vary by race), survey evidence also shows, for example, that white respondents are more likely to hold exaggerated views of Black neighborhoods as crime-ridden (Quillian and Pager 2001). An interaction between neighborhood racial composition and applicant race would thus be consistent with a stereotyping theory, although also consistent with other theories described above (for example, appeals to customer racism). But this interaction is not predicted by "rational" statistical discrimination theories, unless the strength of race as a proxy for the employers' legitimate objective varies sharply by local racial composition. Such scenarios may occasionally be plausible, but (given that appealing to customer or staff prejudice is not itself a legitimate objective), they are probably not common, especially when one controls for community, business, and applicant characteristics. We return to this point in the Discussion below.

B. *Spatial Mismatch*

For decades, Black unemployment has been approximately double that of whites (DeSilver 2013). A longstanding literature explores the role of geography in shaping this gap (see Ihlanfeldt 1994, Kain 2004, and Gobillon, Selod, and Zenou 2007, for reviews). Much of that literature centers on evaluating the "spatial mismatch" hypothesis, developed initially by Kain (1968) (see Stoll and Covington 2012, for a more recent example). In its simplest form, Kain's hypothesis was that housing segregation reduces Black job opportunities and contributes to Black-white gaps in employment. A key premise of this theory is that jobs tend not to be located in Black neighborhoods; the "mismatch" in question is between the location of jobs and the residential location of people of different races (Kain 1968; Ihlanfeldt 1994). Another premise is that people tend to seek jobs close to their homes, to reduce commutes and search costs (Kain 1968; Ihlanfeldt 1994; Gobillo, Selod, and Zenou 2007). Indeed, considerable research suggests that Black residents may be particularly constrained from pursuing distant employment, for reasons including lower rates of car ownership and less public transit access (Kain 1968; Mouw 2000; Raphael and Stoll 2002; Johnson 2006; Gautier and Zenou 2010).

Most of the literature finds that job distributions disfavor Black communities (Raphael and Stoll 2002; Stoll 2006; Stoll and Covington 2012), and some find that this gap has expanded over time (Ihlanfeldt 1994; Mouw 2000; Stoll 2006; Gobillo, Selod, and Zenou 2007; Kneebone and Holmes 2015 give a contrary view). Many scholars have concluded that employment discrimination is a more important explanation than spatial mismatch for persistent disparities (see Ellwood 1986 and Leonard 1986, for seminal examples of this "race, not space" theory).

Kain (1968) and others have suggested that local variation in employment discrimination could exacerbate spatial mismatch, but this idea has rarely been tested empirically. The spatial mismatch literature generally has relied on economic survey data about residential location, employment status, and job location; such surveys do not typically include job application data (see Johnson 2006, for an exception), nor much information about applicant pools, limiting their utility in assessing whether employment gaps are driven by discrimination.

This paper, in contrast, does directly test local variation in discrimination in a causally rigorous way. We also seek to complement the spatial mismatch literature in other ways. Unlike most spatial mismatch studies, ours focuses not on overall job distribution, but on low-skill job vacancies, identified by typical modern job-search methods. We further assess the geographic distribution of employer callbacks—implicitly accounting for neighborhood differences in the applicant pool as well, which existing studies often do not take into account (Mouw 2000). We offer simulations that show how neighborhood differences in job postings and callback rates combine with local discrimination patterns to amplify disparities. We use much more recent data than most of the spatial mismatch literature analyzes. And we use rush-hour driving and public transit commuting-time data rather than aerial distance or driving distance as most studies have used.

II. Data and Experimental Design

The data analyzed in this paper come from a field experiment that we originally designed primarily to investigate the effect on racial discrimination of Ban-the-Box (BTB) laws (Agan and Starr 2018). That paper's online appendix includes considerable detail on the research design. We submitted online job applications to positions in New Jersey and New York City both before and after Ban-the-Box laws went into effect in 2015.[5] The outcome of interest is whether the applicant received a positive response (a "callback") in this period.

Our fictional applicants were Black and white non-Hispanic men in their early 20s. We created detailed applicant profiles, which our research assistants used to apply for jobs. The profiles were created in Black/white pairs, and all other characteristics were randomized. Other than race, the only substantive variations were felony conviction status (drug or property crime or no conviction), GED

[5] Applications were submitted in New Jersey between January 31 and February 28, 2015, and between May 4 and June 12, 2015. Applications were sent in New York City between June 10 and August 30, 2015 and between November 30, 2015 and March 31, 2016

versus regular high school diploma, and whether the applicant had a one-year gap in his prior employment history.[6] Applicant addresses were distributed across 40 towns in New Jersey and 44 neighborhoods in New York City and located in racially diverse, lower-to-middle-class Census blocks. Each business we applied to was assigned to applicants with addresses quite close to the employer; black and white applicants came from the same neighborhoods. This approach minimizes the possibility that racial discrimination in callbacks could be attributed to employers' beliefs about home neighborhoods or concerns about commuting times.

We signal race using applicant names, a ubiquitous practice in audit studies (Bertrand and Mullainathan 2004; Oreopoulos 2011). We identified distinctively Black and white first and last names by analyzing New Jersey birth certificates for men close to our applicants' age. We chose common names that met thresholds for the share of name-holders who were Black or non-Hispanic white.[7]

The jobs we applied to were entry-level positions for which our fictitious applicants were qualified, mostly in restaurants and retail. All had online applications (which most large chains use). We identified jobs by checking the websites of chains meeting size thresholds[8] and via online job boards. We sought to send four applications to each establishment: one Black-white pair before and after BTB.[9] Our job search covered all of New York City and almost all of New Jersey (91% by population). The exceptions were Newark (which already had a

[6] Felony conviction status was not actually conveyed for most applications in the sample; even before Ban-the-Box, 65% of the applications did not include a criminal record question.

[7] To reduce the concern that names also signal socioeconomic status, we picked white names from below the white median in maternal education (the strongest SES indicator available on birth certificates). Our job applications provided extensive socioeconomic information (e.g., work histories, education, address), and ensured that all our applicants were similar on those metrics, further reducing the likelihood that employers would need to rely on SES inferences from names. See Agan & Starr (2018) for further discussion of our approach.

[8] In New Jersey, we targeted chains with at least 30 locations and 300 employees in NJ. In New York, we applied to chains with at least 20 city locations, plus those we applied to in NJ.

[9] This was not always possible (job postings come and go), but the slight differences between the samples in the two application waves should not affect identification in this study.

BTB law), certain rural areas more than 20 miles from the nearest applicant neighborhood, and some very small townships. These exclusions should not threaten the internal validity of our callback-rate analysis, but could affect our analysis of job access gaps, and are thus accounted for in that analysis.[10]

Within these constraints and certain more limited exceptions, we sought to apply to every available job posting in New Jersey and New York City.[11] Our sample thus plausibly approximates the actual geographic distribution of these types of jobs within each of these jurisdictions. Between the two jurisdictions, New York City is somewhat overrepresented, because we had a longer available period to search for jobs there and used an additional search engine.[12]

We linked our experimental data to several outside data sources . Neighborhood racial composition and socioeconomic data (percent poverty, percent unemployed, median household income, and percent attending college) comes from the American Community Survey 2015 5-year estimates for Census block groups (CBGs).[13] Election results from 2016 are reported for New York City at the voting precinct level and for New Jersey at the municipality level.[14]

III. Analysis and Results

We submitted 15,213 applications. Table I provides summary statistics. The racial composition of the businesses' neighborhoods varies widely; both the white share and the Black share range from 0 to 1. The mean non-Hispanic white share is

[10] On balance these exclusions have little effect on the racial composition of the areas we study; Newark is disproportionately nonwhite and the other excluded places are disproportionately white.
[11] We excluded businesses with exceptionally time-consuming applications, those with overwhelmingly female clienteles, and those that required full Social Security numbers and had software that detected the fact that the numbers we provided were (for ethical reasons) invalid.
[12] Fifty-seven percent of our sample is from New York City, although New Jersey has a 5% larger population and a somewhat larger economy by most relevant measures.
[13] For 2% of the sample, these fields could not be coded at the CBG level (e..g,, service jobs in larger neighborhoods); instead we use Census tracts or Zip Code Tabulation Areas (ZCTAs).
[14] In robustness checks, we add fields that were unavailable for some observations: local crime data (from the 2015 Uniform Crime Reports for NJ municipalities, and from NYPD for NYC precincts) and employer characteristics from InfoGroup's BusinessUSA database.

50% (median 55%) and the mean Black share is 14% (median 5%).[15] New York City and New Jersey also have large population shares from other racial or ethnic groups, particularly Hispanic (sample mean 20%) and Asian populations (14%). The sample contains wide socioeconomic variation on all metrics, although businesses tend to be located in wealthier areas. The neighborhoods lean Democratic but vary widely; Donald Trump's local 2016 vote share ranges from 0% to 81% (mean 28%). Forty-six percent of the businesses are retail stores; the balance consists mostly of restaurants.

Overall callback rates were 10.6% and 13% for Black and white applicants—in proportional terms, whites received 23% more callbacks. In neighborhoods where the Black share is less than 1%, white applicants received 30% more callbacks than comparable Black applicants. In neighborhoods that are majority Black, however, Black applicants actually received 6% more callbacks than similar white applicants. Note, however, there are relatively few such neighborhoods in the sample: the 90th percentile of the percent Black distribution is 46% (by contrast 1.25% percent Black is the 25th percentile, and 16% of our applications were sent to businesses in neighborhoods with zero reported Black residents).

A. *Regression Analyses of Racial Differences in Callback Rates*

Does neighborhood racial composition play a causal role in shaping employment discrimination, or are these subsample differences driven by other correlated differences across neighborhoods? This is a question we cannot answer definitively; our experimental approach offers very strong identification of the effect of applicant race, but neighborhood racial composition is not experimentally

[15] Black-share figures include people who identify as both Black and Hispanic, but exclude those who identify as multiracial. The white share is for non-Hispanic whites, also excluding multiracial persons; meanwhile, the Hispanic share does not differentiate by race (so the Black share and the Hispanic share overlap). Our results are not notably affected by adding multiracial persons to their respective shares or by including Black Hispanics in the Black share; these affect the coding of only 1% of the neighborhood population for the average business in the sample.

manipulated, and we rely on a selection-on-observables assumption to infer its role. Still, we can at least disentangle the impact of racial composition from that of a rich set of other observables. In Table II, we show regressions estimating the interaction between neighborhood racial composition and applicant race in predicting callback rates. In the most basic specification (Column 1), we estimate the probability that applicant i to store j receives a callback as:

$$(1) \qquad Callback_{ij} = \alpha + \beta_1 White_i + \beta_2 PercentBlack_j + \beta_3 PercentBlack_j \ x \ White_i + \epsilon_{ij}$$

where $White_i$ indicates applicant race, $PercentBlack_j$ is the percent of the store's CBG that is Black, and $PercentBlack_j \ x \ White_i$ is their interaction. In other specifications, we also include additional neighborhood and business characteristics and their interactions with $White_i$.

In all specifications, the estimated effect of Black population share on racial discrimination is large and statistically significant. The $PercentBlack_j \ x \ White_i$ coefficient ranges from -5.9 points to -7.1 points--large effects, given that the Black baseline callback rate is only 10.6% and the overall white advantage in the sample is 2.4 percentage points. Column 1 shows the simplest specification, with no additional controls, in which the interaction's coefficient is -5.9 percentage points. In entirely non-Black neighborhoods, white applicants have a predicted 3.3 percentage point callback advantage (30.5% over the black baseline), while in all Black neighborhoods, Black applicants have a 2.6 percentage point advantage.

In Column 2, we add neighborhood characteristics, including Hispanic share, whether the store is in New Jersey (versus New York City), a socioeconomic status (SES) index which combines our four SES indicators, and Trump's 2016 local vote share for the voting precinct.[16] Each of these variables is interacted with

[16] We used principal component analysis to combine the 4 SES indicators (median household income, percent unemployed, percent poverty, and percent with a college degree) into a single index. In robustness checks we include them separately.

White. When these variables are added, the estimated $PercentBlack_j \ x \ White_i$coefficient increases slightly in magnitude, to -6.4 points. None of the other variables' interactions with *White* are significant.

Finally, in Column 3, we add chain fixed effects, interacted with *White*, to address the possibility that the racial-composition effect might be shaped by differences in businesses across neighborhoods. The $PercentBlack_j \ x \ White_i$effect only grows larger (-7.1 percentage points). In this specification, in entirely non-Black neighborhoods, we predict white applicants have a 3.4 percentage point callback advantage compared to Blacks, while in all-Black neighborhoods, Black applicants have a 3.7 percentage point advantage. In these (linear) specifications, we predict that Black applicants begin to have a suggestive advantage compared to white applicants at around 50% Black, although their predicted callback rates are not statistically significantly larger than those of white applicants until the population is greater than 80% Black.[17]

In Table III, we show the $PercentBlack_j \ x \ White_i$ coefficient from additional specifications and subsamples; these analyses test robustness and shed light on possible mechanisms. The baseline coefficient comes from Table II, Column 3, and is repeated for comparison in Column 1 of Table III. First, we add in different controls (always interacted with *White)*: *Percent Asian* (Col. 2), all four SES variables underlying our SES factor separately (Col. 3), chain characteristics from the Business USA database in place of chain fixed effects (Col. 4), and controls for property-crime and violent-crime rates (Col. 5).[18] None of these

[17] In analyses not shown here, we considered the possibility of a nonlinear relationship, but concluded that the linear approximation is probably not far off the mark. The callback rate differentials predicted by our linear models are very close to those that we observe when we segment the sample and separately estimate the white effect for very low, evenly mixed, and high Black shares. Our estimates are imprecise at the high end of that range, where the sample is thinner.

[18] The added business characteristics are: number of employees per location, sales per location, size of chain, and whether it is a retail business. Using these rather than Chain fixed effects allows the smallest chains (with only one store in the sample) to be used for identification, but it requires

substantially changes our results. In Columns 6 and 7 we show separate analyses for New Jersey and New York City. These point estimates are fairly similar (-5.7 and -6.8 percentage points, respectively), although not significant in the smaller New Jersey sample. The effect in New York City is substantially larger in proportional terms, given the lower overall callback rate.

In Table III Column 8, we test a variant of the statistical discrimination theory. One might postulate that statistical discrimination *could* explain our results, if the characteristics of Black and white applicants, relative to one another, were systematically different in neighborhoods with different Black shares (in ways known to employers yet not captured by the many fields on the job application). Although we know of no such reason, if one did exist, it would likely be driven by the applicant's neighborhood, not the business's. In Column 8, we add fixed effects for the applicant neighborhood interacted with *White*. Doing so does not change the $PercentBlack_j \ x \ White_i$ effect, suggesting that assumptions shaped by applicant neighborhood cannot explain the interaction; the *employer's* neighborhood is what matters. We discuss Columns 9 and 10 below.

B. Neighborhood Racial Composition and the Effect of Ban-the-Box

In Agan and Starr (2018), we explored the effect of New Jersey's and New York City's adoption of Ban-the-Box laws (BTB), requiring employers to drop criminal records questions on job applications. Among affected employers, we found a spike in racial discrimination: from a 7% white callback advantage (in proportional terms) before BTB to a 43% white advantage after. No such change occurred at companies that did not ask about criminal records even before BTB. In

---

certain chains not found in BusinessUSA to be dropped instead (257 observations) and does not control for unobserved chain characteristics. We omitted crime controls from the main specification because they are missing for a few non-reporting jurisdictions (dropping 89 observations) and because certain features of these data varied from New York City to New Jersey. The coefficient for our main specification in this sample is -0.07 (se=0.024).

Table V, we extend this analysis to examine differences in BTB's effects by neighborhood percent Black of the establishment. This analysis could shed light on a potential mechanism for the neighborhood's effect on racial discrimination: local variation in criminal-record-related stereotyping.

Our approach is a triple-differences regression: we assess the change after BTB in the Black-white gap at affected companies, after "differencing out" what happened over the same period at companies unaffected by BTB. We begin in Column 1 by replicating the simplest variant of the analysis in Agan and Starr (2018). The key term of interest is *White x Box Remover x Post*, where Post is an indicator variable for whether the application was sent after BTB and Box Remover indicates that the business's job application was affected by BTB: that is, that it had the "box" before BTB, and removed it afterward.[19]

In Column 1, the triple-differences estimate is 3.9 percentage points and is marginally significant, with a p-value just over 0.05. That is, we attribute to BTB a 3.9 percentage-point growth in the Black-white gap.[20] In Columns 2 and 3 respectively, we show the results for business neighborhoods with Black shares below and above the sample median. In Column 2 (below-median Black share: under 5.3% Black), the estimated effect is very large (6.8 percentage points) and statistically significant ($p<0.01$). In Column 3 (above-median Black share), the estimate is near zero and insignificant. Similar statistically significant results are obtained by quadruple-difference analyses (not shown here), adding either a high-black share indicator or a continuous black-share variable as a fourth difference.

[19] The sample is limited to the 74% of observations where we were able to send a complete set of four applications (one Black/white pair before BTB and one after), which allows us to ignore some differences in the job postings that were available in the pre- and post-BTB period, simplifying the analysis. Estimates are similar and more precise in the full sample with chain fixed effects.
[20] This causal inference depends on the identifying assumption that absent BTB, changes in disparity over time would have been similar at the *Remover* businesses and at other businesses. We discuss this assumption, and other research design questions, in Agan and Starr (2018).

These results suggest that our principal finding in Agan and Starr (2018)—that BTB greatly increases the Black-white callback gap—is driven overwhelmingly by employers in neighborhoods with low Black shares. In Agan and Starr (2018) we showed that employers' beliefs about racial disparities in criminal-record rates appeared to be exaggerated stereotypes. Our results here suggest that those stereotypes are held particularly by employers in the least Black neighborhoods, which would be consistent with prior survey and implicit-bias-test research suggesting that stereotypes about Black criminality are more common among white Americans (e.g., Quillian and Pager 2001).

However, we do not believe that, conversely, differences in crime-related stereotyping fully explain our main results in this paper—the relationship between neighborhood's racial composition and employment discrimination rates. Returning to Table III, in Columns 9 and 10 we show the $PercentBlack_j \ x \ White_i$ effect in subsamples defined by whether the application contained the criminal-records-question "box." The point estimate for "box" employers is only a little smaller than for "no box" employers (-6.5 versus -7.7 percentage points); it is statistically insignificant, but this is because it is estimated in a much smaller sample. Despite this imprecision, the similar point estimates suggest that neighborhood racial composition strongly predicts employer racial discrimination even when employers have individual data on criminal records and have no reason to make race-based assumptions about them. Some other form(s) of in-group bias, beyond crime-related stereotypes, must be contributing to this result.

C. Neighborhood-Level Disparities in Employment Access

Black applicants may have an advantage in pursuing jobs in Black neighborhoods—but there are fewer Black neighborhoods, and Black neighborhoods offer fewer jobs to apply to.[21] Treating the postings that we found as an approximation of where jobs are available, we illustrate this disproportion by comparing the distribution of postings in our sample and the overall (non-race-specific) distribution of callbacks we received to the distribution of these jurisdictions' Black and white populations.[22] Consistent with the spatial mismatch thesis, we find that employers that posted jobs and called us back are concentrated in whiter and less Black neighborhoods.

New Jersey's Black population share is 12.9%, but in our sample the mean neighborhood Black share is 10.9%, and among observations receiving callbacks, it is 9.3%. That is, job postings in New Jersey, and especially those that resulted in callbacks for our applicants, tended to be in neighborhoods that were less Black than the state average. Majority Black neighborhoods (defined by the Census block group, or CBG) have a particular dearth of job opportunities: only 3.6% of postings and 2.6% of callbacks in our sample were found there, even though 8.5% of New Jersey's CBGs are majority Black. Meanwhile, the distribution of jobs and callbacks favors neighborhoods with high non-Hispanic white shares.

In New York City, the geographic distribution of callbacks even more heavily favors white neighborhoods and disfavors Black neighborhoods. The city has an overall Black population share of 25%, but the mean employer neighborhood in our sample overall, and also among callbacks, had a Black share of 17.1%; meanwhile, white neighborhoods were overrepresented in jobs and callbacks. In

[21] The lower number of Black neighborhoods is obvious given population shares, but it gives context to our discrimination results: even if jobs were equally distributed across neighborhoods, our results suggest most employers would favor white applicants. The analysis below focuses on further *disproportion* in job access, relative to population shares.

[22] We consider the number of callbacks (undifferentiated by applicant race) to be an important measure of whether jobs are truly available. The distribution of postings alone is a starting point, but does not capture variation in employer eagerness to hire or in the strength of competition. Some employers accept applications constantly even if they are not immediately hiring, and some may be flooded with applications for scarce slots.

New York City, 25% of CBGs are majority Black, but that was true for only 12.6% of the employer CBGs in our sample and 13.9% of the callbacks.

D. Extrapolating Our Results to Real World Population Distributions

In our experiment, our fictitious Black and white applicants lived in the same places and applied to the same jobs. But in the real world, Black and white residents have different geographic distributions and will thus presumably apply to jobs in different places. Here, we carry out geographic reweighting exercises to extrapolate our results from the experimental sample to hypothetical Black and white job applicants who are distributed realistically across New Jersey's and New York City's neighborhoods (but who otherwise remain identically qualified).

Because employers in Black neighborhoods appear to favor Black applicants, one might think that the white advantage in callbacks would be *less* substantial given a realistic population distribution. Relative to our fictitious applicants, real Black applicants should apply more often to jobs in neighborhoods where they benefit from discrimination rather than being disadvantaged by it. If so, perhaps this geographic self-sorting mitigates discrimination patterns (much as in Becker's classic model of employer discrimination, Black applicants sort to less-discriminatory firms). This theory does not require assuming that applicants know employers' discrimination patterns—only that people apply to jobs near home.

But even if geographic self-sorting reduces the employment discrimination that black applicants are exposed to, it may not ultimately increase their employment opportunities. As the previous section showed, jobs and callbacks are not equally distributed, and geographic self-sorting will not help residents of Black neighborhoods if few employers are hiring there. Indeed, even setting aside the disproportion in the initial distribution of postings, our estimated callback rate for Black applicants is slightly *lower* (albeit not significantly) in entirely Black neighborhoods than in entirely non-Black neighborhoods (see the "Percent Black" main effect in Table II, Col. 1): the employment discrimination effect reverses, but this effect is canceled out by the fact that employers there rarely call any applicants back. Meanwhile, for white applicants, callback rates are much higher in non-Black neighborhoods. These patterns imply that for Black residents of Black neighborhoods, geographic self-sorting could mean sorting toward local employers that are both more scarce and less likely (or at least not more likely) to call them back. For white residents of white neighborhoods, it could mean sorting toward local employers that are more numerous and call them back much more often.

In Table VII, for each jurisdiction, we show three simulations that make differing assumptions about where people apply. All rely on three simplifying assumptions: (1) the relevant pools of Black and white job searchers (young, male, low-skilled) have the same geographic distribution as the Black and white populations as a whole, (2) our experimental sample's composition and callback rates by race are good proxies for the distribution of available job opportunities and callback rates by race for these pools, and (3) each job searcher applies to every business in our sample within the geographic parameters we define, and no other jobs. We believe the first two assumptions are reasonable as approximations.[23] The third is less realistic—many criteria influence job seekers' choices. Still, if the ways our assumptions diverge from reality are similar for Black and white job seekers, they should not substantially affect the white/Black ratios that we estimate.

In Panel A, we assume applicants apply to all jobs within their own Zip Code Tabulation Area (ZCTA), which is a mappable Census area that in most cases tracks ZIP codes. This represents a relatively strong version of the self-sorting

[23] The basic case for assumption (2) is described above: we sought to apply to every job posting we could find within parameters similar to those many real-world applicants likely use; we gave our large team of RAs many thousands of hours to find those jobs, and we believe they did a reasonably good job of doing so. We have no reason to believe that the businesses excluded by our search parameters would have different callback patterns from those in our data, though we cannot test this. However, using the BusinessUSA database, we can see that included and excluded businesses were similarly distributed by geography; if anything, excluded businesses are slightly *more* concentrated in white neighborhoods.

hypothesis: that people apply to jobs only within their fairly immediate neighborhoods (in New York City and denser New Jersey cities) or in their towns or communities (in less dense parts of New Jersey). In Panels B and C, we assume that they apply to all jobs within 15 and 30 minutes' rush-hour commute respectively, based on driving times in New Jersey and public transit times in New York City, without leaving the jurisdiction (New Jersey or New York City).[24] The thresholds we pick are fairly low but plausible proxies for the distances within which many candidates might search for a low-wage job.[25] Commuting time presumably matters directly to job seekers (and employers), whereas the ZCTA is a looser proxy for proximity.[26] But the commuting-time-based simulations have some drawbacks, especially affecting New Jersey, that the ZCTA approach avoids: a censoring problem for neighborhoods close to the state borders,[27] the need to

---

[24] Driving times are estimated using ArcGIS, and public transit times using the Google Distance Matrix API. We carry out these simulations separately for New Jersey and New York City for several reasons. First, assumption (2) above is more problematic if we combine the jurisdictions, because we oversampled New York City relative to New Jersey. Second, ZCTAs are dissimilar across jurisdictions; the average New York City ZCTA in our sample is smaller in area but has 81% more people than the average New Jersey ZCTA. Third, the Panels B and C simulations use different commuting methods because most New Jersey residents commute by car and most New York City residents commute by public transit. Finally, the simulation results highlight interesting differences across jurisdictions. Because of the differences mentioned here, one should not directly compare the estimated numbers of applications and callbacks across jurisdictions, but the callback rates and the white/Black ratios should be fairly comparable.

[25] These thresholds are fairly typical for New Jersey but somewhat low for New York City. Forman (2016) reports that the national average commute time is 26 minutes, that commutes are lower for low-wage sectors like retail, and that the average retail worker in New York City commutes 42 minutes. Low-wage workers have more geographically focused job searches due to limited transportation acces, search costs, and commuting costs (Kneebone and Holmes 2015; Allard and Danzinger 2002). We err on the side of using a lower threshold because the point of the simulations is to test whether geographic self-sorting would mitigate the effect of discrimination. With less self-sorting, we would expect results more like those in our experimental sample.

[26] ZCTAs also vary in geographic size. Within New York and New Jersey respectively, however, there are no strong correlations between a ZCTA's size and its racial composition, so there is no reason to believe that this variation substantially distorts the results of our simulation.

[27] While ZCTAs are confined within states, the 15- and 30-minute commuting radii are artificially truncated by New Jersey's long land borders, which means that job access will be understated for the periphery of New Jersey (which is disproportionately white) relative to the center. This censoring is not a significant issue for New York; given the city-based public transit system and the island geography, estimated commuting times from almost everywhere in NYC to points

impute missing data for the New Jersey municipalities left out of the original experiment,[28] and reliance on generalizations about commuting modality.[29] Thus, we report both methods. They produce similar patterns in New York City, but the choice of approach is important in New Jersey.

In each simulation, based on the assumptions above, we calculate the average number of job applications sent, callback rates, and total callbacks received for Black and white job searchers in New Jersey and New York City. We assign weights to each sample observation based on the probability that a New Jersey (or New York City) resident of the applicant's race would live near the business (within the specified parameters), and then use the reweighted sample distribution and callback outcomes to extrapolate the expected averages.[30]

Arguably, one could consider the disparity in total callback numbers to be the ultimate disparity of interest, because if job searchers do apply to all nearby

outside the city are rarely under 30 minutes anyway. In addition, across New York's five boroughs, only between 6% and 13% of residents commute out of the city (Forman 2016).

[28] In the ZCTA version, we simply omit Newark and the other municipalities that were excluded from our experiment; this creates no internal validity issues. In the commuting time versions, this approach would not work because very large shares of New Jersey are within 15 or 30 minutes of at least one of the omitted places. Instead, for these simulations, we conducted another job search in summer 2018 for only the omitted places; we did not apply to these jobs, but estimated an imputed callback probability based on the regression estimates for the main New Jersey sample.

[29] In New York City, for example, we assume everybody takes public transit and/or walks. But while this is true for most New Yorkers, there is geographic variation; in all Staten Island neighborhoods and a few outlying Queens neighborhoods, the majority commute by car (Forman 2016). Because car commuting is generally faster there, and because most of the neighborhoods with high car-commuting rates have relatively white populations, our assumption probably downward-biases our estimates of racial disparities in job access.

[30] We assign weights to each observation based on the share of New Jersey's or New York City's population of the applicant's race that lives in the ZCTA or within the given commuting time. After dividing this weight by 2 (because our sample includes two waves of applications to the same businesses), we sum the weights across the Black observations to get the number of expected applications per Black applicant in New Jersey/New York City, and likewise for the white observations. To get the total number of callbacks, we multiply the weights by the callback outcome and similarly sum them. Callback rates are the ratio of callbacks to applications. For the ZCTA version in New Jersey, the population denominator excludes the ZCTAs that were excluded from our search (about 9% of the population). For the other versions, the business data includes the 118 new businesses that we found in summer 2018 in those omitted areas (assigned two white and two Black observations each) and imputed callback results for them, as discussed above.

jobs, this represents the disparity in overall access to employment opportunity. That said, we think the callback *rate* disparity is probably even more important, because in the real world, applying to every nearby job is costly and unrealistic. So for many applicants (especially those in dense areas with many employers nearby), what may matter the most is the success rate when one does apply. The white/Black callback-rate ratio can also be understood as a measure of how many more jobs the average Black job-seeker would have to apply to in order to obtain the same number of callbacks as a white job-seeker with the same qualifications.

Table VII presents the results of these various simulations. In New York City, in all simulation versions, the projected racial disparity in callback rates and (especially) total callbacks is much larger than the 8% white advantage observed in our experimental sample. The projected callback-rate disparity is similar in all versions: compared to identical Black job searchers who apply to the same number of jobs, the average white NYC job searcher receives between 18% and 21% more callbacks. Because both the distribution of job postings and the city's public transit network very heavily favors whiter and less Black neighborhoods, the projected disparity in total callbacks is much more dramatic. Here the projected figures vary across simulations: whites receive more callbacks by a factor of 1.67 (ZCTA version), 2.9 (15-minute commute), or 2.07 (30-minute commute).

In New Jersey, the patterns differ more across simulations. In the ZCTA version, as in New York City, the projected disparities for the realistically distributed population are much larger than the white advantage in our experimental sample, which was 38%. White applicants are projected to have 73% higher callback rates than identical Black applicants, and to receive 96% more total callbacks if they apply to all nearby jobs. In the 15-minute commuting-time version, the callback *rate* disparity is 54%, again much higher than in the experiment; in the 30-minute commuting-time version, it is 39%, very similar to what we saw in the experiment. However, the disparity in *total* callbacks in these simulation versions is far smaller than what we saw in the experiment (a 3% and 5% white advantage for the 15- and 30-minute versions, respectively). This is because we project that although Black applicants will face substantially lower callback rates than white applicants, they will find more postings within a 15- or 30-minute commute.

Why do the New Jersey results, particularly for total callbacks, vary so much across simulations? One possibility is that this represents a real phenomenon: Black residents in New Jersey may tend to have fewer jobs available immediately *in* their ZIP codes but more jobs fairly nearby, perhaps because they live in denser areas. Another possibility is that it represents a data limitation: white New Jersey residents are more likely to live near the state borders, and thus more likely to have our count of nearby jobs artificially truncated (a problem that does not affect the ZCTA version and does not substantially affect NYC, as discussed above). To explore further, we conducted some Census-tract-level regressions (not shown in the tables) using the numbers of nearby jobs and callbacks in our sample as outcome variables. We found that New Jersey Census tracts with higher Black shares indeed had fewer jobs and callbacks within their ZCTAs, but more jobs and callbacks within the 15- and 30-minute commuting thresholds. However, the latter pattern (but not the ZCTA pattern) disappeared when we added fixed effects for the county (effectively controlling for proximity to the border, plus other county-level characteristics including density).[31] This suggests that the apparent Black advantage in commutable jobs may have been mainly an artifact of data censoring.

Overall, the New Jersey results are complicated to interpret. Unlike in New York City, the commuting-time versions project that in a realistically distributed population, disparities in total callbacks will be less than we found in the

[31] Point estimates dropped to close to zero. Adding controls for tract-level population density, instead of fixed effects, reduced but did not eliminate the apparent Black advantage in nearby jobs. Note even within counties and controlling for tract density, there is no evidence that Black NJ residents have *fewer* jobs within 15- or 30-minute commutes—a sharp contrast to the NYC results.

experiment, but this finding may be artificial due to the border censoring problem. That said, we can draw a few conclusions. First, , the callback *rate* disparity is not similarly subject to the censoring problem (which equally affects the numerator and denominator of the rates), and in all the simulations, this disparity is either similar to or substantially larger than what we found in the experiment. This implies that even if Black New Jersey residents *can* obtain almost the same number of callbacks within a 15- or 30-minute commute as similar white residents can, they can only do so by taking on the costs of applying to a far larger number of jobs (between 39% and 73% more, depending on the simulation version). Second, even in New Jersey, job postings are disproportionately not located *within* Black communities (and overall callback rates are much lower there), and that fact contributes to the employment discrimination that Black residents face, given the relationship between neighborhood demographics and discrimination patterns.

Overall, the simulations offer little support for the optimistic view that geographic self-sorting will alleviate the effects of employment discrimination. Indeed, the New York City results illustrate that when the distribution of jobs and transit access to those jobs decisively favors white neighborhoods, geographic self-sorting will sharply exacerbate disparities. In addition, a key assumption of the simulations—that everybody commutes the same way and has the same geographic parameters—is conservative, attenuating disparity estimates. Black residents in reality are less likely to own cars and to have reliable transit, and work closer to home—so real-world disparities among similarly qualified applicants may be larger (Raphael and Stoll 2002; Gautier and Zenou 2010).

VI. Discussion and Conclusions

This study offers robust experimental evidence of employment discrimination that on balance significantly favors white applicants. The magnitude and even the direction of this discrimination is strongly predicted by the racial composition of the employer's neighborhood, after accounting for other observable neighborhood characteristics. Employers in heavily non-Black neighborhoods strongly favored white applicants; this effect shrinks as the Black share grows, and it reverses in predominantly Black neighborhoods. Our sample also provides evidence that employers generally do less hiring in Black neighborhoods. Moreover, our simulations show how discrimination patterns can interact with neighborhood-level hiring disparities to produce large race gaps in job access. Finally, our Ban-the-Box results suggest that anti-Black stereotyping about crime is more common in non-Black neighborhoods.

A. *Mechanisms of discrimination*

Although this core finding may seem unsurprising to many readers, it has not previously been demonstrated in a study that, like ours, offers strong causal identification of racial discrimination. Moreover, it is not predicted by at least one of the major economic theories of discrimination. Our findings are quite hard to reconcile with the idea that this racial discrimination is driven by so-called "rational" statistical discrimination. Employers may well be relying on race-based assumptions about applicants' merits—that is, stereotypes or prejudices. But our findings belie the idea that these assumptions are empirically grounded.

In our study, the employers in different neighborhoods were similar in type—indeed, our fixed-effects specifications are identified based on differences among different locations of the same chain—and thus presumably have generally similar objectives in terms of worker productivity. It is hard to think of a plausibly empirically supportable reason (unrelated to stereotypes or to the racial prejudices of customers, staff, or others) why employers in non-Black neighborhoods would view Black candidates as likely to be less productive than white candidates, while employers in Black neighborhoods do not hold that view or even think the opposite. It is even less conceivable in our experimental setting, where employers were provided with extensive individualized information about applicants' characteristics, and white and Black applicants to jobs in all neighborhoods had an identical distribution of those characteristics. And it makes even less sense given that the effect does not turn on the *applicant's* address, which is presumably more relevant to expectations about the applicant than the business address is.

Instead, some other mechanism must be at work, but our data are potentially consistent with several such mechanisms, so there is a limit to our ability to disentangle them. Hiring managers could be engaging in stereotyping, or indulging sheer animus. Our Ban-the-Box results suggest that one specific stereotyping mechanism (stereotyping about Black criminality, which is more prevalent in non-Black neighborhoods) might contribute, but does not explain *most* of the neighborhood effect on employment discrimination. Employers could be relying on other racial stereotypes as well, or they could be catering to the perceived preferences of customers or existing staff, or they could be assuming that an applicant whose race does not "fit" the neighborhood would be less likely to accept the job or to stay in it. It seems plausible that all of these channels could simultaneously contribute to the strong effects that we found.

B. *Spatial mismatch*

Our simulations show how two spatial dimensions of racial disparity in employment interact: local variation in discrimination patterns and disparities in geographic access to jobs. They illustrate a dilemma facing Black job seekers who live in Black neighborhoods with limited local employment: searching only close to home reduces available openings, but searching in white neighborhoods instead makes adverse racial discrimination more likely. While some aspects of the New Jersey simulations are difficult to interpret given the state-border censoring problem, in New York City at least, the job access disparity is so substantial that (despite the fact that it reduces exposure to discrimination) geographic self-sorting appears to be on balance counterproductive for Black job seekers. Our analyses and simulations are broadly supportive of some of the premises of the spatial mismatch literature. We provide strong evidence for one of the mechanisms originally suggested by Kain (1968): employers favor candidates who share the racial background of the neighborhood. This mechanism suggests that the "race not space" dichotomy may be oversimplified—space mediates the role of race in hiring.

Our analyses do not directly address the spatial mismatch hypothesis itself: that housing segregation drives employment disparities. But we can use our data to give a back-of-the-envelope answer to a simple counterfactual: What disparity would our estimates predict if New Jersey and New York City had no housing segregation (if all else could be held equal)? Suppose every Census block group throughout each jurisdiction had the same Black share, equal to the jurisdiction's mean (12.9% for New Jersey and 25.1% for New York City). Applying the coefficients from our simplest specification (Equation 1) estimated separately in each jurisdiction, we can predict Black and white callback rates.

In this counterfactual exercise, the number of available nearby postings is identical by race, eliminating any disparities in geographic job access. What is left is the racial discrimination effect on callback rates, which also declines, consistent with the spatial mismatch hypothesis. For a completely desegregated New Jersey, white applicants receive 37% more predicted callbacks than identical Black applicants. This is still a large advantage, but it is slightly smaller than what we observed in our New Jersey experimental sample (a 38% white advantage) and smaller than the predicted real-world white callback-rate advantage from our simulations (which ranges from 39% to 73% across the simulations). For a completely desegregated New York City, whites are predicted to have just a 2% higher callback rate, compared to 8% higher in our experimental sample and about 20% higher in the simulations. The reason is straightforward: the racial discrimination effect is estimated conditional on *Percent Black* being fixed at the jurisdictional mean instead of the lower sample mean, and anti-black discrimination is lower in neighborhoods with higher black populations.

In the real world, additional dimensions of disparity can be expected to compound those we identified in our experiment and simple simulations. As noted above, these include racial differences in transportation access, which our simulations ignored. In addition, because our fictitious candidates had identical characteristics across race and space, we ignore other structural contributions to employment disparities beyond spatial mismatch and employment discrimination, such as differences in educational opportunities.

C. *Ban-the-Box*

This paper also adds to our findings concerning Ban-the-Box laws (BTB) in Agan and Starr (2018). Those findings showed, consistent with BTB's premise, that people with records faced a large callback disadvantage, and BTB substantially redressed it. But BTB also had the unintended consequence of triggering a large increase in the Black/white callback gap. The results in Table IV of this paper suggest that the latter consequence was heavily driven by employers in more white and less Black communities; employers in communities with more Black residents appear less likely to stereotype Black applicants based on assumptions about criminality. This distinction suggests that BTB's effects may differ across jurisdictions. Many cities that have adopted BTB are mostly or substantially Black, and our results here suggest that BTB should not be expected to increase racial discrimination in those cities. However, many other mostly white cities or states have adopted or are considering BTB, and for those jurisdictions our result here has the opposite implication. In portions of our sample with similar demographics, BTB seems to have led to an especially dramatic spike in racial discrimination.

D. *Other Limitations*

Beyond the challenges of identifying causal mechanisms discussed above, our study has important limitations. We focused specifically on low-skilled jobs in two jurisdictions, and on young Black and white men exclusively. These sectors are important: the retail and restaurant sectors together employ about 20% of the entire U.S workforce, and are key sources of jobs for the low-skilled workers that are on the margins of unemployment. However, further studies could usefully explore own-group bias and neighborhood effects in other markets, as well as on other racial and ethnic groups and on women and older workers.

Likewise, while we believe that our sample of employers was fairly representative of the jobs available in these sectors in New York City and New Jersey over the study period, we cannot say whether our results (either on callback discrimination or on the location of jobs) are representative of other places. On the other hand, one should not assume disparities are *less* serious elsewhere. If the pattern that employers in more white and less Black communities are more likely to discriminate against Black workers does hold throughout the U.S., many other places in the U.S. are likely to show larger Black/white disparities than we observed in our sample, which was drawn from a racially diverse state and from the United States' most diverse major city.

References

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

Anderson, Lisa, Roland Fryer, and Charles Holt. 2006. "Discrimination: Experimental Evidence from Psychology and Economics." In *Handbook on the Economics of Discrimination*, edited by William M. Rodgers III, Chapter 4: Edward Elgar

Arrow, Kenneth Joseph. 1973. "The Theory of Discrimination," In *Discrimination in Labor Markets* edited by Ashenfelter, Orley, and Albert Rees, Princeton, NJ: Princeton University Press, 3-33.

Azar, José, Ioana Marinescu, and Marshall Steinbaum. 2020. "Labor Market Concentration." The *Journal of Human Resources*, 1218-9914R1

Ballance, Joshua, Robert Clifford, and Daniel Shoag. 2020. "No More Credit Score": Employer credit check bans and signal substitution." *Labour Economics* 63: 101769.

Bartik, Alexander Wickman, and Scott Thomas Nelson. 2016. "*Credit Reports as Resumes: The Incidence of Pre-employment Credit Screening*." MIT Economics Department Working Paper 16-01.

Becker, Gary Stanley. 1957. "*The Economics of Discrimination: An Economic View of Racial Discrimination*." University of Chicago.

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

Bordalo, Pedro, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. 2016. "Stereotypes." *The Quarterly Journal of Economics* 131 (4): 1753-1794.

Bohren, J. Aislinn, Kareem Haggag, Alex Imas, and Devin G. Pope. 2019. "Inaccurate Statistical discrimination." NBER Working Paper 25935.

Bushway, Shawn D. 2004. "Labor Market Effects of Permitting Employer Access to Criminal History Records." *Journal of Contemporary Criminal Justice* 20 (3): 276-291.

Carlsson, Magnus, and Stefan Eriksson. 2019. "In-Group Gender Bias in Hiring." *Economics Letters*, 185: 108686.

Combes, Pierre-Philippe, Bruno Decreuse, Morgane Laouenan, and Alain Trannoy. 2016. "Customer Discrimination and Employment Outcomes: Theory and Evidence from the French Labor Market." *Journal of Labor Economics* 34 (1): 107-160.

Deming, David J., Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz. 2016. "The Value of Postsecondary Credentials in the Labor Market: An Experimental Study." *American Economic Review* 106 (3): 778-806.

DellaVigna, Stefano, and Matthew Gentzkow. 2019. "Uniform Pricing in US Retail Chains." *The Quarterly Journal of Economics* 134 (4): 2011-2084.

Desilver, Drew. 2013. "Black Unemployment Rate Is Consistently Twice That of Whites." *Pew Research Center* 21.

Doleac, Jennifer L., and Benjamin Hansen. 2020. "The Unintended Consequences of "Ban the Box": Statistical Discrimination and Employment Outcomes When Criminal Histories Are Hidden." *Journal of Labor Economics* 38(2): 321-374.

Ellwood, David T. 1986. "The Spatial Mismatch Hypothesis: Are There Teenage Jobs Missing In the Ghetto?." In *The Black Youth Employment Crisis*, 147-190. Chicago: University of Chicago Press.

Epstein, Richard A.1992. *Forbidden Grounds: The Case Against Employment Discrimination Laws*. Cambridge: Harvard University Press.

Fang, Hanming, and Andrea Moro. 2011. "Theories of Statistical Discrimination and Affirmative Action: A Ssurvey." In *Handbook of Social Economics*, vol. 1, edited by Behabib, Jess, Matthew O. Jackson, and Alberto Bisin, 133-200. The Netherlands: North-Holland.

Forman, Adam. 2016. "Fast City, Slow Commute." Center for an Urban Future Data Brief. https://nycfuture.org/data/fast-city-slow-commute.

Gautier, Pieter A., and Yves Zenou. 2010. "Car Ownership and the Labor Market of Ethnic Minorities." *Journal of Urban Economics* 67 (3): 392-403.

Gersen, Jacob E. 2007. "Markets and Discrimination." *New York University Law Review* 82: 689-737.

Gobillon, Laurent, Harris Selod, and Yves Zenou. 2007. "The Mechanisms of Spatial Mismatch." *Urban Studies* 44 (12): 2401-2427.

Giuliano, Laura, David I. Levine, and Jonathan Leonard. 2009. "Manager Race and the Race of New Hires." *Journal of Labor Economics* 27 (4): 589-631.

Greenwald, Anthony G., and Linda Hamilton Krieger. 2006 "Implicit Bias: Scientific Foundations." *California Law Review* 94 (4): 945-967.

Hewstone, Miles, Mark Rubin, and Hazel Willis. 2002. "Intergroup Bias." *Annual Review of Psychology* 53 (1): 575-604.

Holzer, Harry J., and Keith R. Ihlanfeldt. 1998. "Customer Discrimination and Employment Outcomes for Minority Workers." *The Quarterly Journal of Economics* 113 (3): 835-867.

Holzer, Harry J., Steven Raphael, and Michael A. Stoll. 2006. "Perceived Criminality, Criminal Background Checks, and the Racial Hiring Practices of Employers." *The Journal of Law and Economics* 49 (2): 451-480.

Ihlanfeldt, Keith. 1994. "The Spatial Mismatch Between Jobs and Residential Locations Within Urban Areas." *Cityscape* 1 (1): 219-244.

Jackson, C. Kirabo, and Henry S. Schneider. 2011. "Do Social Connections Reduce Moral Hazard? Evidence from the New York City Taxi Industry." *American Economic Journal: Applied Economics* 3 (3): 244-67.

Johnson, Rucker C. 2006. "Landing a Job in Urban Space: The Extent and Effects of Spatial Mismatch." *Regional Science and Urban Economics* 36 (3): 331-372.

Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *The Quarterly Journal of Economics* 82 (2): 175-197.

Kain, John F. 2004. "A Pioneer's Perspective on the Spatial Mismatch Literature." *Urban Studies* 41 (1): 7-32.

Kneebone, Elizabeth and Natalie Holms. 2015. “The Growing Distance between People and Jobs and Metropolitan America”, Metropolitan Policy Program at Brookings. http://www.brookings.edu/research/the-growing-distance-between-people-and-jobs-in-metropolitan-america/

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo. 2013. "Duration Dependence and Labor Market Conditions: Evidence from A Field Experiment." *The Quarterly Journal of Economics* 128 (3): 1123-1167.

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

Lahey, Joanna N., and Ryan A. Beasley. 2009. "Computerizing Audit Studies." *Journal of Economic Behavior & Organization* 70 (3): 508-514.

Leonard, Jonathan S. 1986. “Comment,” in *The Black Youth Employment Crisis* (Freeman & Holzer eds.), University of Chicago Press: 185-189.

Leonard, Jonathan S., David I. Levine, and Laura Giuliano. 2010. "Customer Discrimination." *The Review of Economics and Statistics* 92 (3): 670-678.

Mobasseri, Sanaz. 2019. "Race, Place, and Crime: How Violent Crime Events Affect Employment Discrimination." *American Journal of Sociology* 125 (1): 63-104.

Mouw, Ted. 2000. “Job Location and the Racial Gap in Unemployment in Detroit and Chicago, 1980 to 1990,” *American Sociological Review,* 65: 730-753.

Neumark, David, Roy J. Bank, and Kyle D. Van Nort. 1996. "Sex Discrimination in Restaurant Hiring: An Audit Study." *The Quarterly Journal of Economics* 111 (3): 915-941.

Norman, Peter. 2003. "Statistical Discrimination and Efficiency." *The Review of Economic Studies* 70 (3): 615-627.

Oreopoulos, Philip. 2011. "Why Do Skilled Immigrants Struggle in the Labor Market? A Field Experiment With Thirteen Thousand Resumes." *American Economic Journal: Economic Policy* 3 (4): 148-71.

Pager, Devah. 2003. "The Mark of a Criminal Record." *American Journal of Sociology* 108 (5): 937-975.

Phelps, Edmund S. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62 (4): 659-661.

Phillips, David 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, Lincoln, and Devah Pager. 2001. "Black Neighbors, Higher Crime? The Role of Racial Stereotypes in Evaluations of Neighborhood Crime." *American Journal of Sociology* 107 (3): 717-767.

Quillian, Lincoln, Devah Pager, Ole Hexel, and Arnfinn H. Midtbøen. 2017. "Meta-Analysis of Field Experiments Shows No Change in Racial Discrimination in Hiring over Time." *Proceedings of the National Academy of Sciences* 114 (41): 10870-10875.

Raphael, Steven, and Michael A. Stoll. 2002. "Modest Progress: The Narrowing Spatial Mismatch Between Blacks and Jobs in the 1990s." *Washington, DC: The Brookings Institution*.

Riach, Peter A., and Judith Rich. 2002. "Field Experiments of Discrimination in the Market Place." *The Economic Journal* 112 (483): F480-F518.

Schwab, Stewart. 1986. "Is Statistical Discrimination Efficient?." *American Economic Review* 76 (1): 228-234.

Stoll, Michael A. 2009. "Ex-offenders, Criminal Background Checks, and Racial Consequences in the Labor Market." *University of Chicago Legal Forum*: Vol. 2009: Iss. 1, Article 11.

Stoll, Michael A. 2006. "Job Sprawl, Spatial Mismatch, and Black Employment Disadvantage." *Journal of Policy Analysis and Management* 25 (4): 827-854.

Stoll, Michael A., Steven Raphael, and Harry J. Holzer. 2004. "Black Job Applicants and the Hiring Officer's Race." *Industrial and Labor Relation Review* 57 (2): 267-287.

Stoll, Michael A., and Kenya Covington. 2012. "Explaining Racial/Ethnic Gaps in Spatial Mismatch in the US: The Primacy of Racial Segregation." *Urban Studies* 49 (11): 2501-2521.

West, Jeremy. 2018. "Racial Bias in Police Investigations." Working Paper. Retrieved from *University of California, Santa Cruz* website: https://people. ucsc. edu/~ jwest1/articles/West_RacialBiasPolice. pdf.

Wozniak, Abigail. 2015. "Discrimination and the Effects of Drug Testing on Black Employment." *Review of Economics and Statistics* 97 (3): 548-566.

Table I: Summary Statistics

A. Employer Geography	**Mean**	**SD**	**Min**	**Max**
NJ	0.43			
Percent Black	0.14	0.22	0.00	1.00
Percent White (Non-Hispanic)	0.50	0.29	0.00	1.00
Percent Asian	0.14	0.16	0.00	0.95
Percent Hispanic	0.20	0.20	0.00	0.97
Median Household Income	$79,200	$40,600	$9,000	$250,000
Percent Poverty	0.14	0.13	0.00	1.00
Percent College Educated	0.49	0.24	0.00	1.00
Percent Unemployed	0.05	0.04	0.00	0.36
Trump Percent Local Vote	0.28	0.20	0.00	0.81
B. Callback Rates	**Black**	**White**	**Ratio (W/B)**	**N**
Overall	0.106	0.130	1.23	15213
Black share <=1%	0.113	0.147	1.3	7601
1%<Black share <=50%	0.103	0.128	1.24	10433
Black share >50%	0.108	0.102	0.94	1323

Notes: The *Percent Black*, *Percent Non-Hispanic White*, *Percent Asian*, *Percent Hispanic White*, *Median Household Income*, and *Percent Poverty* variables are drawn from 2011-2015 ACS data on the employer's Census Block Group; tract-, county-, or municipality-level data was used for a small number of cases in which the census block was nonresidential. *Trump Percent Local Vote* is reported for the 2016 general election at the voting precinct level in New York City and the municipality level in New Jersey. The table has 15,213 observations.

Table II. Effect of Black Share and Other Variables on Black-White Callback Gap

	(1) Callback	(2) Callback	(3) Callback
Percent Black X White	-0.059***	-0.064***	-0.071***
	(0.016)	(0.021)	(0.024)
Percent Black	-0.010	0.062*	0.035
	(0.029)	(0.036)	(0.024)
White	0.033***	0.034**	-0.056***
	(0.007)	(0.014)	(0.013)
Percent Hispanic X White		-0.053**	-0.041
		(0.026)	(0.026)
Percent Hispanic		-0.008	-0.007
		(0.030)	(0.025)
White x NJ		0.039***	0.044***
		(0.012)	(0.014)
NJ		0.014	-0.005
		(0.016)	(0.016)
Constant	0.108***	0.108***	0.108***
	(0.015)	(0.015)	(0.015)
Observations	15213	15213	15213
Nbhd Char (X White)	No	Yes	Yes
Chain FE (X White)	No	No	Yes

Notes: Standard errors clustered on the chain in parentheses. Race data are from the 2011-2015 ACS data on the employer's Census Block Group; tract-, county-, or municipality-level data was used for a small number of cases in which the census block was nonresidential. Nbhd Char are characteristics of the neighborhood: SES factor (combining median household income, percent unemployed, percent poverty, and percent with a college degree) and Percent voting for Trump in 2016 (for voting precincts in NYC and municipalities in NJ). *p<0.1 **p<0.05 ***p<0.01.

Table III. Interaction of Black Population Share with Race Gap in Callbacks: Alternative Specifications and Samples

	(1)	(2)	(3)	(4)	(5)
Percent Black x White	-0.071***	-0.068***	-0.074***	-0.057***	-0.066***
	(0.024)	(0.026)	(0.025)	(0.024)	(0.026)
Variant	Main	Add % Asian	Separate SES Vars	Business Chars	+ Crime Ctrls
Observations	15213	15213	15213	14849	15104

	(6)	(7)	(8)	(9)	(10)
Percent Black x White	-0.057	-0.068***	-0.071***	-0.065	-0.077**
	(0.056)	(0.029)	(0.024)	(0.043)	(0.031)
Variant	NJ	NYC	Center FE	Has Box	No Box
Observations	6600	8613	15213	3110	12103

Notes: Standard errors clustered on the chain in parentheses. Only the *Percent Black X White* coefficient is shown, but the regressions parallel those shown in Table II with the variations as indicated. And added control is also interacted with White. Column 1 repeats Table II. Column 3; Column 2 adds in % Asian; Column 3 includes each SES indicator separately rather than the index (median household income, percent unemployed, percent poverty, and percent with a college degree); Column 4 adds in business characteristics: number of employees per location, sales per location, size of chain, and whether it is a retail business, this information is missing for 257 observations not found in the BusinessUSA data and mimics Table II Column 2 instead of 3 as it does not include Chain Fixed Effects; Column 5 adds in controls for property crime and violent crime rates (which is missing for some non-reporting jurisdictions); Column 6 is only in NJ; Column 7 only in NYC; Column 8 adds in fixed effects for the area the *applicant* lived in; Column 9 is amongst applications that had a criminal record check box and column 10 is amongst those that did not.
*p<0.1 **p<0.05 ***p<0.01.

Table IV. Triple-Differences Estimates of Ban-the-Box's Effects on Race Gap in Callbacks: Subsamples of the Percent Black Distribution

	(1)	(2)	(3)
Subset of Balanced Sample	*All*	*Percent Black Below Median*	*Percent Black Above Median*
Post x Box Remover x White	0.039*	0.068***	0.009
	(0.020)	(0.024)	(0.024)
Post x White	-0.002	-0.018	0.014
	(0.014)	(0.017)	(0.015)
Post X Box Remover	-0.019	-0.034	-0.003
	(0.023)	(0.028)	(0.022)
Box Remover X White	-0.017	-0.038**	0.005
	(0.015)	(0.018)	(0.018)
Box Remover	0.017	0.028	0.006
	(0.028)	(0.028)	(0.030)
White	0.024**	0.033**	0.014
	(0.012)	(0.014)	(0.013)
Post	0.016	0.034	-0.003
	(0.017)	(0.023)	(0.013)
Constant	0.090***	0.093***	0.087***
	(0.021)	(0.021)	(0.024)
Observations	11184	5648	5536

Notes: Standard errors clustered on the chain in parentheses. *Post* is an indicator for the post-Ban-the-Box (BTB) wave of applications. *Box Remover* is an indicator for businesses whose job applications were changed by BTB: those who had the criminal records "box" before BTB and then removed it. *White* indicates applicant race. All regressions are conducted within the sample of businesses to which we sent complete sets of four observations (one black/white pair before and after BTB), and for which we are able to code *Percent Black*. See Agan and Starr (2018) Table V Column 1.

Table V. Neighborhood Racial Composition and Job Locations

A. New Jersey

	Our Sample Job Postings	Our Sample Callbacks	New Jersey*
Mean Black Share in CBG	10.9%	9.3%	12.9%
Mean White Share in CBG	58.9%	63.2%	57.0%
Percent of CBGs Majority Black	3.6%	2.6%	8.5%

B. New York City

	Our Sample Job Postings	Our Sample Callbacks	New York City
Mean Black Share in CBG	17.1%	17.1%	25.1%
Mean White Share in CBG	43.2%	45%	33%
Percent of CBGs Majority Black	12.6%	13.9%	25%

Notes: Population shares are reported at the Census block group level based on the 2011-2015 5-year American Community Survey. The white share is non-Hispanic only. The baseline comparison figures for New Jersey are drawn from the same portion of New Jersey that our job search covered, including about 91% of the state's population. "Job postings" refer to observations in the sample, such that employers to which we applied more than once (as intended by the research design) are reported more than once, because the main reason some employers received fewer applications than others is that they were hiring less often.

Table VI. Simulating Job Access for Realistic Racial Geographic Distribution

A. Applications to All Postings Within Zip Code Tabulation Area of Residence

	Applications Per Capita	Mean Callback Rate	Callbacks Per Capita
New Jersey			
Black	4.86	10.5%	0.51
White	5.49	18.2%	1.00
White/Black Ratio	1.13	1.73	1.96
New York City			
Black	11.96	8.6%	1.03
White	16.70	10.3%	1.72
White/Black Ratio	1.40	1.20	1.67
B. Applications to All Postings within 15 Minute Commute Time (within NJ/NYC)			
New Jersey			
Black	266	9.40%	25.03
White	176	14.55%	25.69
White/Black Ratio	0.66	1.54	1.03
New York City			
Black	23.41	8.49%	1.99
White	56.26	10.28%	5.78
White/Black Ratio	2.40	1.21	2.90
C. Applications to All Postings within 30 Minute Commute Time (within NJ/NYC)			
New Jersey			
Black	623	11.22%	70.00
White	474	15.55%	73.71
White/Black Ratio	0.76	1.39	1.05
New York City			
Black	167	8.65%	14.45
White	293	10.18%	29.87
White/Black Ratio	1.75	1.18	2.07

Notes: Figures projected, based on our sample distribution of businesses and callbacks, for a counterfactual population in which black and white applicants are geographically distributed to mirror the real population, but otherwise remain identical. They assume that applicants apply to every business in our sample within their Zip Code Tabulation Area [Panel A] and within 15 and 30-minute commutes within NJ or NYC respectively (using driving in NJ or public transit in NYC) [Panels B and C]. The number of applications and callbacks per capita are calculated by reweighting our sample observations by the probability that a person of the applicant's race in that jurisdiction would live within the ZCTA or the commuting-time threshold; weights are deflated by half because we attempted to apply to each business twice for each race. Applications per capita and callbacks per capita are the sum of the weighted observations or callbacks)across the sample (within jurisdiction and race). Callback rates are then calculated arithmetically.