NBER WORKING PAPER SERIES

ALGORITHMIC BIAS AND HISTORICAL INJUSTICE:
RACE AND DIGITAL PROFILING

Abigail Matthew
Amalia R. Miller
Catherine Tucker

Working Paper 32485
http://www.nber.org/papers/w32485

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2024

We are grateful to Ellora Derenoncourt, Ginger Jin, Ben Leyden, Michael Luca, and various conference and seminar participants for helpful feedback and to Yutong Chen for excellent research assistance. We acknowledge financial support from the Bankard Fund for Political Economy at the University of Virginia. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w32485

Algorithmic Bias and Historical Injustice: Race and Digital Profiling
Abigail Matthew, Amalia R. Miller, and Catherine Tucker
NBER Working Paper No. 32485
May 2024
JEL No. J15,J78,K24,M37

ABSTRACT

This paper studies the implications of attempts at "ethnic-affinity" profiling on Facebook that reflects users' engagement with content on Facebook. Profiling by ethnic-affinity is highly correlated with Census estimates of population race by geography. However, more users were profiled as African-American in former slave states relative to the baseline population. This occurs because the targeting algorithm was better at identifying Black users through differentiated engagement with cultural content in these states. This implies that policies restricting the collection of racial identity data will be unsuccessful due to the existence of proxies, and that relying on proxies may introduce troubling biases.

Abigail Matthew
University of Wisconsin-Madison
Madison, WI 53715
abigail.l.matthew@gmail.com

Amalia R. Miller
Department of Economics
University of Virginia
P. O. Box 400182
Charlottesville, VA 22904
and NBER
armiller@virginia.edu

Catherine Tucker
MIT Sloan School of Management
100 Main Street, E62-533
Cambridge, MA 02142
and NBER
cetucker@mit.edu

1 Introduction

A key policy concern arising from the increasing use of machine learning and AI is the potential for intentional discrimination or unintended biases that worsen inequality and harm historically disadvantaged groups (Lee et al., 2019; Rambachan et al., 2020). The EU's landmark 2023 AI Act includes regulation to prevent racial and gender bias in AI systems.[1] In the US, scholars have identified disproportionate risks to Black and African American individuals (Noble, 2018; Allen, 2022), and presidential executive orders have directed federal agencies to combat algorithmic discrimination and develop tools and practices to address AI-related equity and civil rights challenges.[2]

One approach to preventing discrimination is to block or obscure information on protected group membership (Goldin and Rouse, 2000), as in laws restricting information flows about health (Miller and Tucker, 2018) and criminal backgrounds (Doleac and Hansen, 2020; Agan and Starr, 2018). Similarly, the European General Data Protection Regulation (GDPR) provisions treat racial and religious identity as sensitive information and restrict its collection and use.[3] However, limiting the flow of racial information will not be effective at making algorithms race-neutral if the underlying data contain other variables that are highly correlated with race (Allen, 2022; Gillis and Spiess, 2019). Whether hiding racial category information is an effective anti-discrimination strategy depends on whether and how racial information can be inferred using other available data.

This paper addresses these questions with data collected from a setting in which racial identity is inferred by sophisticated algorithms, using records of activity and engagement with content, on Facebook. We collected our data by purchasing a variety of national ad campaigns that were shown over a million times to over 531,000 unique Facebook users in August 2017. During this time, Facebook allowed advertisers to create target audiences based on including or excluding "ethnic affinity" groups such as African American, Asian, and Hispanic. This practice has been criticized (Angwin and Parris, 2016; Angwin et al., 2017) in the press. The company later removed

[1] https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683.

[2] https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/

[3] https://commission.europa.eu/law/law-topic/data-protection/reform/rules-business-and-organisations/legal-grounds-processing-data/sensitive-data/what-personal-data-considered-sensitive_en

advertisers' ability to exclude based on multicultural affinity groups[4] and agreed to eliminate racial targeting.[5] Our data therefore provide unique insight into the operation of ethnic targeting on a large advertising-supported platform that is not otherwise available to external observers.

A key feature of our setting is that Facebook did not collect information from users directly about their racial or ethnic identities. The company built ethnic-affinity groups using onsite activity such as liking, posting or engaging with different types of content. Using reported metrics for our various ad campaigns, we are able to track the distribution of where the ads were shown and clicked across 192 media markets.

Our first result is that ads targeted at different demographic groups were more likely to be shown in areas with higher populations of those subgroups, which is consistent with Facebook's racial affinity categories predicting racial and ethnic group identification. This provides empirical support for the expectation that racial information can be inferred from other correlates on a broad population. It also confirms that hiding a demographic category is not enough to mask identity.

Our second result is that, above and beyond the rates predicted by population shares, ads targeted at Facebook users with Black and African American affinity, but not with other ethnic affinities, were significantly more likely to be shown in states where slavery was legal at the start of the American Civil War. This appears to be driven by variation in the effectiveness of algorithms at identifying Black audience members based on patterns of engagement with content. Facebook is able to predict more users as having a Black or African American affinity in states with this history of slavery because of greater racial differentiation in what cultural material is engaged with in those states.

In contrast to the first finding, which emphasizes the power of machine learning to impute missing classifiers, this finding emphasizes its limitations. Members of a minority group are easiest to identify when they possess characteristics that are distinctive to that group, and not common in the general population. More generally, the more the minority group overlaps with other groups, the less discerning the algorithm can be with proxy variables. Achieving a low false positive rate,

[4](https://www.axios.com/2017/12/15/facebook-disabling-multicultural-affinity-group-advertising-tool-1513307261

[5]https://www.theverge.com/2018/7/24/17609178/facebook-racial-dicrimination-ad-targeting-washington-state-attorney-general-agreement

with high confidence of correctly tagging individuals as group members, may require emphasizing traits (or combinations of traits) that are not representative of the group as a whole.

This means that indirect inference of group membership based on other data, even exceptionally rich data on digital activity, can create another level of bias based on classification. These features are often not intentional expressive choices, but the digital detritus of living in the world, consuming content, shopping, and socializing. The underlying traits used to categorize people are complex and necessarily opaque, which creates challenges for individuals who want to access or correct this information.

Taken together, our results caution policymakers against a simple approach of excluding or masking information on race or other sensitive categories from datasets in the hopes of preventing intentional discrimination. For example, recent rule-making has focused on banning the use of protected characteristics in the targeting of advertising.[6] It is possible that hiding information makes it less salient, or that discrimination can be prevented by raising its costs, but these tradeoffs need to be examined directly, and weighed against the unintentional effects of relying on proxies. Our results suggest that data masking is ineffective at reducing racial discrimination. We explore what happens when direct measures of racial identity are excluded from the data, but firms still wish to segment users on that basis. Group membership is correlated with many individual characteristics, including many that have legitimate non-discriminatory purposes. In the world of big data and machine learning, there is no way to completely obscure an individual's group membership from a firm.

We make contributions to two bodies of literature. The first is the literature on algorithmic fairness which reflects concerns that algorithms used to categorize individuals and predict outcomes may embed or generate biases, and therefore exacerbate social and economic inequality (Rambachan et al., 2020; Cowgill and Tucker, 2020). Prior research examining algorithmic fairness has uncovered systematic differences in predictions by race, even when race is not fed into the model (Obermeyer et al., 2019). This approach to detecting bias is only available to people with access to data inputs and algorithmic outputs, and when the data include a measure of race. Instead, this paper considers

[6] https://www.justice.gov/crt/case/united-states-v-meta-platforms-inc-fka-facebook-inc-sdny

what happens when data on race are not directly available, examining algorithmic predictions of racial group affiliation.

We also contribute to the literature that helps inform algorithmic auditing (Lee et al., 2019; Kallus et al., 2022; Kleinberg et al., 2020), which is becoming becoming central to governmental approaches to algorithmic fairness.[7] This literature emphasizes that omitting information on group membership makes it impossible to audit a process to detect disparate *ex post* outcomes or to adjust it to remove unintentional biases (Lee et al., 2019; Kallus et al., 2022; Kleinberg et al., 2020; Obermeyer et al., 2019). These results support policies, such as the Home Mortgage Disclosure Act, that mandate data collection and reporting around protected categories. Our paper highlights an additional risk. When auditing is based on biased proxies, it can produce inaccurate results. Unequal treatment can remain hidden if the proxy draws from a relatively-advantaged minority population; if the reverse is true, gaps may appear with the proxy that are not present overall.

2 Data and Setting

This study examines data from a series of national ad campaigns, conducted on Meta, then known as Facebook, for one week in August 2017. We focus on Meta because of its coverage in the US in this time period – around 79% of all Americans online used Facebook and 59% of 18-29 year-olds used Instagram.[8] In 2017, Tiktok had not yet emerged as a competitor, meaning that Meta had not yet suffered a decline in its usage by young people, and Facebook and Instagram were the most frequently used social networks by people under the age of 35.[9]

Our campaigns presented visually identical ads (Figure A-1) to groups of targeted users with different "ethnic affinities." Individuals who clicked the ad were directed to the webpage of a federal government program that helps people find pathways into federal jobs. The study focused on younger people, aged 20-29, because advice about career pathways is particularly relevant for people in this age group. This age group also had universally high Facebook and Instagram adoption

[7] https://www.ftc.gov/business-guidance/blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai and https://www.gov.uk/government/publications/findings-from-the-drcf-algorithmic-processing-workstream-spring-2022/auditing-algorithms-the-existing-landscape-role-of-regulators-and-future-outlook

[8] https://www.pewresearch.org/internet/2016/11/11/social-media-update-2016/

[9] https://www.pingdom.com/blog/social-media-in-2017/

rates across racial groups, which allows us to use Census population data to measure demographic and economic characteristics of the Meta audience. We separately targeted younger (aged 20-24) and older (aged 25-29) male and female users.

Our main target categories are by ethnic affinity. Figure A-2 shows the Facebook ads interface that we used to order ad campaigns with this targeting. The available categories were African-American, Asian-American, and Hispanic (with sub-groups for dominant language). We purchased separate ads targeted at three of the ethnic affinity groups (using "All" for Hispanic) and another set targeted at Facebook users not in these three categories.[10]

Meta does not collect data on race when people register for Facebook or Instagram. Therefore, these ethnic affinity targeting categories are inferred. Meta instead informed advertisers that these categories were inferred from users' actions on its website.

Targeting by ethnic affinity can be beneficial if it allows sellers of niche ethnic or racially focused products or services to more efficiently reach potential customers. The ability to target minority group audiences can be particularly valuable for businesses that focus on minority customers, including small minority-owned businesses, who might benefit from Facebook advertising to achieve scale. Another socially beneficial use of targeting could be for affirmative action and diversity initiatives aimed at increasing opportunities for groups that have been historically excluded (Miller and Segal, 2012; Miller, 2017). At the same time, racial targeting, inclusive or exclusive, can be used for discriminatory or predatory purposes, and violate equal protection laws in areas of housing, banking and employment.

Although this was not a job ad, clicking on the link could be beneficial to the viewer if they learn more about available jobs and find a better employment match. This makes racial targeting concerning, because it could exacerbate racial disparities. At the same time, racial targeting of advertisements using proxies (such as attendance of a historically Black university or college (HBCU) or social group membership) is a common strategy for affirmative action plans aimed at improving diversity when broad advertising approaches produce non-diverse pools of applicants.

[10] Given that African-American as an "ethnic affinity group" was being used by Facebook as synonymous with being Black, when referring to Facebook's ethnic affinity groups with the United States, we use 'African American' and 'Black' interchangeably.

Recent evidence suggests that the diversity of those in federal employment has not progressed, and has in some cases regressed.[11]

The data generated by these campaigns is a series of daily measures, for each campaign across 192 media markets. For each market, day and demographic target group, we observe the number of times the ad was shown (impressions), the number of unique viewers (reach), the number of actions taken to engage with our ads (clicks), and the advertising costs (based on clicks, for our campaigns).

We later collected supplemental information on the target audiences using the Audience Insights tool that Facebook made available to advertisers. At the time that we collected this information (August 2019), the demographic targets based on "ethnic affinity" were called "multicultural affinity" but were otherwise unchanged. We used Audience Insights to collect data on what types of content were considered predictive of ethnic affinity. However, it was no longer possible to exclude any of the designated ethnic affinity groups from audiences. By 2022, this targeting was no longer available to Facebook advertisers.

Because these options are no longer available, our data provide unique insight into how Facebook itself categorized users racially.

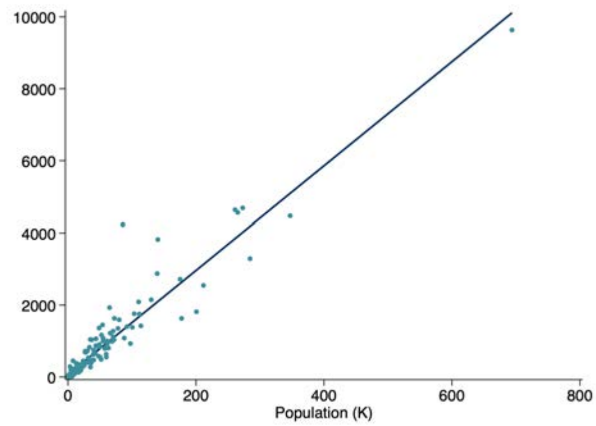
2.1 Geographic Measure of Historic Inequality

We are interested in the question of whether algorithms can reflect historic inequality and injustice. We do this by splitting the sample into states where slavery was legal during the Civil War and all others. Our Slave State group includes the 15 states, plus the District of Columbia, where slavery was legal in 1861, at the start of the war: Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, North Carolina, South Carolina, Tennessee, Texas, and Virginia. In addition to these 16 states, we also include West Virginia, which was part of Virginia at the outbreak of the Civil War (Woods, 2015). Our Slave State group therefore includes all 11 states in the Confederacy as well as five Union Slave States (Delaware, Kentucky, Maryland, Missouri, and West Virginia) and DC. The designation also aligns closely with the current Census division for the South, but including Missouri and excluding Oklahoma.

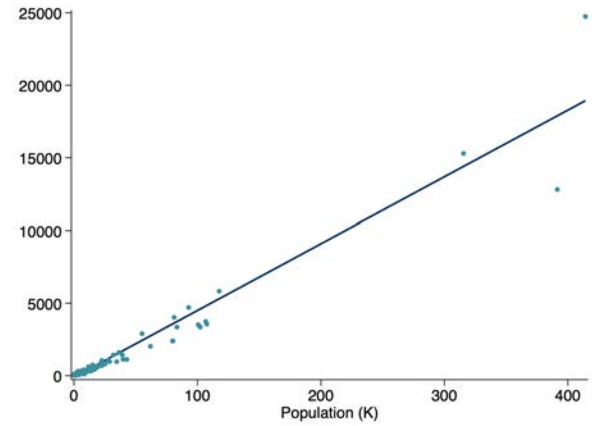
[11] https://www.opm.gov/policy-data-oversight/diversity-and-inclusion/reports/feorp-2016.pdf

This split based on slavery captures a fundamental injustice shaping the historical Black experience in the US (Logan, 2022). Following the abolition of slavery and the end of Reconstruction, state-sanctioned racial injustice continued to be codified for generations in laws mandating racial segregation and disenfranchising sizable Black populations. Research has found evidence of lasting place-based legacies of slavery, reflected in persistent links between historical prevalence of slavery in a location and the subsequent economic, social, and political outcomes of its residents (Jung, 2023; Berger, 2018; Acharya et al., 2016; Reece and O'Connell, 2016; Bertocchi and Dimico, 2014).

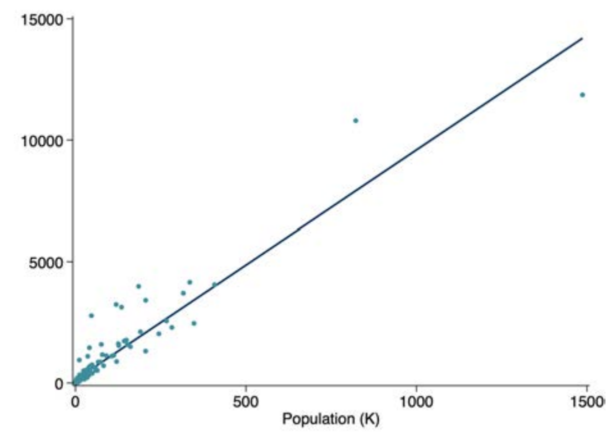
The Slave State designation is associated with both historical and contemporary measures of the Black population distribution within the US (Table A-1). In the first decennial Census following the Civil War, the Slave States had substantial Black population shares, with over three-quarters of states having populations that were more than 20% Black and no states with less than 1%. The reverse was true for non-Slave States, where the majority of states had <1% Black population and none had 10% or higher. This relationship has become less stark over time, as a result of the Great Migration and increasing Black immigration from other countries, which significantly increased the Black population in states such as New York, Illinois, Michigan, New Jersey, and Ohio. Nevertheless, more than 150 years after the Civil War, the 2020 Census shows the majority of former Slave States still have Black population shares above 20%, a threshold not reached in any of the non-Slave States. Median household income is somewhat lower in Slave States for both Black and White households, but the differences are not statistically significant. Intergenerational household income mobility (Chetty et al., 2014) is significantly higher in non-Slave States for both Black and White households, as is individual income mobility for White men and women. Black incarceration rates are similar between Slave and Non-Slave States, but overall incarceration rates are higher in Slave States. These patterns suggest that population density and household income mobility may contribute to differences in algorithmic performance between the two types of states, which we explore in Section 4.1.



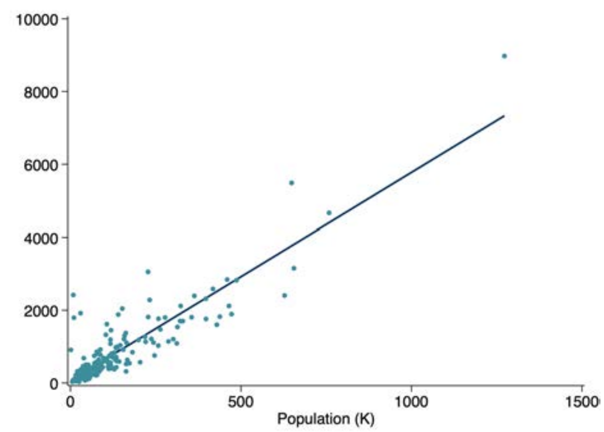
(a) Black Affinity



(b) Asian Affinity



(c) Hispanic Affinity



(d) No Affinity (White)

Figure 1: Ad Reach and Census Population Across Markets by Ethnic Affinity

Notes: Each panel is for campaigns targeted at different ethnic affinity group. Dots on the scatter plots each represent a media market (out of 192). The y-axis is the total reach of campaigns (aggregated across gender, age, and dates) for that market. The x-axis is the total Census population (in thousands, aggregated across gender and age) for that racial and ethnic group.

3 Results

3.1 Ethnic Affinity Is Strongly Related to Race and Ethnicity

Our first result is that the geographic distribution of the ad reach produced by ethnic affinity targeting in Facebook is strongly correlated with the racial and ethnic distribution of the US population. The scatterplots in Figure 1 show, for each ethnic affinity group, that ad campaigns targeting that group were shown systematically more in areas with higher estimated populations of the targeted demographic group using Population counts from US Census Bureau (2017). These are not artifacts of DMAs with larger total populations also having larger populations of different demographic subgroups. We confirmed this by running a regression of ad campaign reach for each affinity group on the total Census population (summed across racial and ethnic groups) and the sub-population in the targeted race or ethnicity (Appendix Table A-3).

3.2 Ads Targeting Black Affinity Are Shown More in Slave States

After finding that the actual racial distribution of a media market is highly predictive of ad reach of ethnically targeted campaigns, we next evaluated whether our campaigns targeting Black or African American affinity were delivered differently across Slave States and non-Slave States. We show this in Figure 2, which plots the reach of our ad campaigns targeting Black affinity, relative to the size of the local Black population in the relevant gender and age group (in thousands). Ads targeted at Black affinity are shown significantly more in Slave States, but the same is not true for other ethnic affinity groups.

Regression model estimates for ads targeting Black affinity confirm the finding in Figure 2 for ad reach per population (an increase of 0.2 in Slave States, significant at the 1 percent level) and also show that ad impressions and clicks are also significantly higher in Slave States, even after controlling for population (Appendix Table A-4). Estimates from pooled models that include other campaigns confirm the higher reach in Slave States for ads targeting Black affinity is not present for other types of ethnic targeting. There is a positive and significant interaction between Black and Slave State.[12]

[12] Appendix Table A-5 shows comparisons to all other ethnic groups (columns 1-4) and to the group excluding multicultural affinities (columns 5-8).

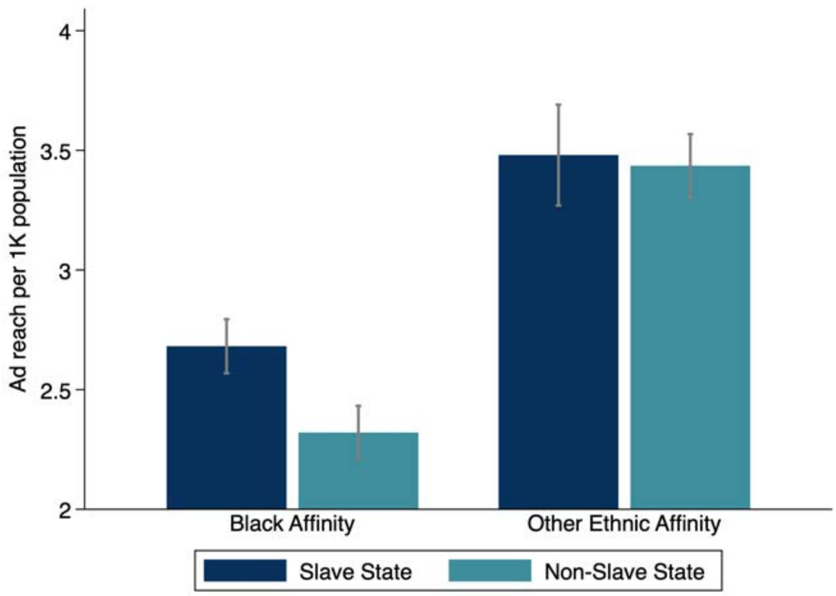


Figure 2: Ad Reach per Population by Black Affinity

Notes: This figure plots average ad reach per population across media markets of Facebook ad campaigns targeted by racial affinity, gender (male or female) and age group (20-24 or 25-29). Ad reach is the number of unique individuals shown the ad and estimated population is from 2016 Census estimates for the corresponding age-race-gender group. We split the demographic target groups based on racial affinity (Black or other) and assigned locations to Slave State groups based on the dominant state in the media market. 95% confidence intervals shown for each bar.

It is possible that the algorithm used to deliver ads to people within the target audience explains our effect (Lambrecht and Tucker, 2019). Ads are only delivered to Facebook users when the algorithm places a winning bid in the auction for that advertising slot on the Facebook platform. Details of the bidding algorithm are a black box from the point of view of researchers, but we were able to specify the ad objectives and payment basis. The objective for all of our campaigns was to maximize clicks on an embedded link (to the jobs page). We set a maximum $20 daily budget. However, costs of ad delivery were not lower, either in absolute terms, or relative to other ethnic affinity target groups, for Black affinity users in Slave States.[13] The magnitudes are small in absolute terms and relative to the mean costs (of about $1.50). None are statistically significant. The comparisons in the later columns yield somewhat larger and positive point estimates. This implies that ads in Slave States are more expensive for Black affinity, and points against our finding. The only highly significant estimate in the table is the indicator for Black affinity (Appendix Table

[13]Estimates are in Appendix Table A-7.

A-7, column 3). This indicates that it is significantly more expensive to serve ads to users identified as having Black affinity than to those with no multicultural ethnic affinity (corresponding to White non-Hispanic users).

3.3 Direct evidence on digital profiling

These results are based on a single campaign run for a specific product. To ensure that our results apply more broadly, we directly collected data on the reported size of potential target audiences on Meta. Specifically, we collected data separately for Slave States and non-Slave States on the number of people profiled as having African American ethnic affinity. This measure comes from the Audience Insights tool on Facebook, that allows advertisers to estimate the audience they will reach if they target different segments on Facebook. We collected data for each state, by each gender, and age group (20-24 and 25-29) combination. Then, we downloaded the estimated Facebook audience size for people with Black affinity. We were not able to use this tool to generate audience sizes for all media markets in our full sample, as Meta restricts access to the tool, which is why we use states instead. Even with this more aggregated unit of analysis (50 states instead of 192 media markets), we still found several target groups for which Facebook provided "No Information." These were typically places with small Census population counts for the target group. For larger audiences, we set the size equal to the midpoint of the range provided by Facebook.

We were not allowed to select an audience that *excludes* any ethnic affinity group, so we are not able to compare Black affinity audience size to other ethnic affinities. We therefore downloaded estimated audience size measures for each state, gender and age group without specifying any affinity (the "All" group in Figure 3). This means that the All group provides a benchmark that is dominated by other racial and ethnic groups, but that is still affected by our target group of interest - the Black share of the Census population in our age groups is 14.4%.

Figure 3 shows the size of audience by Black affinity designation and no affinity group designation. When scaled to the population of these states, the audience size calculated by Facebook is the same as the actual Census population when no ethnic affinity designation is chosen. This was true across both historic free and Slave States. However, this was not the case for users with African American affinity. The Facebook audience size surpasses the population by 34% in Slave

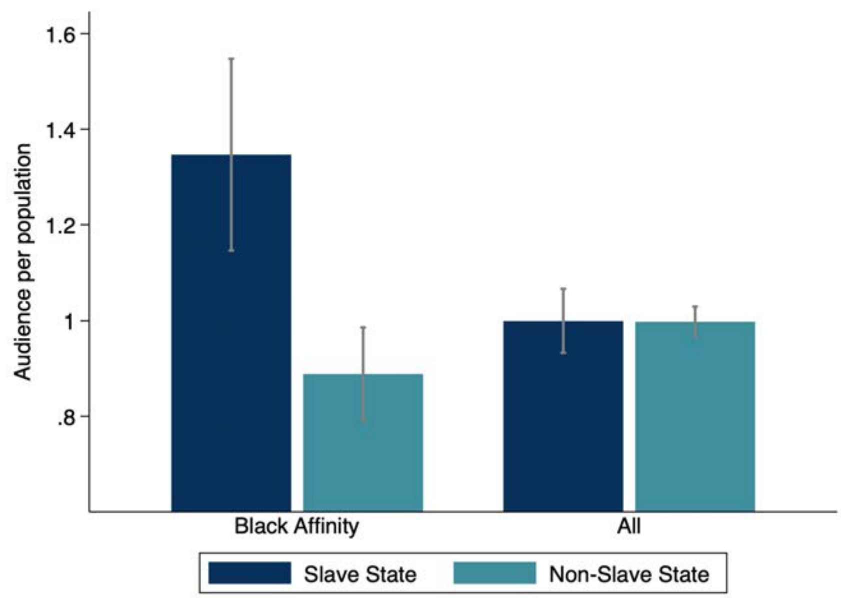


Figure 3: Facebook Audience per Population by Black Affinity

Notes: This figure compares Facebook audience size reports for different target audiences, based on racial affinity (Black affinity or unrestricted), gender (male or female), age group (20-24 or 25-29) and location (state). The audience size is scaled to the estimated Census population for the corresponding gender-age-race group in the state. The Slave State and non-Slave State bars are averaged across states and demographic groups. 95% confidence intervals shown for each bar. We treat missing observations as having an audience size of zero.

States, but is only 88% of the population in non-Slave States.

The significantly larger audience size for Black affinity groups in Slave States corresponds to an average of about 10K additional people per demographic subgroup, after controlling for the state's population of that sub-group (which has an average slope of 0.92).[14] This gap is smaller if we exclude observations with missing audience size, which happens less often for Slave States, rather than including them as zeros. By contrast, we find no significant difference between former slave and free states in the Facebook audience size when we do not target by ethnic affinity. This suggests that the relatively higher rate of ad delivery for Black affinity in former Slave States, conditional on population, could stem from the greater size of the potential target audience in those states.

4 Exploring the Sources of Disparity Between Slave and Non-Slave States

4.1 What variables explain this disparity?

We explore what might drive the lasting effect of historical oppression on the outcomes of the modern-day advertising algorithms targeting users with Black affinity by adding controls in our regression model for ad reach scaled to population. The results in Table 1 show the overall estimate with no covariates (column 1) is relatively stable after adding controls for socioeconomic status (columns 2-5), despite several measures of economic disadvantage (unemployment, low-income population, economic immobility) predicting higher ad volume. Column 6 shows the effect is not explained by variation in the proportion of the Black population (alone or in combination) that reports a single race.

The variables that do account for the Slave State effect are indicators for Black population shares above different thresholds (column 7). This is consistent with population density being the variable that drives the relationship between digital engagement with content and historical oppression. Therefore, a potential mechanism that could explain our effect is that members of an ethnic group are more identifiable through their digital activity in instances where they constitute a larger share of the local population, because they are more likely to develop distinctive communities with shared interests that leads them to more consistently engage with similar digital content. However, because the population share distributions differ sharply across the sets of states - there

[14] See Appendix Table A-6 for full estimates.

Table 1: Accounting for the Slave State Effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Slave State	0.196**	0.201**	0.294***	0.195**	0.159*	0.273**	-0.0763
	(0.0757)	(0.0769)	(0.0772)	(0.0754)	(0.0748)	(0.0937)	(0.0872)
Pct. College Grad.		-0.00421					
		(0.00779)					
Pct. Unemployed			0.0630***				
			(0.0150)				
Pct. Low Income				0.0102*			
				(0.00482)			
Income Mobility					-0.0468**		
					(0.0155)		
Share Black Only						-1.157	
						(0.731)	
Pop. Sh. 1-5%							0.488***
							(0.109)
Pop. Sh. 5-10%							0.681***
							(0.130)
Pop. Sh. 10-20%							0.256*
							(0.118)
Pop. Sh. >20%							0.946***
							(0.137)
Observations	5376	5376	5376	5376	5376	5376	5376
R-Squared	0.00365	0.00383	0.00878	0.00623	0.00991	0.00484	0.0315

Notes: Dependent variable is ad reach (unique viewers) per thousand population. Sample is restricted to ads targeting Black affinity. Each observation is a gender (male or female) by age group (20-24 or 25-29) by media market (192) by date (7) combination. The omitted population share is <1 percent. Regressions are weighted by population. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

are only Slave State observations with shares above 20% – it is not possible to tease the effect of population shares apart further econometrically.

4.2 Mechanism: Variation in intensity of engagement with digital content

We next obtained direct evidence about what may drive this effect, by looking at the types of Facebook content that were highly predictive of race, and how this differed across Slave and non-Slave states.

"Affinity" is defined by Facebook as the degree to which a page is more popular with users from the specific audience target than it is with Facebook users overall, as a multiple of baseline popularity. An affinity score of 10X means that a page is 10 times more popular among members of that audience than the average Facebook user. We collected information on the digital content with the highest affinity scores for users that Facebook labeled as having an African-American affinity. We did this separately for each age groups (20-24 and 25-29) and gender categories and for each state.

The specific content of the top-liked pages with the highest affinity scores for Black audiences in our Facebook data shows some expected results. Among Black males ages 20-24, the top affinity score page was the hair care company WaveBuilder, with 275,200 likes, and a Black affinity audience of 38,400. WaveBuilder itself is a hair care and maintenance line for particular waved hair styles that are predominantly worn by Black men. The WaveBuilder Facebook page is unlikely to interest non-Black men or women who cannot wear this kind of hairstyle and it had an affinity score of 212X. Two top pages for Black women are Jess Hilarious and E'TAE Natural Products. Jess Hilarious is a Black female comedian who has made television appearances on channels like the Black Entertainment Television (BET) channel. The BET and BET Music networks also appear among the top pages for Black men and women. E'TAE Natural Products is a hair and skin care product line geared towards "natural and relaxed hair"—terminology that is frequently used to describe Black hair that is either unaltered by chemical straightening or straightened or texturized using a relaxer.

However, other pages with high affinity scores, such as Champs Sports, Irish Spring (soap), and Wet Seal (apparel and makeup), are less obvious candidates for high scores in the Black affinity

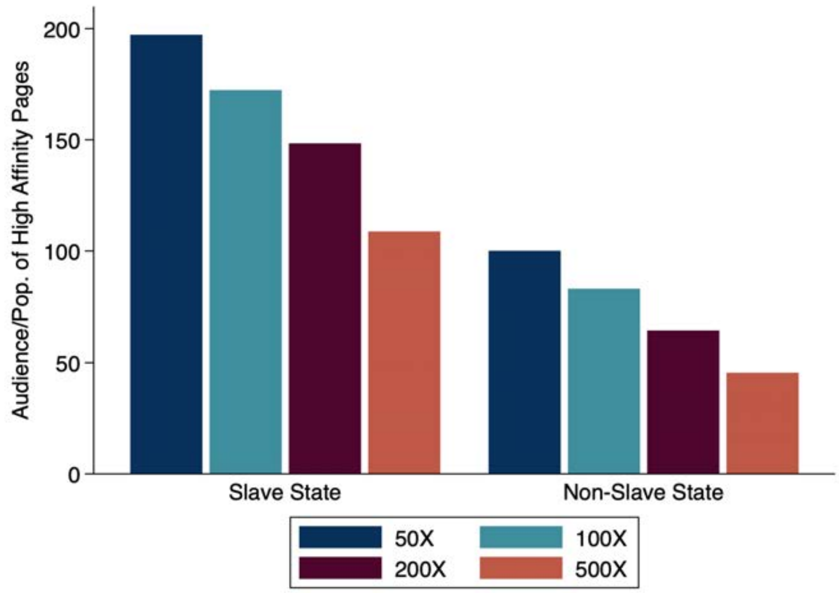


Figure 4: Black Target Audience Size of High Affinity Pages, Scaled to Population
Notes: This figure shows the average value (across groups within each set of states) of the maximum audience size for the target group for Facebook pages with affinity scores above each threshold (50X, 100X, 200X and 500X), scaled to the Census population for the target demographic group in the state.

designation. Consistent with the relatively young age groups we study, every state has one or more state universities among the high affinity content, but it is notable that no pages for HBCUs appear on the lists. These more unexpected correlations illustrate the challenge inherent in attempting to remove all digital proxies for racial identity. It might be obvious to be concerned if a liking for Black hair products were used to exclude someone from seeing an ad for housing, or to target a payday loan. It is less obvious to be concerned about a liking for 'Champ Sports.' Digital engagement with content that is predictive of race includes content that is not itself racially targeted, and may be impossible to predict (or purge) without racial identity data.

We examine differences in Black target audience sizes for pages that Facebook identifies as having high affinity values for Black users between Slave and non-Slave states in Figure 4. The figure plots the average value (across states and sex-age groups) of the maximum audience size for each target group and state, across all of the high-affinity pages, using varying cutoffs for high-affinity (50X to 500X). This audience size measure corresponds roughly to the number of users that Facebook would find from the specified affinity group by focusing only on the high-affinity page with the largest set of users from the group.

The figure shows that, within each state group, the maximum audience size decreases as the affinity threshold is raised, and the maximum is taken over a smaller set of pages. Although the weakly negative effect is fully expected, the steep decline in audience size (from 10K to 3K) reveals that the highest affinity pages – the ones that are most distinctively Black – are not necessarily the ones that are most popular among users with Black affinity.

The main result of Figure 4 is that, for each affinity threshold, the audience size, scaled to Census population, in Slave States is substantially higher than in other states. This illustrates a feature of liking behavior that enables the algorithm to find more users with Black affinity in those states. These differences are statistically significant: the Slave State coefficient is positive and significant for predicting audience sizes from the Black affinity target groups among top pages with high affinity levels, but it is not significant for predicting total (all race) audience sizes (Appendix Table A-8). We similarly find significantly more distinct pages with high affinity for Black audience target groups in Slave States (Appendix Figure A-3 and Table A-9).

Therefore, it appears our results occur because in Slave states there is more differentiated engagement with content that machine learning identifies as associated with a African American ethnic affinity. One explanation of this phenomenon comes from cultural economics, and the finding that cultural transmission is facilitated by population density (Bisin and Verdier, 2011). For example, it is easier to maintain a culture as an immigrant if there is a large population of immigrants from the same country living locally. However, we recognize there may be many other cultural and sociological reasons why there is more distinctive engagement with digital content thought to be linked to an African-American ethnic affinity in states with a history of slavery.

5 Implications and Conclusions

This paper studies the use of racial proxies to infer race in the digital content world. It uses unique data from Meta, and finds that while data on engagement with content does predict race well, there are systematic biases in how well it predicts. The algorithms are more likely to identify someone as being Black, relative to the baseline population, in states with a history of slavery. We present evidence that this occurs because higher population shares of a racial group are associated with engagement with more differentiated online content.

Our findings confirm concerns about the limitations of masking or excluding information on racial groups as a way of preventing discrimination. Not only do we find that engagement with digital content is predictive of racial identity, but also that there is no easy way to purge the data of proxies. The relationship between race and zip code is well-noted (Lee et al., 2019), but avoiding detailed geographic targeting is not enough to remove informative racial classification. Some of the categories of digital content engagement that predicts a Black audiences, such as BET and WaveBuilder, may be expected, but others are not. This makes the project of identifying and removing any individual page or category of pages unlikely to be effective at removing all proxies.

This paper departs from prior literature on algorithmic fairness in economics that focuses on labor markets (Cowgill et al., 2020), healthcare delivery (Obermeyer et al., 2019), or criminal justice (Stevenson and Doleac, 2022), to explore algorithmic decision-making in the context of digital profiling. Predatory marketing targeted at African American communities has been documented for mentholated cigarettes (Li, 2019) and sub-prime loans (Hawkins and Penner, 2020) and in campaigns aimed at suppressing voter turnout.[15] The use of racial proxies for targeting of advertising that promotes products that are not welfare enhancing is particularly concerning when it disproportionately reaches Black populations that have been been subject to more historical racial injustice, such as those in former Slave States, or that show persistent harms from systemic inequities in their worse contemporary measures of economic outcomes. We also find a significant interaction between our two measures of disadvantage. The algorithms' preferential ability to identify African American people in former Slave States is concentrated in lower-income areas (Appendix Table A-10), which suggests a further potential for imperfect racial proxies to exacerbate the harms from racially targeted advertising.

Our results also highlight the challenges in using proxies to predict race. Without information on actual racial identification, it will be impossible for researchers, regulators or firms to identify all of the potentially concerning racial proxies included in digital activity data, which might be a significant share of all data. Furthermore, our finding that inferred race using data on engagement

[15] https://www.justice.gov/usao-edny/pr/social-media-influencer-douglass-mackey-convicted-election-interference-2016.

with digital content produces a proxy measure that is informative, but that also suffers from systematic bias, shows that enforcement of anti-discrimination rules and auditing of algorithms for bias can both be compromised if data on actual racial identity are fully excluded.

There are of course limitations to our study. First, we study the inference of race on one platform at one point in time. Second, though we prevent suggestive evidence that the mechanism for the systematic bias in the prediction of race is due to the fact that there is more differentiated engagement in content in states with with more concentrated black populations we cannot test this directly. Notwithstanding these limitations, we believe that this paper is a useful contribution in understanding both how racial proxies can exist and be systematically biased in the digital world.

References

Acharya, A., M. Blackwell, and M. Sen (2016). The political legacy of American slavery. *The Journal of Politics 78*(3), 621–641.

Agan, A. and S. Starr (2018). Ban the box, criminal records, and racial discrimination: A field experiment. *The Quarterly Journal of Economics 133*(1), 191–235.

Allen, A. L. (2022). Dismantling the "Black opticon": Privacy, race equity, and online data-protection reform. In *Yale Law Journal Forum*.

Angwin, J. and T. Parris, Jr. (2016). Facebook lets advertisers exclude users by race. *ProPublica 28*.

Angwin, J., A. Tobin, and M. Varner (2017). Facebook (still) letting housing advertisers exclude users by race. *ProPublica 21*.

Berger, T. (2018). Places of persistence: Slavery and the geography of intergenerational mobility in the united states. *Demography 55*(4), 1547–1565.

Bertocchi, G. and A. Dimico (2014). Slavery, education, and inequality. *European Economic Review 70*, 197–209.

Bisin, A. and T. Verdier (2011). The economics of cultural transmission and socialization. In *Handbook of social economics*, Volume 1, pp. 339–416. Elsevier.

Carson, E. A. (2018). Prisoners in 2016. NCJ Report 251149.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The Quarterly Journal of Economics 129*(4), 1553–1623.

Cowgill, B., F. Dell'Acqua, S. Deng, D. Hsu, N. Verma, and A. Chaintreau (2020). Biased programmers? Or biased data? A field experiment in operationalizing AI ethics. In *Proceedings of the 21st ACM Conference on Economics and Computation*, pp. 679–681.

Cowgill, B. and C. E. Tucker (2020). Algorithmic fairness and economics. *Columbia Business School Research Paper*.

Doleac, J. L. and B. 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.

Gillis, T. B. and J. L. Spiess (2019). Big data and discrimination. *The University of Chicago Law Review 86*(2), 459–488.

Goldin, C. and C. Rouse (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American Economic Review 90*(4), 715–741.

Hawkins, J. and T. C. Penner (2020). Advertising injustices: Marketing race and credit in America. *Emory LJ 70*, 1619.

Jung, Y. (2023). Formation of the legacy of slavery: Evidence from the US South. *European Economic Review 154*, 104409.

Kallus, N., X. Mao, and A. Zhou (2022). Assessing algorithmic fairness with unobserved protected class using data combination. *Management Science 68*(3), 1959–1981.

Kleinberg, J., J. Ludwig, S. Mullainathan, and C. R. Sunstein (2020). Algorithms as discrimination detectors. *Proceedings of the National Academy of Sciences 117*(48), 30096–30100.

Lambrecht, A. and C. Tucker (2019). Algorithmic bias? An empirical study of apparent gender-based discrimination in the display of STEM career ads. *Management Science 65*(7), 2966–2981.

Lee, N. T., P. Resnick, and G. Barton (2019). Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. *Brookings Institute: Washington, DC, USA 2*.

Li, O. (2019). From housing to health: Imagining antidiscrimination provisions for menthol cigarette marketing. *Colum. J. Race & L. 9*, 369.

Logan, T. D. (2022). American Enslavement and the Recovery of Black Economic History. *Journal of Economic Perspectives 36*(2), 81–98.

Miller, A. R. and C. Segal (2012). Does temporary affirmative action produce persistent effects? a study of black and female employment in law enforcement. *Review of Economics and Statistics 94*(4), 1107–1125.

Miller, A. R. and C. Tucker (2018). Privacy protection, personalized medicine, and genetic testing. *Management Science 64*(10), 4648–4668.

Miller, C. (2017). The persistent effect of temporary affirmative action. *American Economic Journal: Applied Economics 9*(3), 152–190.

Noble, S. U. (2018). Algorithms of oppression. In *Algorithms of oppression*. New York University Press.

Obermeyer, Z., B. Powers, C. Vogeli, and S. Mullainathan (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science 366*(6464), 447–453.

Rambachan, A., J. Kleinberg, J. Ludwig, and S. Mullainathan (2020). An economic perspective on algorithmic fairness. In *AEA Papers and Proceedings*, Volume 110, pp. 91–95.

Reece, R. L. and H. A. O'Connell (2016). How the legacy of slavery and racial composition shape public school enrollment in the American South. *Sociology of Race and Ethnicity 2*(1), 42–57.

Stevenson, M. T. and J. L. Doleac (2022). Algorithmic risk assessment in the hands of humans. *SSRN Working Paper 3489440*.

US Census Bureau (2017). Annual county resident population estimates by age, sex, race, and Hispanic origin. Section: Government.

Woods, M. E. (2015). Mountaineers becoming free: Emancipation and statehood in west virginia. *West Virginia History 9*(2), 37–71.

Online Appendix: Supplementary Figures and Tables

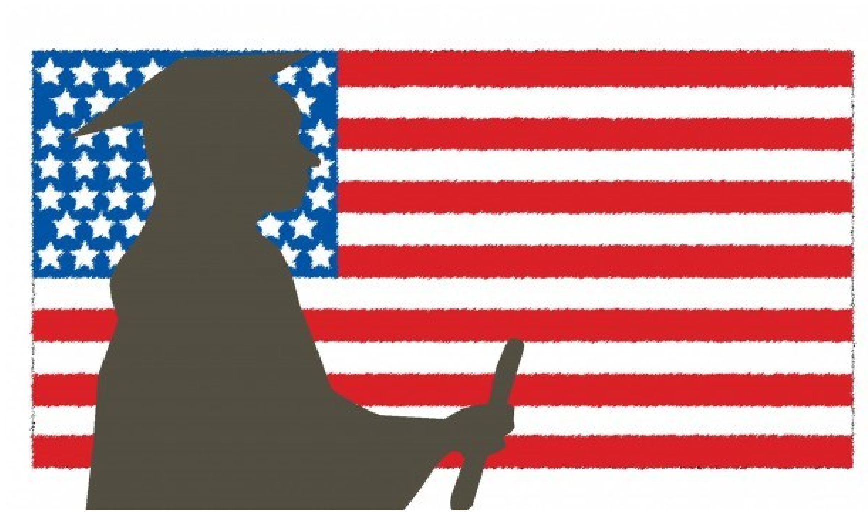


Figure A-1: Sample Ad Image

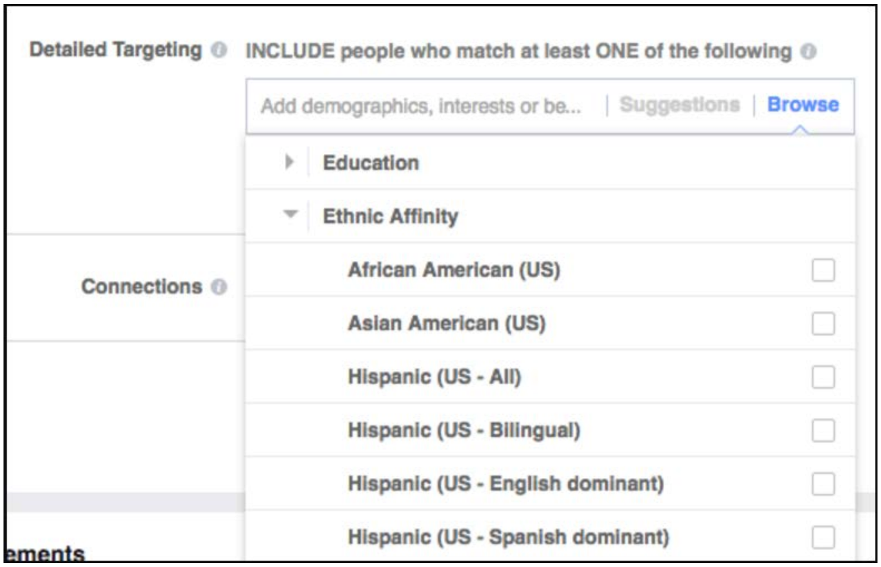


Figure A-2: Facebook Targeting by "Ethnic Affinity"

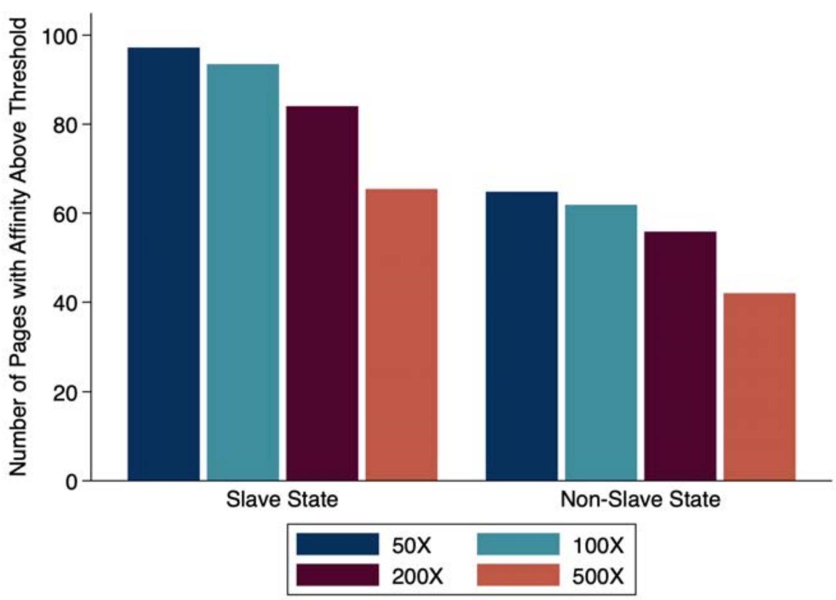


Figure A-3: Number of Facebook Pages with High Affinity

Notes: This figure shows average number of Facebook pages with affinity scores above each threshold (50X, 100X, 200X and 500X). Affinity scores reflect the degree to which the target audience is more likely to like the page than are Facebook users in general. A score of 500X means that target group members are 500 times more likely to like the page. The maximum number of pages per group is 100.

Table A-1: Population Differences between Slave and Non-Slave States

	(1) Slave State	(2) Non-Slave	(3) Diff.
Historical Black Share of State Pop. <1%	0.0	61.3	61.3***
Historical Black Share of State Pop. 1-5%	5.9	35.5	29.6*
Historical Black Share of State Pop. 5-10%	5.9	3.2	-2.7
Historical Black Share of State Pop. 10-20%	11.8	0.0	-11.8
Historical Black Share of State Pop. >20%	76.5	0.0	-76.5***
Current Black Share of State Pop. <1%	0.0	0.0	0.0
Current Black Share of State Pop. 1-5%	5.9	41.2	35.3**
Current Black Share of State Pop. 5-10%	5.9	32.4	26.5*
Current Black Share of State Pop. 10-20%	29.4	26.5	-2.9
Current Black Share of State Pop. >20%	58.8	0.0	-58.8***
Any HBCU in State	100.0	11.8	-88.2***
Median Black Household Income (K)	45.6	48.2	2.6
Median White Household Income (K)	74.3	76.6	2.3
Black Household Income Mobility	32.6	34.8	2.2**
White Household Income Mobility	42.8	46.7	3.9***
Black Female Individual Income Mobility	40.9	41.1	0.3
Black Male Individual Income Mobility	38.9	39.7	0.8
White Female Individual Income Mobility	37.4	41.0	3.6***
White Male Individual Income Mobility	46.7	50.9	4.2***
Black Prisoners Per K Pop.	11.9	11.5	-0.4
Total Prisoners Per K Pop.	5.1	3.5	-1.6***
Observations	17	34	51

Notes: Each state (plus the District of Columbia) is an observation. Historical population shares are from the 1870 Census counts for states and territories (and therefore exclude Alaska, Hawaii, and Oklahoma). Current population shares are from the 2020 Census. Presence of historically black colleges and universities (Any HBCU) is for 2018. Median household income is measured in thousands of dollars. Income mobility measures are the adult income rank for a child born to parents at the 25th household income percentile. Data at the commuting zone from Chetty et al. (2014) are aggregated to the state level using population weights. Prisoner counts are for state and federal facilities from Carson (2018) and exclude DC. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-2: Ad Reach Predicted by Population in Ethnic Group

	Black (1)	Asian (2)	Hispanic (3)	No Affinity (4)
Black Pop. (K)	1.829***			
	(0.131)			
Asian Pop. (K)		7.107***		
		(0.547)		
Hispanic Pop. (K)			1.073***	
			(0.103)	
White Pop. (K)				0.586***
				(0.0490)
Total Pop. (K)	0.0538*	-0.107*	0.119***	0.103***
	(0.0215)	(0.0544)	(0.0307)	(0.0239)
Observations	5376	5376	5376	5376
R-Squared	0.673	0.747	0.750	0.606
Dep. Var. Mean	22.80	25.58	21.20	29.24

Notes: This table relates ads reach (number of unique individuals shown the ad) for ads targeting different affinity groups to the Census population of the targeted racial/ethnic group and the total Census population across groups in the media market. Each observation is a gender (male or female) by age group (20-24 or 25-29) by media market (192) by date (7) combination. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-3: Ad Reach Predicted by (Single Race) Population in Ethnic Group

	Black (1)	Asian (2)	Hispanic (3)	No Affinity (4)
Black Pop. (K)	1.856***			
	(0.130)			
Asian Pop. (K)		7.714***		
		(0.651)		
Hispanic Pop. (K)			1.073***	
			(0.103)	
White Pop. (K)				0.586***
				(0.0490)
Total Pop. (K)	0.0760***	-0.0778	0.119***	0.103***
	(0.0202)	(0.0579)	(0.0307)	(0.0239)
Observations	5376	5376	5376	5376
R-Squared	0.674	0.732	0.750	0.606
Dep. Var. Mean	22.80	25.58	21.20	29.24

Notes: This table relates ads reach (number of unique individuals shown the ad) for ads targeting different affinity groups to the Census population of the targeted racial/ethnic group and the total Census population across groups in the media market. Each observation is a gender (male or female) by age group (20-24 or 25-29) by media market (192) by date (7) combination. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-4: Ads Targeting Black Affinity Shown More in States with a History of Racial Injustice

	(1) Reach Per Pop.	(2) Impressions	(3) Reach	(4) Clicks
Slave State	0.196**	8.165***	2.960**	0.0215**
	(0.0757)	(1.949)	(0.920)	(0.00781)
Population (K)		3.910***	2.053***	0.00568***
		(0.196)	(0.0943)	(0.000601)
Observations	5376	5376	5376	5376
R-Squared	0.00365	0.624	0.671	0.152
Dep. Var. Mean	2.307	45.03	22.80	0.0612

Notes: This table examines ads targeting Black affinity. Each observation is a gender (male or female) by age group (20-24 or 25-29) by media market (192) by date (7) combination. The outcome in column 1 is the ad reach (number of unique individuals shown the ad) per thousand people in the Census population estimate for the corresponding gender-age-race group in the counties that comprise the media market. Outcome variables in the remaining columns are: impressions (number of times the ad was shown), reach, and clicks (the number of times that people clicked on the link in the ad). Column 1 is weighted by population. The remaining columns are unweighted and include controls for the estimated population. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-5: Comparisons to Other Ethnic Affinities and No Ethnic Affinity

	All others				No Affinity			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Reach Per Pop.	Impressions	Reach	Clicks	Reach Per Pop.	Impressions	Reach	Clicks
Black × Slave State	0.507***	23.03***	9.900***	0.0425***	0.227**	9.806***	4.834***	0.0379***
	(0.0881)	(2.394)	(1.119)	(0.00868)	(0.0790)	(2.257)	(1.103)	(0.0104)
Black Affinity	0.557***	-17.45***	-7.740***	-0.0248***	1.300***	-6.612***	-2.488*	-0.0232**
	(0.0720)	(1.446)	(0.696)	(0.00511)	(0.0621)	(1.743)	(1.026)	(0.00886)
Slave State	-0.310***	-14.86***	-6.940***	-0.0210***	-0.0307	-1.641	-1.873**	-0.0165*
	(0.0449)	(1.390)	(0.636)	(0.00378)	(0.0225)	(1.138)	(0.608)	(0.00689)
Population (K)		2.072***	1.180***	0.00314***		1.165***	0.803***	0.00182***
		(0.0805)	(0.0380)	(0.000190)		(0.0478)	(0.0316)	(0.000241)
Black × Pop.		1.838***	0.873***	0.00254***		2.745***	1.250***	0.00386***
		(0.212)	(0.102)	(0.000630)		(0.202)	(0.0994)	(0.000647)
Observations	21504	21504	21504	21504	10752	10752	10752	10752
R-Squared	0.0223	0.308	0.395	0.138	0.137	0.585	0.642	0.108

Notes: This tables compares ads targeting Black affinity to ads targeting other ethnic affinity groups. Each observation is a combination of gender (male or female), age group (20-24 or 25-29), media market (from 192 markets), date (7 days) and ethnic affinity group. The ethnic affinity groups in columns 1 to 4 are Black, Asian, Hispanic, and none; in columns 5 to 8, they are Black and none. Outcome variables are as in Table A-4. Columns 1 and 5 are weighted by population. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-6: Facebook Target Audience Size by State and Affinity

	Black (1) Imputed	(2) Dropped	(3) Non-Missing	All (4)
Slave State	9.962*	9.733*	0.103***	0.808
	(3.890)	(3.918)	(0.0226)	(17.04)
Population (K)	0.919***	0.914***	0.00191***	0.942***
	(0.0644)	(0.0656)	(0.000375)	(0.0677)
Observations	200	176	200	200
R-Squared	0.837	0.821	0.112	0.880
Dep. Var. Mean	36.35	41.30	0.880	222.8

Notes: This table reports separate regressions predicting Facebook audience size (in thousands), or missing audience size, for state-gender-age groups with or without racial affinity targeting. Columns 1 to 3 are focused on Black affinity, while column 4 is not restricted based on ethnic affinity. Column 1 imputes missing audience size with zero, while column 2 omits those observations. The dependent variable in column 3 is an indicator for non-missing audience size. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-7: Costs of Engagement for Black and Other Affinity Groups

	Black Affinity Only (1) Click Cost	All Affinity Groups (2) Click Cost	Black or No Affinity (3) Click Cost
Black × Slave State	-0.00214	0.0362	0.0689
	(0.0622)	(0.0788)	(0.0790)
Population (K)	0.000236	-0.0000331	0.000496
	(0.000634)	(0.000241)	(0.000307)
Black Affinity		0.0505	0.245***
		(0.0691)	(0.0699)
Slave State		-0.0383	-0.0711
		(0.0486)	(0.0487)
Black × Pop.		0.000269	-0.000260
		(0.000676)	(0.000704)
Observations	291	1156	645
R-Squared	0.000465	0.00321	0.0715
Dep. Var. Mean	1.536	1.482	1.403

Notes: This dependent variable in this table is the costs per click (the objective of the campaign). Each observation is a combination of gender (male or female), age group (20-24 or 25-29), media market (from 192 markets), and date (7 days). Column 1 focuses on campaigns targeting Black affinity, while column 2 provides a comparison to all other ethnic affinity target groups, and column 3 compares Black affinity to no affinity. Observations are missing when clicks are zero and there was no spending in the market. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-8: Facebook Audience Size of High Page Affinity, Scaled to Population

	Black			All		
	(1)	(2)	(3)	(4)	(5)	(6)
	> 50X	> 100X	> 200X	> 50X	> 100X	> 200X
Slave State	97.17***	89.35***	83.91***	-25.43	-23.09	-18.50
	(17.71)	(17.19)	(17.23)	(17.89)	(17.30)	(14.73)
Observations	200	200	200	200	200	200
R-Squared	0.167	0.159	0.151	0.00746	0.00651	0.00783
Dep. Var. Mean	131.2	111.6	91.34	179.4	160.4	132.0

Notes: This table reports estimated Slave State coefficients from regressions of the average value (across groups within each set of states) of the maximum audience size (across pages) for the target group for Facebook pages with affinity scores above each threshold (50X, 100X, and 200X), divided by the Census population of the target demographic group in the state. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-9: Number of Facebook Pages with High Affinity

	Black			All		
	(1)	(2)	(3)	(4)	(5)	(6)
	> 50X	> 100X	> 200X	> 50X	> 100X	> 200X
Slave State	32.40***	31.63***	28.19***	0.0928	-2.886	-6.724
	(3.942)	(3.885)	(4.659)	(0.858)	(2.498)	(4.363)
Observations	200	200	200	200	200	200
R-Squared	0.139	0.141	0.110	0.0000330	0.00694	0.0136
Dep. Var. Mean	75.19	72.00	64.88	96.27	90.90	83.89

Notes: This table reports the average number of Facebook pages with affinity scores above each threshold (50X, 100X, 200X and 500X). Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A-10: Racial Targeting and Low Income Population Share

	High Share Low-Income	Low Share Low-Income
	(1)	(2)
Slave State	0.584***	-0.0326
	(0.0833)	(0.102)
Observations	2688	2688
R-Squared	0.0294	0.000106
Fraction Slave State	0.479	0.406

Notes: Dependent variable is ad reach (unique viewers) per thousand population. Sample is restricted to ads targeting Black affinity. Each observation is a gender (male or female) by age group (20-24 or 25-29) by media market (192) by date (7) combination. Regressions are weighted by population. Robust standard errors in parentheses. + $p < .1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$