Distinctively Black Names and Educational Outcomes

Daniel Kreisman

Georgia State University

Jonathan Smith

Georgia State University

Names can convey information about race or ethnicity and therefore can be used to discriminate against protected groups; many researchers have demonstrated as much through audit studies. Yet few studies link life outcomes with names using observational data. We use administrative data from over 3 million Black students to ask whether those with more statistically Black names have differential educational outcomes. We find that while test scores, college enrollment, and college completion are negatively correlated with Black names net of background characteristics, this relationship is absent when we compare across siblings within households.

I. Introduction

Centuries of racial discrimination and a lasting legacy of slavery and segregation in the United States have resulted in racial disparities in nearly all measures of schooling. This is a persistent and concerning reality that

The conclusions of this paper are the authors' alone and do not reflect those of any affiliated institutions or data providers. We thank the many colleagues who have commented on earlier versions of this paper. All errors are our own. Replication files are available in a zip file. This paper was edited by James Heckman.

Electronically published March 3, 2023

Journal of Political Economy, volume 131, number 4, April 2023.

https://doi.org/10.1086/722093

underscores a continuing need for policies that both address racial discrimination and foster equity in schooling.

We add to our understanding of this reality by asking whether names, often a signal of race or ethnicity, play a role. If prior work is any guide, one might expect that they do. A host of correspondence, or audit, studies show differences in everything from job application callbacks (Bertrand and Mullainathan 2004; Pager, Bonikowski, and Western 2009; Gaddis 2015; Kline, Rose, and Walters 2021) to housing applications (Page 1995; Ewens, Tomlin, and Wang 2014), to response rates from local public service providers (Giulietti, Tonin, and Vlassopoulos 2019), to recommendations for Advanced Placement coursework (Francis, de Oliveira, and Dimmitt 2019) for those with first names that clearly identify the applicant as Black.[1]

While these studies might prelude differences in long-run outcomes for individuals with statistically Black names, the limited evidence connecting names with outcomes in observational data fails to find negative effects. For example, Cook, Logan, and Parman (2016) find that African American men who had distinctively Black names lived approximately 1 year longer than other African American males in the late nineteenth and early twentieth centuries.[2] Aura and Hess (2010), using the General Social Survey, find that neither family income nor completed schooling are lower for Black respondents with statistically Black names, net of a limited set of controls. Notably, using data from California, Fryer and Levitt (2004) find that Black women with names that are more common among Black Americans are born under less advantageous circumstances but ultimately have long-run outcomes similar to those for women with names more common among White Americans, in terms of completed schooling or age at first birth, for example. The following can be viewed as a replication and extension of that study in which we reach similar conclusions, with some caveats.

We make use of administrative data to explore whether test scores and long-run educational outcomes vary among Black students with the statistical degree to which a name is Black or White. This might result from discrimination on names in schools or from discrimination on names in other domains of life that are manifested in differential schooling outcomes. Our data come from records for over 15 million Black and White exam takers in the high school graduating cohorts of 2004–14 that are linked with college enrollment and completion records, allowing us to observe educational outcomes for a large proportion of the US college-going population. Where Fryer and Levitt (2004) observe differences in years of schooling only for

[1] See Quillian et al. (2017), Baert (2018), and Gaddis (2018) for excellent overviews of audit studies. See Lieberson and Bell (1992) and Lieberson and Mikelson (1995) for sociological work on naming conventions, race, and class.

[2] Cook, Logan, and Parman (2014) and Cook, Parman, and Logan (2021) further consider the origins of distinctively Black names in the early twentieth century.

women who were born in California and who later gave birth in California while in their early 20s, we observe several test scores and college outcomes, along with which high schools students attended and measures of their parents' income and education. Our sample also has far wider geographic variation and spans a larger swath of the academic achievement distribution, allowing us to observe what role geography and academic ability play. We also estimate effects of names on educational outcomes, net of household differences, through a large sample of siblings. Related to our work on names in an educational context, Foster (2008) uses administrative data from one university and finds no effect of names on course grades. Conversely, Figlio (2005), who uses data from one school district in Florida, finds negative correlations between names associated with low socioeconomic status—and to a lesser degree, names that are more common among Black students—and test scores, even among siblings.[3]

We find that Black students with more statistically Black names have lower exam scores, and rates of college enrollment and graduation, than Black students with less statistically Black names. To illustrate, figure 1 plots standardized exam scores (the PSAT [Preliminary SAT])[4] for over 3 million Black students by a Black name index (BNI), which ranges from 0 (a name found only among White students) to 1 (a name found only among Black students). In raw means, Black students with the most statistically Black names perform nearly 0.6 standard deviations lower, on average, than their Black peers with the least statistically Black names. When we compare among peers from similar households, as measured by student-reported parental income and education, who attended the same high school (or who lived in the same ZIP code) in the same year, Black students with the most statistically Black names perform 0.2 standard deviations lower than those with the least. That is, of a raw 0.6–standard deviation disparity, only two-thirds can be explained away by observable household factors, neighborhood, or high school effects. Further, net of these same factors, Black students with more statistically Black names are less likely to either attend or graduate from college. Moving across the full BNI distribution corresponds to a 3.5% decrease in college enrollment and a 16% drop in the likelihood of college completion. These results run counter to the existing literature, which finds no relationship net of controls, highlighting the value of a large and diverse sample.

When we turn to our sample of over 720,000 Black siblings, allowing us to compare across students with different names in the same household,

[3] Figlio (2005) uses several definitions of a "Black name" that are different from those in other papers in the literature.

[4] The PSAT is a national standardized exam, often considered practice for the SAT, a college entrance exam.

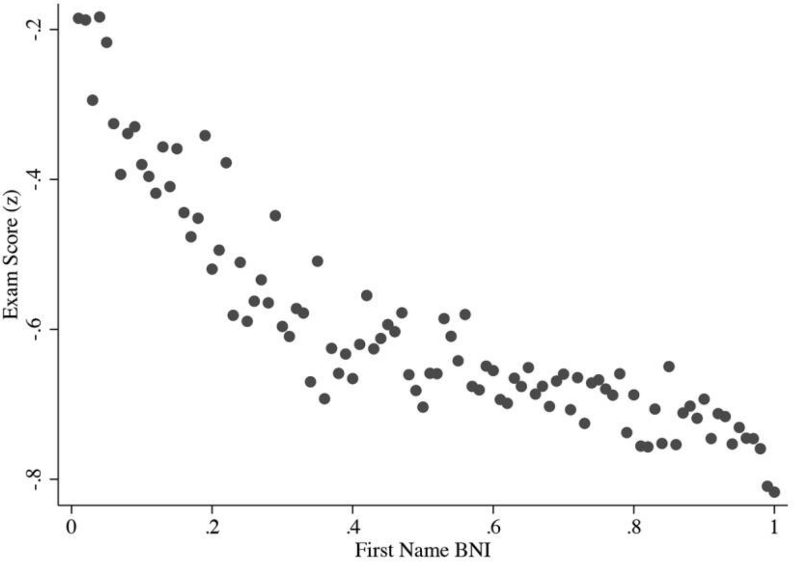

FIG. 1.—BNI and exam scores. Sample is Black students with standardized exam scores in the graduating high school cohorts 2004–14. BNI ranges from 0 to 1, with BNI = 1 for first names used only by Black students.

we find that the relationship between names and educational outcomes disappears. This is not because our sibling sample is selectively different or because there exists insufficient variation in names or test scores within households. Rather, results suggest that unobserved factors captured by household fixed effects, beyond what we can observe in our data, explain the initial negative relationship, potential explanations for which we discuss at length.

There is, of course, much we cannot rule out. For example, we cannot rule out sibling spillover effects within the household or differential treatment across names in other aspects of life, such as contact with law enforcement or in applications for jobs, that might affect schooling outcomes at the household level. Likewise, our results do little to explain the many results in the literature documenting discrimination across names outside of education. Rather, our results might highlight the fundamental difference between our study and audits—that teachers and others in the education system know their students and that names reveal little about race, unlike audit studies, where names are a primary indicator. Thus, we caution that these results do not provide evidence contrary to discriminatory practices against Black or other students from protected classes, in schooling or elsewhere. That discrimination exists in America is an unassailable fact. We find only that educational outcomes that vary with names across households do not do so within households.

We conclude our analysis by contextualizing these results. We do so first with respect to Fryer and Levitt (2004); while we ultimately reach a similar conclusion, differences in interim results merit consideration. Second, in the spirit of that paper, we discuss potential implications for audit studies. Finally, we discuss limitations, policy implications, and potential directions for future work.

II. Data

The majority of our analysis focuses on the set of 3 million Black students from high school graduating cohorts 2004–14 who took the PSAT/National Merit Scholarship Qualifying Test (henceforth PSAT). The PSAT is commonly considered a practice exam for the SAT, which we observe for a smaller number of students.[5] We also make use of the 12 million White test takers to define a BNI and for supplemental analyses in an appendix. Long-run educational outcomes come from a link to the National Student Clearinghouse, which contains information on whether and where students enroll in college for approximately 94% of all college students in the United States, in addition to whether they earned a college degree.[6]

Exam registration requires filling out basic information, including first and last names, race/ethnicity, and gender. Students have the option to self identify as White, Black, Hispanic, Asian, Native American and Pacific Islander, or other.[7] Students (and sometimes the schools themselves) also report their high school and ZIP code. We make use of a sibling identifier that relies on home address, high school, and last name.[8] For the smaller subset of students who take the SAT, we have a categorical assessments of parental income and education, both of which are student reported.

Table 1 shows summary statistics for the over 3 million Black students who took the PSAT and constitute our main analytic sample.[9] A large share of our PSAT sample have missing parental income and/or education,

[5] During the sample period, the PSAT was offered once a year, most frequently at students' high schools. The cost of the exam can range from free to approximately $20, and students can and often do take the exam multiple times.

[6] The most notable deficiency is for-profit college enrollment.

[7] Race/ethnicity is missing for 2.6% of the test takers, and another 3.3% are of "other" race/ethnicity.

[8] Goodman et al. (2015) detail the process and the conservative approach used that will not find all siblings, since addresses can change and misspellings can occur.

[9] Table A1 displays summary statistics for Black and White students.

TABLE 1
Summary Statistics (Black Students; 3,300,428 Observations)

Parent/Household	Mean (SD)	Student	Mean (SD)
Parent highest education:		Male	.46 (.50)
No college	.13 (.33)	PSAT (*Z*-score)	−.70 (.81)
Some college	.15 (.36)	SAT (*Z*-score)	−.75 (.86)
BA+	.18 (.38)	Enrolled in any college	.69 (.46)
Education missing	.54 (.50)	Enrolled in 4-year college	.43 (.50)
Household income:		College degree	.16 (.36)
<$50,000	.19 (.40)	4-year college degree	.13 (.33)
$50,000–$100,000	.08 (.28)	BNI (first name)	.72 (.29)
≥$100,000	.03 (.17)	BNI (last name)	.72 (.20)
Income missing	.69 (.46)	Name count (000s)	25.9 (50.5)

Note.—Sample is Black students who took the PSAT and graduated high school between 2004 and 2014. Parental education is higher of either parent. Some later cohorts have not yet had enough time to earn a college degree. Table A1 shows statistics for the full sample of Black and White students.

though we show that limiting to those with nonmissing values or using the SAT does not meaningfully alter results. The average PSAT exam *Z*-score for Black students, normalized over all Black and White test takers, is −0.7, evidence of a well-documented difference in test scores across Black and White students. A similar story is true across all national exams, including the SAT, NAEP (National Assessment of Educational Progress), and ACT (American College Test).[10] Sixty-nine percent of Black students in our sample enroll in any college, and 43% first enroll in a 4-year college. These numbers are higher than national averages because those students who take these exams are more likely to enroll in college than those who do not. Roughly 16% of Black students receive a degree. These numbers are somewhat lower than the national averages simply because the later part of the sample are still enrolled in college (whom we exclude from analyses on college graduation).

III. Correlates of Statistically Black Names

We measure the statistical relationship between a name and whether someone is Black using a BNI—the same method as Fryer and Levitt (2004). The index measures the relative commonality of a name among Black and White individuals. This is different than a Black-“sounding” name insofar as individuals may act differently toward names they believe are Black

[10] See NAEP’s “Achievement Gap” reports and ACT’s “The Condition of College and Career Readiness” reports.

as opposed to those that statistically are.[11] BNI for first name j is calculated as follows:

$$\mathrm{BNI}_{\mathrm{name}_j} = \frac{\Pr(\mathrm{name}_j|\mathrm{Black})}{\Pr(\mathrm{name}_j|\mathrm{Black}) + \Pr(\mathrm{name}_j|\mathrm{White})}. \tag{1}$$

The numerator is the probability of observing name j among Black test takers, and the denominator is the sum of that probability and the likelihood of observing the same name among White test takers. By construction, the index ranges from 0 to 1. BNI equals 0.5 when a name is equally common among White and Black students. If Black students are twice as likely to have a particular name j than White students, BNI equals 0.67 $(0.2/(0.2 + 0.1))$; if they are half as likely, BNI equals 0.33. We calculate this measure separately for first and last names.

We construct BNI from the sample of all 19 million Black and White students who took any of the three College Board exams (PSAT, SAT, and Advanced Placement). This sample is nearly identical to the sample of 15 million Black and White PSAT takers in table A1 but provides a larger sample to construct the index. White students have an average first-name BNI of 0.28, while Black students' average BNI is 0.72. This is nearly identical to the average, 0.71, for Black mothers of children born between 1989 and 2000 in California in data used by Fryer and Levitt. Last-name BNI is nearly identical to first-name BNI, though we note an important distinction: whereas first names are a choice that may reflect tastes and preferences, last names are largely hereditary and may therefore may reflect long-standing racial animus toward Black Americans over many generations.[12]

Figure 2 plots the entire distribution of first-name BNI for Black and White students. This figure provides three pieces of information. First, there exists common support for Black and White students at nearly all points in the distribution, 0 and 1 notwithstanding by definition. Second, there are mass points for common names, many of which are shared for Black and White students in the middle of the distribution. Third, a much larger share of Black students have uniquely Black names than White students, an issue we explore in sensitivity tests with no change to results.

Table 2 investigates characteristics associated with statistically Black names by regressing first-name BNI on gender, parental income and education, cohort, and first-name count to control for uniqueness. We add ZIP code fixed effects in the second column to show differences among students

[11] Gaddis (2017) shows that names given by more educated Black mothers are less likely to be perceived as Black than names given by less educated Black mothers from a survey study.

[12] Darolia et al. (2016) conduct a correspondence study in which they use race-neutral first names and attempt to signal race using only last names. They find no differences in callbacks.

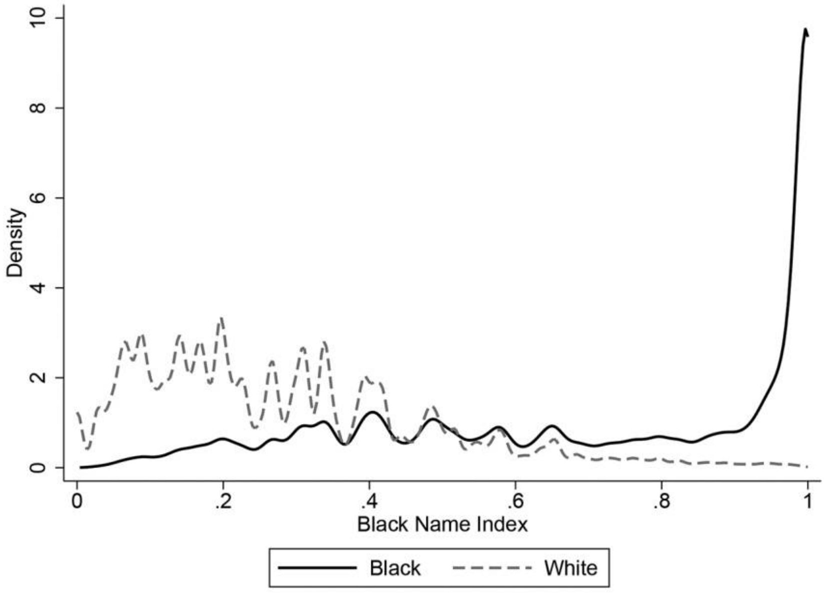

FIG. 2.—Distribution of BNI by race. Sample is all Black and White students who interacted with the College Board in the graduating high school cohorts 2004–14. BNI ranges from 0 to 1, with BNI = 1 for first names found only among Black students.

within a relatively small geography (results are the same using high school fixed effects, which is a larger unit). Black female students have higher BNIs, on average, than Black males and are more likely to have unique names. Students from families with more schooling and higher incomes have lower BNIs. The coefficient on year, which we enter here linearly to detect trends, is small in magnitude but suggests a slight increase in BNI over time. The coefficient on first-name count is negative, which implies that more common names, which can appear tens of thousands of times in the sample, have marginally lower BNIs. Comparing across the two columns, with and without ZIP code fixed effects, reveals that even within a relatively small geography, parental education and income are strong predictors of BNI. Coefficients on those predictors are diminished by only about one-quarter when we absorb unobserved location factors.

IV. Names and Educational Outcomes

Our main analysis begins by asking whether educational outcomes vary with names among Black students. We do so in three parts. First, we focus on the relationship between test scores and BNI, asking whether the large raw differentials we observe in figure 1 can be explained by observable

TABLE 2
Correlates of BNI (Black Students; 3,300,428 Observations)

	(1)	(2)
Male	−.013***	−.013***
	(.000)	(.000)
Parent some college	−.020***	−.015***
	(.001)	(.001)
Parent BA	−.059***	−.045***
	(.001)	(.001)
Parent income \$50,000–\$100,000	−.029***	−.022***
	(.001)	(.001)
Parent income >\$100,000	−.064***	−.047***
	(.001)	(.001)
Year (linear)	.004***	.004***
	(.000)	(.000)
Name count (000s)	−.003***	−.003***
	(.000)	(.000)
ZIP code fixed effects		X
Adjusted R^2	.354	.373

Note.—The dependent variable is first-name BNI, which ranges from 0 to 1. Model is estimated by ordinary least squares. Sample includes all Black students who took the PSAT. Omitted categories are parent with less than (any) college and parental income less than \$50,000. Indicators for missing parental education and income are included. Standard errors (in parentheses) are clustered at the ZIP code level.
*** $p < .01$.

household characteristics and neighborhood effects. We then turn to our sample of siblings to isolate potential name effects net of fixed household factors. Finally, we examine long-run educational outcomes in the same fashion.

A. BNI and Test Scores

Our analysis relies on the following regression equation:

$$Y_i = \alpha + \beta \mathrm{BNI}_i + X_i'\Gamma + \tau_t + \phi_s + \varepsilon_i, \tag{2}$$

where Y_i is an educational outcome, for example PSAT score, for individual i in high school cohort t in ZIP code s. In our sparsest model, we include only first- and last-name BNIs, an indicator for students identified as male, first-name count, and cohort fixed effects, τ_t. We then add categorical parental income and education, X_i. We then include ZIP code fixed effects, ϕ_s. In all cases, we cluster standard errors at the ZIP code level. We are primarily interested in β, the coefficient on first-name BNI, which informs us whether students with names that are more common among Black individuals have different test scores than those with less statistically Black names, conditional on those observable characteristics.

Column 1 of table 3 shows a strong negative relationship between first-name BNI and test scores. Accounting only for high school cohort, name

TABLE 3
Black Names and Test Score Differences (PSAT) within and across Households

	All Black Students			Black Siblings		
	(1)	(2)	(3)	(4)	(5)	(6)
First-name BNI	−.406***	−.285***	−.189***	−.190***	−.005	−.154***
	(.004)	(.003)	(.002)	(.004)	(.007)	(.004)
Sibling's first-name BNI						−.148***
						(.003)
Household fixed effects					X	
Birth order				X	X	
ZIP code fixed effects			X	X		X
Parent controls		X	X	X		X
Last-name BNI	X	X	X	X		X
Male, count	X	X	X	X	X	X
Cohort fixed effects	X	X	X	X	X	X
Observations	3,300,428	3,300,428	3,300,428	720,592	720,592	720,592
Adjusted R^2	.021	.179	.261	.280	.500	.282

Note.—The dependent variable is PSAT Z-score, normalized to all Black and White test takers. Sample is all self-identified Black PSAT test takers in the 2004–14 graduating high school cohorts. Parental controls include categorical income and highest education, including indicators for missing values. If multiple siblings, the sibling first-name BNI is the average of all other siblings' first-name BNIs. Standard errors (in parentheses) are clustered at the ZIP code level.

*** $p < .01$.

count, gender, and last-name BNI, column 1 shows that going from the least to most statistically Black first name among Black students corresponds to just over a 0.40–standard deviation decrease in scores. Column 2 adds parental education and categorical household income. While this attenuates the negative relationship, we still find nearly a 0.3–standard deviation difference in PSAT score between Black students with the most and least statistically Black names. Column 3 adds ZIP code fixed effects, allowing us to compare across Black students from similar socioeconomic backgrounds who live in the same ZIP code.[13] The coefficient on first-name BNI suggests that even within this narrow comparison group, Black students with the most statistically identifiable Black names perform nearly 0.2 standard deviations lower on our test measure than those with the least statistically Black names.

In table A2, we show a comprehensive set of robustness tests, each of which shows that results are not sensitive to a host of choices. These include (coefficient on first-name BNI in parentheses) dropping students with missing education or income (−0.177); dropping those with unique names (−0.187); limiting to students where 90% or more of the high school took the PSAT (−0.187); using only states where SAT is the dominant college entrance exam (−0.174); using only self-reported males (−0.160) or only females (−0.210); changing the dependent variable to SAT (−0.220),

[13] High school fixed effects yield equivalent results.

PSAT math (−0.145), or PSAT verbal (−0.208); and using ZIP code–by-cohort fixed effects (−0.182). We highlight that the relationship is stronger for female students and is larger when we use the SAT, noting that correlations with parental income and education are stronger for female students and that differences are larger among those at the top end of the PSAT test score distribution, which we discuss further below.

B. *Isolating Name Effects within Households*

We next seek to isolate effects of first names within households. We limit our sample to approximately 720,000 Black siblings and reestimate equation (2) with a household fixed effect. Now, X_i includes only reported gender, sibling order, and commonality of first name, as parental income, education, geographic location, and last-name BNI do not vary within household; β now shows test score differences across students in the same household with different first names.

Column 4 of table 3 reestimates column 3 limited to our sibling sample, demonstrating that the coefficient on first-name BNI is nearly identical. In column 5, we add the household fixed effect. The resulting coefficient is a precisely estimated zero. This result suggests that within households, test scores do not vary with names. Yet, we do find score differences by name across otherwise observationally equivalent households. To illustrate, in column 6 we show that a student's sibling's BNI (or the average of her siblings' BNIs) is equally predictive of her own test scores as her own BNI.[14] This suggests that households with higher average BNIs have lower test scores, on average.

One potential explanation for null effects within the household is that we are simply working with less variation than across households. While this is true, there still exists significant variation. For example, the mean range (maximum less minimum) of first-name BNI within Black households is 0.29 (SD = 0.25). This mean range, 0.29, is equivalent to 1 standard deviation of the total variation in the full sample of Black students. For PSAT score, the same average within-household range is 0.60 (SD = 0.57), where 1 standard deviation of PSAT is approximately equal to 1 in the full sample, and the mean Black-White gap in PSAT is 0.70. This demonstrates that while both test scores and first-name BNI are highly correlated within households, there still exists considerable within-household variation in each. For example, in our household fixed effects specification, we can still detect a precise gender difference of −0.07 on the coefficient for male.

[14] Results are similar if we choose one random sibling in each household as the main observation for concerns about a reflection problem. Results are also similar, but of about two-thirds the magnitude, if we use only households with exactly two siblings (who are not twins) and add a full set of controls for the sibling's gender, test scores, cohort, and commonness of name.

A second possible explanation could be spillover effects, where one sibling faces discrimination on his or her name and families are forced to compensate by directing resources toward that sibling at the expense of others. This would result in large across-household differences and potentially small or no difference within them, which is consistent with our data. Similarly, since siblings typically attend the same school, teachers' perceptions of younger siblings might be influenced by their interaction with their older sibling. While this effect is difficult to test, table A2 shows that our results hold regardless of whether the oldest or youngest sibling has a higher BNI, which might be one indication of spillovers. A third consideration could be that families face discrimination on names outside of schooling. Fryer and Levitt (2004) show that parent and child BNIs are strongly correlated. If parents face discrimination on names, for example in job applications, this would affect both children equally, resulting in across- but not within-household differences. Unfortunately, without parent BNI, we have no way to test this.

A final consideration is the role of schools. In figure 3, we ask whether the relationship varies with the share of students in a high school who are Black. In figure 3*A*, we plot coefficients from our main regression separately by the share of a student's high school that is Black (in bins of 0.1). We find that the across-household relationship is roughly twice as large for students in schools where 20% of students are Black, compared with those where 20% of the population is not Black. Yet, doing this for a within-household estimate shows no relationship, though standard errors are large. Thus, while we uncover heterogeneity in the relationship across households, we find no evidence that the racial composition of schools masks underlying heterogeneity of the within-household relationship.[15]

C. *Long-Run Educational Outcomes*

We extend this analyses to long-run educational outcomes regarding college. This tests whether the relationship extends to more consequential educational outcomes and allows us to replicate the line of inquiry in Fryer and Levitt (2004), who find no impacts of names on educational attainment. For these analyses, we restrict the sample in two ways—first, we focus on siblings, and second, we use students in cohorts we observe at least 6 years after leaving high school to allow sufficient time to attend and graduate college.

The first column of coefficients in table 4 shows estimates from equation (2) for college outcomes. We begin again by replicating our main PSAT result for this subsample of siblings to demonstrate that limiting

[15] See app. D for a protracted discussion and analysis on the "household effect" and the relative role of neighborhoods, families, and schools.

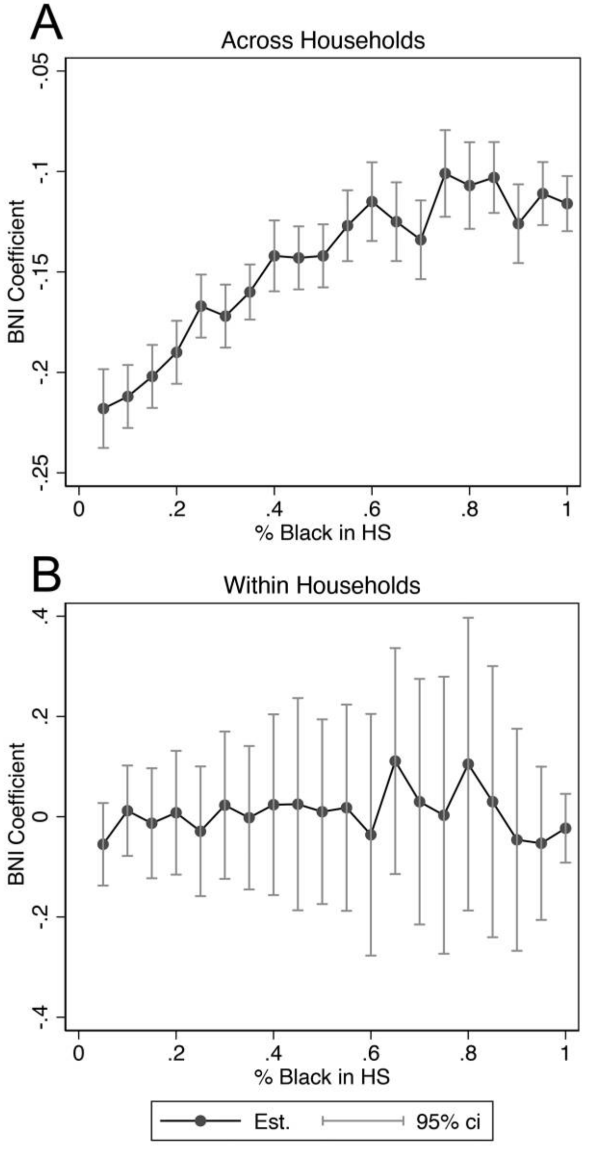

FIG. 3.—BNI and test scores by percentage of Black students in high school. Figures plot coefficients for separate regressions by percentage of Black students in a high school (in bins of 0.1) for the full model in equation (2) (*A*) and a model with household fixed effects (*B*). Percentage of Black students in high school comes from National Center for Educational Statistics's Common Core of Data and Private School Survey. ci = confidence interval.

TABLE 4
BNI and Long-Run Educational Outcomes across and within Households (444,786 Observations)

Outcome	Across-Household βBNI (SE)	Within-Household βBNI (SE)	Mean of Dependent Variable
PSAT	−.192*** (.005)	−.010 (.012)	−.53
Enrolled in any college	−.027*** (.003)	.003 (.007)	.78
Enrolled in 4-year college	−.045*** (.003)	−.002 (.008)	.52
College degree	−.052*** (.003)	−.004 (.008)	.33
4-year college degree	−.051*** (.003)	−.004 (.007)	.29

Note.—Sample is all Black students whom we observe 6+ years after exiting high school and with a sibling. The across-household regressions include all family background controls, ZIP code and cohort fixed effects, gender, name count, and birth order. Within-household regressions exclude ZIP code and parental controls. Standard errors are clustered on ZIP codes.
*** $p < .01$.

to later cohorts is not driving results. The remaining rows in this column show meaningful differences across Black students with the most and least statistically Black names. For example, moving across the BNI distribution is associated with a 2.7 percentage point (3.5%) decrease in the likelihood of enrolling in college and a 5.2 percentage point (16%) decrease in the likelihood of graduating. These relationships are more pronounced for 4-year colleges. The second column of coefficients shows that, like test scores, these differences are zero within households.

V. Discussion, Limitations, and Conclusions

In this final section, we reconcile our results with those of Fryer and Levitt (2004) and with the resume audit literature; summary remarks follow.

A. *Reconciling with Fryer and Levitt*

Ultimately, we find that while educational outcomes vary with names across observationally similar households, we do not observe an effect of names within them. While this result is largely consistent with conclusions reached by Fryer and Levitt (2004), meaningful differences exist. We find a large unconditional relationship between BNI and educational outcomes and further find that observable household factors cannot explain away the result. Fryer and Levitt (2004) find no unconditional relationship between BNI and education (years of schooling) and find that including household controls in their analysis has little impact on that outcome.[16] Across a range of

[16] See table III, col. 4, top panel in Fryer and Levitt (2004). Estimates for going from 0 to 100 on BNI for education are −0.0006 (SE: 0.0005) without controls and 0.0008 (SE: 0.0005) with controls.

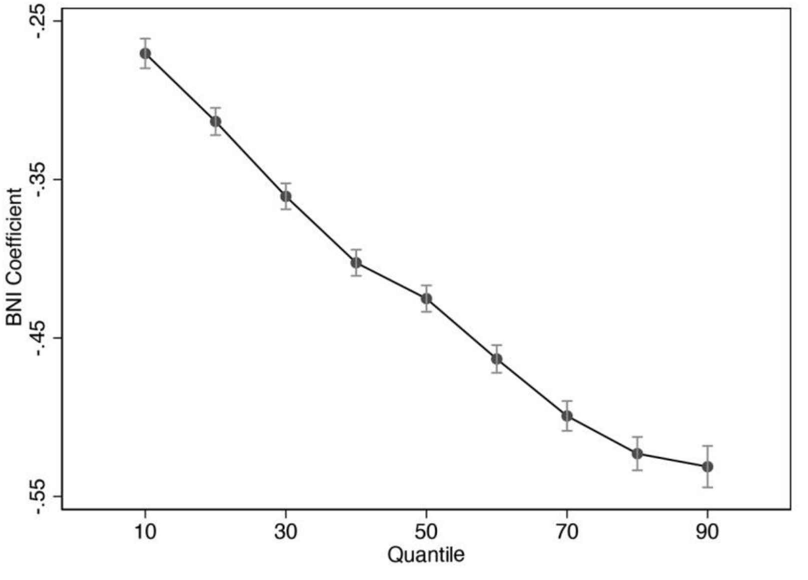

FIG. 4.—Quantile regression of test scores on BNI. Figure plots coefficients from an unconditional quantile regression of standardized PSAT *Z*-score on BNI. The 99% confidence intervals are plotted from standard errors clustered at the ZIP code level.

outcomes, their raw differences are relatively small, and after controls for childhood circumstances are included, nearly all are zero.

One potential explanation for differences across studies is that one or both of our data sets are from a selected sample. Fryer and Levitt observe a sample of women who gave birth at relatively young ages and completed less schooling than those in our sample; average years of schooling in their sample is only 12.4 years, and age at first birth is just over 19. In our sample, more than 78% of students we observe for 6 years after high school attend at least some college, and one in three graduate. To illustrate how sample selection might affect estimates, figure 4 plots results from an unconditional quantile regression of BNI on test scores. The figure shows that the relationship is roughly twice as large at the 90th percentile as at the 10th. This implies that the unconditional relationship between names and outcomes is stronger for those in the right-hand (higher-score) portion of the PSAT score distribution, potentially explaining differences across studies.[17]

[17] We also account for some selection issues in our robustness checks by using subsamples where nearly all students in a high school (or most in a state) took the PSAT. Results are similar and are shown in table A2.

A second explanation is that Fryer and Levitt have different controls. They have circumstances at birth, where we observe only circumstances in high school. If early-life circumstances are more predictive of teenage schooling outcomes than those in the teenage years, our household effect could simply be picking up what we would have observed with measures of early-life circumstances. Additionally, education is self-reported on birth certificates, implying that classical measurement error in Fryer and Levitt could further bias their estimates toward zero. We cannot say for certain whether these facts are the driver of differences between results or whether simply observing different samples at different times led to these results.

B. Implications for Resume Audit Studies

By sending identical resumes where only names differ, resume audit studies show dramatic differences in callback rates. These studies make clear that applications with Black names face discrimination and are far less likely to receive callbacks.[18] By showing that names are correlated with other circumstances, Fryer and Levitt (2004) argue that it becomes difficult to determine the degree to which employers are simply using names as a signal of race or are making decisions based on other factors correlated with names in addition to race. It is important to note that both are illegal, and both directly affect communities of color. Those authors note that one "cannot directly test between these competing hypotheses. . . . [because their] data set lacks clean measures of productivity. . . . [and they] do not have in [their] data all of the information on a resume" (Fryer and Levitt 2004, 798–99).

We consider the latter of these with a regression that includes data from both Black and White students and includes what employers might or might not observe on resumes (see app. B for a longer discussion and full results). For example, we find that conditional on what employers might observe on resumes, including gender, which high school and college a student attended, and his or her college major if a graduate, moving across the full BNI distribution corresponds to a -0.56 standard deviation difference in test scores. We also show that this result holds for a subsample of our data matched to actual resumes scraped from an online jobs board and holds even if we include information about skills and work experience on those resumes.[19] This aligns with Fryer and Levitt's concern that names are correlated with attributes net of what employers might observe on resumes.

[18] See also Heckman (1998), Pager (2007), and Neumark (2012, 2018) for a discussion of what can and cannot be learned from audit studies.

[19] See app. C for analysis that uses a subsample of actual resumes, as detailed in Kreisman, Smith, and Arifin (2021).

How can these results contextualize what we have learned from resume audit studies? One possibility, noted by Fryer and Levitt, is that employers are discriminating on race and are using names to infer it at the initial screening. Those same discriminatory employers might then further screen out Black applicants without statistically Black names when they learn of race later in the process. This would lead to large audit study effects, similar long-run outcomes across names, and a wide Black-White gap in observational data. Another possibility is that employers may be acting on information other than race in names. For this to explain results in existing studies, one would have to consider (1) that employers know the correlation between names and productivity and (2) that residual differences across names, conditional on what is on a resume, are predictive of workers' productivity. We take each of these in turn.

First, it is unlikely that employers can accurately relate first names to measures of productivity. This theory might suggest that larger organizations, which have larger samples to learn from, would statistically discriminate (on names) more often. Banerjee, Reitz, and Oreopoulos (2018) conduct a test of this and find precisely the opposite. More, in a recent study of over 100 large employers, Kline, Rose, and Walters (2021) find that statistically Black names reduce contact rates by 2.1 percentage points (9%). Importantly, they find that high rates of discriminatory behavior are concentrated in a small number of industries for which information about test scores or childhood circumstances might add limited value relative to others (auto services, sales, and retail). Along these lines, in our pooled regression of Black and White students described above, controlling only for gender, cohort, and name, the BNI coefficient is -0.90. It decreases to -0.56 when college and major are included, suggesting that information about completed schooling should reduce the callback gap. Yet Bertrand and Mullainathan (2004) found that adding more skills or schooling to resumes did not decrease the callback rate. Similarly, Gaddis (2015) found that listing an "elite" university on a resume, as opposed to a less selective one, did not close the application contact rate. In fact, Black resumes from elite universities only yielded callback rates similar to those for White applicants from less selective universities. If employers were using names to infer productivity, one might expect callback differences to decline with skills.

Second, it is unclear to what degree negative correlations with test scores, or childhood circumstances, are useful proxies for expected productivity, conditional on resume information. Residual negative differences in test scores or childhood circumstances, for example, might in fact imply that students who had fewer advantages in childhood, or who had lower test scores going into the same college but who still graduated with the same degree and major, put in more effort or had higher noncognitive skills that are in fact highly valuable in the labor market, as in Heckman, Stixrud, and Urzua (2006) and Deming (2017). This would imply an even higher degree of racial animus in hiring than existing audit studies uncover. More, we note that results here provide no explanation for the many results in the literature documenting discrimination on names in contexts outside of schooling or work, where test scores should play no role at all, for example, housing applications (Page 1995; Ewens, Tomlin, and Wang 2014) or local public service provision (Giulietti, Tonin, and Vlassopoulos 2019), among many others.

In short, our results suggest that interpreting audit studies is no simple task. By taking names at the extremes of the Black and White naming distribution, audits are not comparing applications of modal Black and White applicants. They are in fact comparing individuals who, on average, are quite different in both schooling and measured academic ability, even conditional on observable characteristics. Those with the most statistically Black and White names rarely have similar educational backgrounds and outcomes, likely a product of a very unequal educational system. This can make it difficult to clearly determine what factors employers respond to on resumes.

C. *Limitations and Concluding Remarks*

We document substantial educational disparities across Black students according to their first names. These disparities, ranging from test scores to college completion, are concerning. Yet within households, Black students with more statistically Black first names have test scores and long-run outcomes similar to those for their siblings with names that are less statistically Black.

We caution that there is much we cannot rule out. For example, it is possible that the lack of differentials we observe across names within households is the result of a fundamental difference between our study and audit studies—that teachers and others in the education system know their students and likely do not judge them by their names. This is not true in the labor market, where discrimination on names is well documented. Additionally, there could exist discrimination on names in schooling, leading families to compensate by directing additional resources toward children with more statistically Black names at the expense of their siblings. Further, we note that our main outcome measure is a standardized test. If teachers or others in the educational system discriminate on names, results might be more prominent in classroom-level setting, such as grades or discipline, noting that Foster (2008) does not find grading effects in a similar effort. We also cannot rule out discrimination in a multitude of other aspects of life, for example, in contact with law enforcement or applications for jobs, that might affect schooling outcomes at the household level. Likewise, our results do little to explain the many results in the literature documenting discrimination across names outside of education and employment.

In sum, we caution strongly that these results, taken together or independently, do not in any way invalidate or provide evidence contrary to the existence of discriminatory practices against Black or other minority students, job seekers, or workers. That discrimination exists, toward not only Black Americans but many other minority groups, is a well-established fact. The necessity of workplace-discrimination laws in America, including the continuing importance of the Civil Rights Act of 1964, is but one piece of evidence. There are countless others. Rather, our results are intended to add a new insight into the complicated relationship between race and schooling in America.

To this end, results here open the door for future inquiry and potentially for policy. Concerning the former, the next logical step is to examine long-run consequences, particularly wages and employment over the life cycle. Increasing access to data that include names, and potentially test scores, linked with administrative earnings records is an obvious candidate. The ability to observe siblings will help. Concerning policy, our results suggest that employers might make better hiring decisions with more personal information and personal interaction than with simply screening resumes, which is increasingly common with online jobs boards and has the potential to exacerbate racial discrimination in hiring. In a step further, Onwuachi-Willig and Barnes (2005) argue that Title VII protections should extend to cover discrimination against names—what they call "proxy discrimination"—even regardless of an applicant's race. Further, one might consider anonymous resumes. This would inhibit employers' ability to discriminate on names, though other information on resumes might signal race. It turns out that an experiment in France conducted precisely this exercise (Behaghel, Crépon, and Le Barbanchon 2015). Unfortunately, the researchers found that anonymous resumes decreased the rate at which firms hired minority candidates, by over 10 percentage points in interviews and nearly 4 percentage points in recruitment. The authors suggest that this made it difficult for employers who favored minority candidates to seek them out. In short, the mechanisms that create the world we observe are not necessarily straightforward, nor are the solutions.

References

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

Baert, Stijn. 2018. "Hiring Discrimination: An Overview of (Almost) All Correspondence Experiments since 2005." In *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*, edited by S. Michael Gaddis, 63–77. Methodos series, vol. 14. Cham: Springer.

Banerjee, Rupa, Jeffrey G. Reitz, and Phil Oreopoulos. 2018. "Do Large Employers Treat Racial Minorities More Fairly? An Analysis of Canadian Field Experiment Data." *Canadian Public Policy* 44 (1): 1–12.

Behaghel, Luc, Bruno Crépon, and Thomas Le Barbanchon. 2015. "Unintended Effects of Anonymous Résumés." *American Econ. J. Appl. Econ*. 7 (3): 1–27.

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

Cook, Lisa D., Trevon D. Logan, and John M. Parman. 2014. "Distinctively Black Names in the American Past." *Explorations Econ. Hist*. 53:64–82.

———. 2016. "The Mortality Consequences of Distinctively Black Names." *Explorations Econ. Hist*. 59:114–25.

Cook, Lisa D., John Parman, and Trevon Logan. 2021. "The Antebellum Roots of Distinctively Black Names." *Hist. Methods* 55 (1): 1–11.

Darolia, Rajeev, Cory Koedel, Paco Martorell, Katie Wilson, and Francisco Perez-Arce. 2016. "Race and Gender Effects on Employer Interest in Job Applicants: New Evidence from a Resume Field Experiment." *Appl. Econ. Letters* 23 (12): 853–56.

Deming, David J. 2017. "The Growing Importance of Social Skills in the Labor Market." *Q.J.E*. 132 (4): 1593–640.

Ewens, Michael, Bryan Tomlin, and Liang Choon Wang. 2014. "Statistical Discrimination or Prejudice? A Large Sample Field Experiment." *Rev. Econ. and Statis*. 96 (1): 119–34.

Figlio, David N. 2005. "Names, Expectations and the Black-White Test Score Gap." Working Paper no. 11195 (March), NBER, Cambridge, MA.

Foster, Gigi. 2008. "Names Will Never Hurt Me: Racially Distinct Names and Identity in the Undergraduate Classroom." *Soc. Sci. Res*. 37 (3): 934–52.

Francis, Dania V., Angela C. M. de Oliveira, and Carey Dimmitt. 2019. "Do School Counselors Exhibit Bias in Recommending Students for Advanced Coursework?" *BE J. Econ. Analysis and Policy* 19 (4): 20180189.

Fryer, Roland G., Jr., and Steven D. Levitt. 2004. "The Causes and Consequences of Distinctively Black Names." *Q.J.E*. 119 (3): 767–805.

Gaddis, S. Michael. 2015. "Discrimination in the Credential Society: An Audit Study of Race and College Selectivity in the Labor Market." *Soc. Forces* 93 (4): 1451–79.

———. 2017. "How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies." *Sociological Sci*. 4:469–89.

———, editor. 2018. *Audit Studies: Behind the Scenes with Theory, Method, and Nuance.* Methodos series, vol. 14. Cham: Springer.

Giulietti, Corrado, Mirco Tonin, and Michael Vlassopoulos. 2019. "Racial Discrimination in Local Public Services: A Field Experiment in the United States." *J. European Econ. Assoc*. 17 (1): 165–204.

Goodman, Joshua, Michael Hurwitz, Jonathan Smith, and Julia Fox. 2015. "The Relationship between Siblings' College Choices: Evidence from One Million SAT-Taking Families." *Econ. Educ. Rev*. 48:75–85.

Heckman, James J. 1998. "Detecting Discrimination." *J. Econ. Perspectives* 12 (2): 101–16.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." *J. Labor Econ*. 24 (3): 411–82.

Kline, Patrick M., Evan K. Rose, and Christopher R. Walters. 2021. "Systemic Discrimination among Large U.S. Employers." Working Paper no. 2021-94, Becker Friedman Inst., Chicago.

Kreisman, Daniel, Jonathan Smith, and Bondi Arifin. 2021. "Labor Market Signaling and the Value of College: Evidence from Resumes and the Truth." *J. Human Resources* https://dx.doi.org/10.3368/jhr.0119-9979R2.

Lieberson, Stanley, and Eleanor O. Bell. 1992. "Children's First Names: An Empirical Study of Social Taste." *American J. Sociology* 98 (3): 511–54.

Lieberson, Stanley, and Kelly S. Mikelson. 1995. "Distinctive African American Names: An Experimental, Historical, and Linguistic Analysis of Innovation." *American Sociological Rev.* 60 (6): 928–46.

Neumark, David. 2012. "Detecting Discrimination in Audit and Correspondence Studies." *J. Human Resources* 47 (4): 1128–157.

———. 2018. "Experimental Research on Labor Market Discrimination." *J. Econ. Literature* 56 (3): 799–866.

Onwuachi-Willig, Angela, and Mario L. Barnes. 2005. "By Any Other Name?: On Being 'Regarded as' Black, and Why Title VII Should Apply Even If Lakisha and Jamal Are White." *Wisconsin Law Rev.* 2005:1283–343.

Page, Marianne. 1995. "Racial and Ethnic Discrimination in Urban Housing Markets: Evidence from a Recent Audit Study." *J. Urban Econ.* 38 (2): 183–206.

Pager, Devah. 2007. "The Use of Field Experiments for Studies of Employment Discrimination: Contributions, Critiques, and Directions for the Future." *Ann. American Acad. Polit. and Soc. Sci.* 609 (1): 104–33.

Pager, Devah, Bart Bonikowski, and Bruce Western. 2009. "Discrimination in a Low-Wage Labor Market: A Field Experiment." *American Sociological Rev.* 74 (5): 777–99.

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." *Proc. Nat. Acad. Sci. USA* 114 (41): 10870–75.