Hou, Xiaolong; Jiao, Yang; Shen, Leilei; Chen, Zhuo

Working Paper

The Lasting Impact of the Tuskegee Syphilis Study: COVID-19 Vaccination Hesitation among African Americans

GLO Discussion Paper, No. 1397

Provided in Cooperation with:
Global Labor Organization (GLO)

Suggested Citation: Hou, Xiaolong; Jiao, Yang; Shen, Leilei; Chen, Zhuo (2024) : The Lasting Impact of the Tuskegee Syphilis Study: COVID-19 Vaccination Hesitation among African Americans, GLO Discussion Paper, No. 1397, Global Labor Organization (GLO), Essen

This Version is available at:
https://hdl.handle.net/10419/283882

Accepted by The Journal of Population Economics on February 13, 2024.

The Lasting Impact of the Tuskegee Syphilis Study: COVID-19 Vaccination Hesitation among African Americans[#]

Xiaolong Hou[#*], Yang Jiao[†] Leilei Shen[‡], and Zhuo Chen[§¶]

[#]Acknowledgement: The authors have no conflicts of interest to declare. All co-authors agree with the contents of the manuscript and there is no financial interest to report. We certify the submission is original work and is not under review at any other publication. We thank editor Xi Chen and three reviewers for their helpful comments.

[#]Xiaolong Hou and Yang Jiao contributed equally to this work.

[*]John Munro Godfrey, Sr. Department of Economics, University of Georgia. Email: xh54056@uga.edu.

[†]Department of Accounting, Economics, and Finance, Texas A&M University-Texarkana. Email: yjiao0207gmail.com.

[‡]Senior Data Scientist. Airbnb. Email: Leilei.shen@airbnb.com.

[§]Department of Health Policy and Management, University of Georgia. Email: zchen1@uga.edu.

[¶]School of Economics. University of Nottingham Ningbo China.

The Lasting Impact of the Tuskegee Syphilis Study: COVID-19 Vaccination Hesitation among African Americans

Abstract

It is widely recognized that African Americans have a higher level of mistrust towards the medical and health care sector, which results in insufficient utilization of public health services, low participation in clinical research, and vaccination hesitancy. While the Tuskegee Syphilis Study has been identified as a key factor in this mistrust, its specific influence on COVID-19 vaccination uptake among African Americans remains unexplored. Our paper fills this research gap. Our results suggest that the difference in COVID-19 vaccination rates between communities with low and high proportions of Black residents decreases during the study period, but the gap persists. Notably, counties closer to Tuskegee exhibit a slower rate of progress in reducing the racial disparity in COVID-19 vaccination, indicating that the lingering mistrust stemming from the Tuskegee Study has contributed to unequal vaccination rates between African Americans and the rest of America.

JEL: I1, N9

Keywords: Vaccination, COVID-19, Racial Disparity, Tuskegee

1 Introduction

Starting in 1932, the United States Public Health Service enlisted 600 African American men from Tuskegee, Alabama, to participate in the Tuskegee Study of Untreated Syphilis in the Negro Male (henceforth the Tuskegee Study). Treatment was withheld even though penicillin was proven effective and became the standard of care by 1940. In addition, the men were actively discouraged from seeking medical advice from doctors outside the study (Brandt, 1978).[1] The experiment ended in 1972 when a whistle-blower, Peter Buxtun, leaked information about the project to the New York Times.[2] The news release of the Tuskegee Study sparked a nationwide outcry, prompting the National Association for the Advancement of Colored People to file a class-action lawsuit against the United States Public Health Service.[3] The Tuskegee Syphilis Study "became a symbol of their [African American's] mistreatment by the medical establishment, a metaphor for deceit, conspiracy, malpractice, and neglect, if not outright genocide" (Corbie-Smith et al., 1999, p. 542).

The COVID-19 pandemic has disproportionately affected African Americans, both in terms of their economic and health outcomes.[4] The negative impact on this particular population has raised concerns among policymakers and medical professionals regarding vaccination hesitancy within African American communities. Studies have extensively explored the racial disparity in vaccination rates and the factors contributing to it (Ndugga et al., 2021; Nguyen et al., 2021).[5] However, there is a lack of quantitative research at sub-national levels within the United States that examines the underlying causes of these variations, despite the significance of potential racial differences in COVID-19 vaccinations. Furthermore, there is a scarcity of research investigating the influence of historical events on African American vaccination practices. To address these gaps, our study focuses on

[1] Brandt (1978) detailed a thorough summary of the Tuskegee study and racism in medical research.

[2] The paper published the story on the front page on November 16^{th}, 1972, and "dropped a bomb into the laps." Only 74 study participants were alive then; 128 patients had died of syphilis or its complications, 40 of their wives had been afflicted, and 19 of their children had had congenital syphilis. The digitized version of the article was retrieved from the New York Times Archives, Aide Questioned Syphilis Study.

[3] Survivors of the study later reported that the doctors diagnosed them with "bad blood," and they thought they were being treated when in fact they were only given the placebo.

[4] When compared with non-Hispanic Whites in the United States, African Americans are three times more likely to get COVID-19 and up to six times more likely to die from it (Yancy, 2020). A more recent study by Aburto et al. (2022) found that life expectancy fell more for Black men (3.6 years) compared with White men (1.5 years). Black Americans saw increases in cardiovascular diseases and "deaths of despair" over this period. These changes dramatically increase the already large gap in life expectancy between Black and White people.

[5] According to a National Association for the Advancement of Colored People study in November 2021, only 14% of the Black survey respondents trust the vaccine's safety, and only 18% trust the vaccine's safety and plan to get vaccinated. The full report can be retrieved here: Coronavirus Vaccine Hesitancy in Black and Latinx Communities.

the long-term effects of the Tuskegee Study on vaccine uptake among the Black population during the COVID-19 pandemic. By exploring the impact of historical injustices, our findings provide new insights into how such events can contribute to suboptimal health behaviors and emphasize the government's responsibility to advance equity in public health.

The present study evaluates a causal relationship between African Americans' hesitation to receive the COVID-19 vaccine and their exposure to the Tuskegee Study. Building on the theory of social identification (Bhattacharya et al., 1995; Ashforth and Mael, 1989), we gauge the exposure to the Tuskegee Study by assessing the geographic proximity to Macon County, Alabama, the site of the Tuskegee Study. African Americans who live in close proximity to Tuskegee may be more knowledgeable about the event and, therefore, may have a higher level of medical mistrust and a lower level of vaccination uptake.[6] Empirically, we utilize a three-way fixed effect design, and the coefficient of interest is the interaction between a time indicator, the percentage of the Black population, and the exposure to the Tuskegee study based on the proximity to Macon County. In all specifications, we compare the change in vaccination rates across counties with large proportions of the Black population and those near Tuskegee and those with low proportions of the Black population but far from Tuskegee.

Our findings show that, while the overall vaccination gap between communities with low and high proportions of Black residents gradually closes over time, the vaccination rate — for both those who are fully immunized and those who have received at least one dose of the vaccine — increases more slowly in counties with a higher proportion of African Americans. Additionally, the rate of increase is substantially correlated with exposure to the Tuskegee Study, measured by the distance between the county and Tuskegee, indicating that the gap closes more quickly due to less mistrust when there is less exposure to the Tuskegee Study. Our findings indicate that COVID-19 vaccine hesitancy is more common in locations with higher percentages of Black residents.

There are two potential threats to our identification. The first comes from the reporting of COVID-19 vaccination rates because the Centers for Disease Control and Prevention (CDC) released COVID-19 vaccination rates for all races. In all specifications, we add both the interaction between week dummies and the share of the White population, and the interaction between week dummies and the population share of Hispanics to control for differentiated trends in vaccination growth for other race groups. Another challenge to our identification is the likelihood that vaccine spillover effects are substantially stronger

[6] Proximity between agents, as described by (Tabellini, 2010, p. 680), "could refer to geography, but also to social or economic dimensions such as religion, ethnicity, and class."

for persons living in the same community (Bruhin et al., 2020), especially among people of various races. We divide counties into distinct racially separated residential regions and compare the estimated results of separate regressions for high- and low-segregated areas to exploit the existence and extent of such spillover effects.[7] The estimates for high- and low-segregated areas, on the other hand, are very similar, showing that spillover effects are not responsible for our findings.

In addition, we perform a series of robustness tests. First, we utilize the intensity of news coverage regarding the Tuskegee Study in different counties to establish an alternative measure of Tuskegee exposure. The results corroborate that in counties with higher levels of exposure to Tuskegee news reporting, the vaccination rate is indeed lower. Second, these findings apply exclusively to African Americans and do not extend to other racial groups. Moreover, the results hardly exhibit any variation, and in fact, they become more pronounced in size when we restrict our sample to the Southern regions. Third, given that our treatment variables (the Black population share and distance to Macon, Alabama) remain constant over time, there is a potential for changes in demographic composition due to migration. The migration during the pandemic may also have an impact on our results. To address this concern, we have conducted two tests. In the first test, we replace the Black population percentage in 2010 with the percentage between 1990 and 2005, resulting in small changes in the results. Our second test involves excluding counties with the highest in and out-of-state migration during the pandemic, and our findings remained consistent and robust.

Additionally, our findings are robust to estimation on a nationwide sample and the baseline coefficients for vaccination rate based on geographic proximity to Macon County are larger than 96.7% of placebo tests when substituting proximity to other falsely assigned counties. We have run another robustness test by analyzing the effects of the Tulsa Massacre, another incident that specifically targeted African Americans. The estimates show that the Tulsa Massacre has no impact on vaccination rates. Finally, in the spirit of Altonji et al. (2005), we provide a measure to gauge the strength of the likely bias arising from unobservables. Our findings indicate that it is unlikely that unobserved heterogeneities alone provide a complete explanation.

We have also performed several heterogeneity analyses. Several findings emerge. First of all, counties characterized by higher proportions of young African Americans exhibit not only lower vaccination rates but also encounter slower progress and face a greater challenge

[7] The idea behind such an experiment is that if our results are predominantly driven by spillover effects between people, the impacts in less segregated areas (with more social interaction) should be more significant than in highly segregated areas.

in closing the racial vaccination gap. Political ideology is also important, and the impact of the Tuskegee Study is stronger among Republican-leaning counties. Education plays a crucial role, with a more pronounced effect observed in counties where a higher proportion of residents have attained higher levels of education. Lastly, vaccination eligibility is crucial because the vaccination gap narrowed after vaccine eligibility expanded.

This article builds on and contributes to two strands of literature in economics and public health. First, our research adds to the body of research on medical mistrust and health consequences. Alsan and Wanamaker (2018) find that the disclosure of the Tuskegee experiment led to an increase in medical mistrust and mortality, a decrease in the life expectancy of Black men, and declines in both outpatient and inpatient visits for older Black men. In a subsequent paper using a randomized controlled trial, Alsan et al. (2019) find that Black men are more likely to get a flu vaccine when paired with a Black doctor. Lowes and Montero (2021) find that greater exposure to colonial medical campaigns that resulted in deaths and severe side effects reduces vaccination rates and trust in medicine in Central Africa.

Moreover, our paper is also related to a broader literature on the historical origin of (mis)trust. We contribute to the literature on mistrust in several ways. First, we show that the mistrust from suffering historical wrongdoing could have an intergenerational impact (Knack and Keefer, 1997; Algan and Cahuc, 2010; Nunn and Wantchekon, 2011). The aftermaths of the Tuskegee Study include vaccine hesitation among African Americans to this day. Second, we find that the mistrust could have a spatial pattern, as the closing of the gap in vaccine takeup rates correlates to the distance of the county to Tuskegee. Third, we show that historical traumas unrelated to health, such as the Tulsa Race Massacre, while with a lasting impact, have little bearing on future generations' health-seeking behavior (Dupas and Miguel, 2017).

2 Data and Estimation Strategy

2.1 Data Source

Vaccination Coverage

The vaccination coverage data are retrieved from the CDC COVID-19 Vaccine Tracker.[8] Weekly vaccine rates for each county between January and December 2021 are used.[9] Table A1 summarizes the vaccination rate for 52 weeks over the study period and Figure A1 shows the percentage of each state's population aged 18 and older who are fully vaccinated in mid-May, early October, and mid-December, individually. As expected, more people get vaccinated over time, but there is a wide variation in reported vaccination rates by county across the United States.

Racial Composition and County-level Variables

The fractions of populations of different races aged 18 years and older at the county level are computed based on data from the 2010 Census.[10] Figure 1 shows the distribution of the Black population aged 18 years and older at the county level where darker-shaded areas represent a higher Black population share. The pronounced clustering of the Black population in the Southern United States is readily evident. The Southern region exhibits a distinctive pattern of Black population distribution, which has deep-rooted historical underpinnings and continues to shape the demographic landscape.

[Insert Figure 1 Here]

Data regarding county-level demographic and socioeconomic information is obtained from 2008-2012 American Community Survey (ACS) 5-year estimates. Data on county-level unemployment rates are obtained from the U.S. Bureau of Labor Statistics. Additionally, we collect information about primary care physicians and nurses based on data from the 2019 Area Health Resources Files. The U.S. Census Bureau's County Business Patterns

[8]The link to CDC COVID-19 Vaccinations in the United States can be found here: COVID-19 Vaccination Tracker.

[9]Counties in Alaska and Hawaii are excluded from the analysis due to missing values. Additionally, CDC vaccine coverage data does not include vaccine coverage rates for some counties in certain weeks. Therefore, our baseline analysis is based on an unbalanced panel.

[10]We classify people into non-Hispanic White, non-Hispanic Black, Hispanic/Latino, non-Hispanic Asian, and others. In addition, we calculate the racial composition by age groups and gender.

include information on pharmacy locations. We also utilize data from Bazzi et al. (2020), which quantifies the extent of time each county spent on the frontier during westward expansion in the 18th and 19th centuries.

Vaccine Hesitancy

Based on the Household Pulse Survey (HPS), a rapid-response survey of persons aged 18 and up conducted by the U.S. Census Bureau in collaboration with seven other federal statistical agencies, we calculate vaccination hesitancy for each county. The study collects detailed information about household experiences during the COVID-19 pandemic.

2.2 Summary Statistics

Table 1 compares counties that are close to Tuskegee and have a high percentage of the Black population with those counties that are far from Tuskegee and have a low percentage of the Black population. A county is defined as having a low (or high) share of Black individuals if the proportion of Black individuals falls within the bottom 30^{th} (or top 30^{th}) percentile of the distribution. Furthermore, if a county's distance to Tuskegee falls within the bottom 50^{th} percentile of the distribution, it is classified as "close"; otherwise, it is considered a "far" county. The results suggest that these two groups of counties differ significantly in almost all characteristics, including the demographics of residents, the educational attainment of residents, local economic and labor market conditions, and access to medical resources. It is not surprising that counties with a lower proportion of Black residents and located far from Tuskegee tend to be more wealthy, have a more educated population, and have a low unemployment rate.

[Insert Table 1 Here]

2.3 Empirical Specifications

The main estimation in our study is to explore the dynamic impact of Black population share, the impact of the Tuskegee Study, and their interaction on a county's vaccination rate, using the following regression model:

$$Y_{ct} = \alpha + \beta_1 PB_c \times Dist_c \times Week_t + \beta_2 PB_c \times Week_t + \beta_3 Dist_c \times Week_t + \gamma X_{ct} + \tau_c + \tau_t + \epsilon_{ct}, \quad (1)$$

where Y_{ct} represents the vaccination rate of people at least 18 years old for county c at week t, and PB_c measures the share of Blacks in county c. The exposure to the Tuskegee Study is assessed by $Dist_c$, the geographic distance between county c and Macon, Alabama, where the Tuskegee Study was conducted. As described by Tabellini (2010), "distance between agents could refer to geography, but also to social or economic dimensions such as religion, ethnicity, and class." As a result, being closer to Macon County meant being more exposed to the Tuskegee Study revelation. The interaction term, $PB_c \times Dist_c \times Week_t$, which demonstrates the heterogeneous effect of Tuskegee exposure to communities with varying percentages of Black people, is of particular relevance.

The term X_{ct} includes a linear trend of the share of the White population, a linear trend of the share of the Hispanics, a linear trend of the share of high school graduates, and a linear trend of the unemployment rate. These allow us to control for a linear trend of county specific factors affecting its population vaccination rate. In alternative specifications, we also control for the non-linear trend and the results are barely changed.

County fixed effect τ_c controls for time-invariant county-specific factors, such as the stock of medical facilities, the number of primary care physicians, and the number of drug stores; while week fixed effect τ_t controls for a common economic/social impact which affects all counties at the same time. In the analyses, all regressions are weighted by the total population age 18 and older at the county level, and the standard errors are clustered by state.

3 Empirical Results

3.1 Baseline Results

We begin our study by showing the weekly trend of vaccination rate for those who have received at least one dose of vaccine in Figure 2. In Panel (a), counties are categorized into two groups based on the proportion of Black residents, while in Panel (b), counties are classified according to their proximity to Tuskegee, similar to the grouping shown in Table 1. The bottom panel depicts the interplay of the Black population and distance. Only the vaccination patterns for these two groups are displayed in Panel (c), where the Black line represents the former and the red line, respectively, represents the latter. The pattern is striking. Counties close to Tuskegee with significant Black populations initially had lower

vaccination rates but gradually caught up to the latter.[11]

[Insert Figure 2 Here]

It is essential to highlight that the key specification in Equation 1 focuses on comparing counties in a way that goes beyond just looking at differences between places with a high or low Black population or the distinction between those in proximity to or far from Tuskegee. The county and time fixed effects have addressed these comparisons. Our analysis is instead focused on examining the time variation in vaccination rates between counties with high percentages of the Black population and those with low percentages, with a particular emphasis on the influence of the Tuskegee Study.

The first two columns of Table 2 report our baseline results. For both the fully vaccinated rate and the rate for at least one vaccine dose, the vaccination rate increases more slowly in counties with a higher Black population. However, the growth rate correlates with the distance to Tuskegee. More specifically, for two counties with the same proportion of Black people, the one further away from Tuskegee has a higher growth rate. In other words, if the proportion of Black people in the population rises from 1^{st} to 99^{th} percentiles, the vaccination rate falls by 67.38 percentage; however, the distance to Tuskegee is associated with a gradual reduction of the difference.[12] The difference is narrowed by 16 percentage for every 100 miles traveled. Furthermore, we conduct two alternative estimations to account for the within-county variation over time, as presented in Table 2. In the first approach, we incorporate county-by-month fixed effects, enabling us to control for unobserved factors that are specific to each county and month but remain constant over time. Additionally, we introduce a county-by-week linear trend, which captures the linear changes in vaccination coverage at the county level over time. The results from both sets of regressions align with our primary specification, affirming the consistency of our findings.

[Insert Table 2 Here]

The implications of our estimated results are twofold. To begin with, it improves our

[11] We have also depicted the trend for each of the six groups that were present in every possible combination of the categories between Panel (a) and Panel (b). The pattern suggests that vaccination uptake is inversely related to the distance from Tuskegee. The higher the vaccination rate, the further a county is from Tuskegee, regardless of its racial composition. Furthermore, within the "near" and "far" distance groups, counties with a higher proportion of Black people had lower vaccination rates at first but eventually caught up with others. The results are available upon request.

[12] Black people make up 0.00001 percent of the population in the bottom 1^{st} percentile and 56.15 percent of the population in the top 99^{th} percentile.

knowledge of geographical heterogeneity in vaccination rates.[13] Education, income and wealth gap (McLaughlin et al., 2022), and access to health facilities (Murthy et al., 2021), are among the most important factors, according to recent studies. Our findings reveal that the Tuskegee effect is significant and sizable, implying that the difference in COVID-19 vaccination rates between counties is suggested to be more of a function of culture than of traditional economic factors. Second, our findings enable us to better understand the vaccination disparity between Black and White people. According to a recent Kaiser Family Foundation (KFF) report, Blacks and Hispanics have been less likely than their White and Asian counterparts to receive a COVID-19 vaccine throughout the rollout. However, these disparities have narrowed over time and reversed for Hispanics.[14] Our findings also lend credence to the notion that the vaccination rate gap between communities with varying percentages of Black populations, particularly those located far from Tuskegee, becomes smaller. Consider two hypothetical communities, one entirely made up of only White people and the other entirely made up of Black people. The distance needed to close the racial disparity between these two counties is 750 miles, assuming that vaccination rates grow in a linear fashion.[15]

3.2 Results by Cohorts and Gender

Given that the subjects in the Tuskegee Study were Black men aged 25 and above in 1932, the influence of this study is expected to be more pronounced within the older Black male demographic. We first ran separate regressions for age groups, ranging from 18 to 24, 25 to 34, 35 to 44, 45 to 54, 55 to 64, and 65 and older. We then run separate regressions for men and women in various age groups.

The results are displayed in Table 3. We have three main findings. First, counties with a higher share of the young Black population have a lower vaccination rate. This is consistent with the study by Baack et al. (2021), which was published in the CDC's Morbidity and Mortality Weekly Report and demonstrates that non-Hispanic Black adults have the lowest

[13]Up to May 2022, at least 257 million people or 78% of the population have received at least one dose. While vaccination coverage has increased, it remains uneven across the country. In five of the six New England states, for example, more than 60% of residents are at least partially vaccinated. In the South, however, Mississippi, Alabama, Arkansas, Georgia, Louisiana, and Tennessee have the lowest rates of residents who have had at least one vaccination in the country.

[14]The link to the full description of the KFF report is Latest Data on COVID-19 Vaccinations by Race/Ethnicity.

[15]The first-order derivative with regard to the share of Black people yields the distance between two counties with the same percentage of Black people. Therefore, it requires ((1.2/0.16)×100) miles to close the gap. In terms of magnitude, there are around 1205 counties with a distance greater than 750 miles, making up 38.8% of all counties (3,107 counties).

reported vaccination coverage and intent to get vaccinated, along with those with less education, no health insurance, and lower family incomes. Second, when older Black persons are included in calculating the proportion of Black people in the population, the vaccination gap between high and low-density areas narrow. As shown in column (1) of Table 3, the Tuskegee Study's effect is smaller for younger people, implying that a highly Black-concentrated county must converge to a county with a low Black share if it is located 894 miles away from Tuskegee. In comparison, column (6) indicates that convergence occurs when a county with a larger senior Black population is roughly 451 miles away from Tuskegee. Third, the vaccination rate is lower in counties with more Black men, and the convergence distance is likewise greater than in counties with more Black women. There are a couple of possible explanations, including Black males tend to have lower educational attainment and worse market outcomes than Black females. Furthermore, males have a more elastic need for health services (Cubbin et al., 2000) and are thus more readily deterred from visiting physician and pharmaceutical websites.

[Insert Table 3 Here]

4 Robustness Checks

4.1 Non-linearity and Racial Residential Segregation

Because the vaccination data includes all races, a few assumptions are necessary in order to determine that changes in the full vaccination rate are predominantly driven by changes in the Black vaccination rate. First, the rate of variation in Whites and other races should have remained constant or followed a consistent pattern over time. This problem is addressed by controlling the share of other races and their interactions with time. We also allow non-linear growth rates for other races. The results for the nonlinear trend are shown in Table A2. The point estimates remain nearly unchanged, and our conclusions are similarly corroborated by the estimations.

Additionally, there should be no cross-group spillover, meaning that Blacks' vaccination behavior is unrelated to Whites' in the same area. To determine the size of the spillover effect, we divide counties into high- and low-racially segregated residential regions, assuming that high-racially segregated areas have less interaction across races. We then examine the results of separate regressions for high- and low-segregated areas. If a county is on

the top (bottom) 25^{th} of the distribution, it is considered high (low) segregated.[16] The findings of the racial segregation classification are presented in Table A3.[17] The results for the least racially segregated regions are displayed in columns (1) and (4), whereas the results for the most segregated regions are displayed in columns (2) and (5). The minimal differences observed among regions with diverse levels of racial segregation provide strong assurance that our findings remain unaffected by any potential racial spillover effects. A more rigorous test indicates that the top and bottom racially divided regions do not statistically differ from one another.

4.2 Tuskegee Study's Impact on Other Races

One possibility that could undermine our causal relationship between the Tuskegee Study and Black vaccination practice is that public exposure to the Tuskegee Study could affect people of all races. We re-estimate the regression in Equation 1 by substituting the proportion of the Black population with the population of other racial groups. This is done to investigate whether the impact of the Tuskegee Study is similar for individuals of different racial backgrounds. The theory behind this test is that the Tuskegee study should only impact the COVID-19 vaccination among African Americans while having little or no effect on people of other races. The results are presented in Table 4. This table is split into two parts: the left segment displays results for individuals from various racial backgrounds, while the right segment showcases findings for individuals of Black ethnicity. Furthermore, the table has been divided into three panels, each reporting results specific to Whites, Hispanics, and Asians.

There are several noteworthy discoveries to highlight. To start, as expected, the coefficients for the White community are both statistically insignificant and of a relatively small magnitude. This implies that vaccination patterns among White individuals appear to remain relatively consistent across different levels of White population density, and the influence of the Tuskegee Study does not seem to be prominent. Second, Hispanics and Asians are concentrated in specific geographic areas that differ significantly from the geographical distribution of Black individuals. In Figure A2, we present the percentages of Hispanics, and Asians along with based on 2010 Census Data for each county. Black populations

[16] The information of racial segregation can be found here: residential segregation – Black/White. Index of dissimilarity where higher values indicate greater residential segregation between Black and White county residents.

[17] Because around 800 counties do not have a segregation index, we discard them and re-estimate the regression using Equation 1. After deleting counties with missing segregation index, the point estimates are extremely near to the baseline values, supporting the sample's representation.

are primarily concentrated in the Southern regions; Hispanics are primarily located in the southwestern states, and Asians are concentrated on the eastern and western coasts. To eliminate the possibility that the outcomes are influenced by geographic separation or the choice of residence for Hispanic and Asian populations, we conduct our analysis by excluding states with the highest proportion of Hispanics (Asians). The findings show no significant impact on the vaccination behavior of Hispanic and Asian communities, reinforcing the notion that the Tuskegee Study has minimal influence on these groups.

[Insert Table 4 Here]

4.3 Alternative Measurement of Tuskegee Exposure

Our major estimate of Tuskegee Study exposure is based on a county's geographic distance from Macon County, assuming that people who lived closer to the study's victims were more significantly influenced by the news coverage. We are still curious to see how our results hold up against a different measure of Tuskegee Study exposure.

To develop an alternative metric, we carefully examine regional variations in any news or report containing the Tuskegee Study in the newspapers circulated between 1972 and 1973 and create an index that captures varied levels of Tuskegee news exposure across areas. In particular, we search all pages of all papers in the database for mentions of "Tuskegee Syphilis Study" or 'Tuskegee Syphilis Experiment," resulting in about 27,012,238 pages of newspapers from about 2,000 different publications.[18] Albright et al. (2021) constructed a similar metric to analyze the gradient impact of the Tulsa Massacre on Black communities across the United States.

Since information regarding news coverage is obtained at the state level, we adjusted it to the county level by multiplying the percentage of state news that featured the Tuskegee Study between 1970 and 1972 by the proportion of the county population (as per the 1970 Census) among the total articles published. Greater newspaper coverage would mean that the disclosure of the Tuskegee Study would have been more extensively and clearly communicated among the public. In particular, our measure is constructed based on the following formula.

$$TK_c = \frac{Pop_{cs,1970}}{Pop_{s,1970}} \times News_{s,1972-73}, \tag{2}$$

[18]The historical distribution of news was retrieved here: Newpaper. The source does not offer comprehensive or representative coverage of all newspapers, but the selection does give an idea of how quickly the news of the Tuskegee Study traveled across the nation.

where $News_{s,1972-73}$ is the ratio obtained by dividing the number of news articles about Tuskegee that were published in the state of s between 1972 and 1973 and contained the keywords "Tuskegee Syphilis Study" or "Tuskegee Syphilis Experiment" by the total number of articles published in that state.[19] To map the state exposure to each county within a state, we multiply it by $\frac{Pop_{cs,1970}}{Pop_{s,1970}}$, the share of county c's population in the state s.[20] As shown in Figure A3, there is vast variation in exposure to the Tuskegee study across regions.

We replace the distance to Tuskegee with the news exposure in the main specification of Equation 1 and report the results in Panel A of Table 5. It is worth noting that the newspapers' political inclination appeared to be in line with their readers' political ideologies and preferences (Gentzkow and Shapiro, 2006; Iyengar and Hahn, 2009). Despite the fact that the county fixed effect accounts for such a time-invariant influence, political ideology and preferences may evolve at different rates throughout time. To capture time-varying trends, we interact with a county's majority political vote with a time dummy.[21]

As demonstrated in Panel A of Table 5, when a county is exposed to more Tuskegee news, the vaccination rate is lower. Specifically, a one percentage point increase in the presence of Tuskegee news relative to all news reported in a county results in an 8.83% reduction in the vaccination rate for at least one dose and a 9.31% reduction for those fully vaccinated. However, exposure to Tuskegee news does not appear to have differing effects on counties with varying proportions of Black residents. The newspaper measure is based on the assumption that Black residents and individuals of other racial backgrounds of a county have equal access to newspapers and are, therefore, equally exposed to shocks when the Tuskegee Study is made public. While it is true that newspapers have limited reach among Blacks, particularly among those who are poor, we lack nationwide data to compare the readership between Blacks and people of other races during 1970. We admit that newspaper coverage may have exaggerated the actual influence of the Tuskegee Study.

[Insert Table 5 Here]

[19] The results have barely altered despite the addition of other ratios that were created by gradually extending the 1972–1980 era.

[20] The data source for calculating such population share is from County Intercensal 1970-1972.

[21] We divide the votes for the Democratic (Republican) party by the total votes in that county to get county-level data on support for the Democratic (Republican) party in the 2020 presidential election from the MIT Election Data and Science Lab. Figure A4 shows the county level and vote share results of the 2020 U.S. Presidential Election. The darker the blue, the more Democratic a county voted, and the darker the red, the more Republican a county voted. If a county receives more than half of the Democratic vote, it is classified as Democratic; otherwise, it is coded as Republican.

4.4 Is It Exclusively in the South?

One could argue that our findings are predominantly influenced by unique characteristics of the Black communities in the Southern region, such as their tendency to exhibit lower trust in modern medicine and their reluctance towards receiving vaccinations. In such a scenario, we would expect to observe a statistically insignificant coefficient, as the vaccination rate should not vary based on the proximity to Tuskegee if these specific factors are the primary drivers. To investigate it, we restrict our analysis to counties in the Southern region. As shown in Panel B of Table 5, our major findings remain consistent and robust. Moreover, the influence of the distance to Tuskegee, conditional on the proportion of the Black population, becomes even more pronounced.

4.5 Impact of Tulsa Massacre

We perform another robustness test and investigate the impact of another historical event, the Tulsa Massacre, that targeted African Americans. Both the Tuskegee Study and the Tulsa Massacre were two significant events in the history of the Black community in the United States. Although each event had distinct repercussions on the community, they had varying impacts on health and trust. In this study, we focused on the Tulsa Massacre due to several reasons.

Firstly, the Tulsa Massacre took place in 1921 within Tulsa, Oklahoma's Greenwood District, often referred to as "Black Wall Street." It involved a violent assault perpetrated by a White mob on the thriving Black community, destroying over 1,200 homes and the loss of hundreds of Black lives. The consequences of the Tulsa Massacre were profoundly devastating for the Black community, leading to the destruction of property, businesses, and the loss of lives. The trauma inflicted by this event lingered for generations (Albright et al., 2021). However, it is worth noting that the economic aftermath of the Tulsa Massacre overshadowed its impact on public health trust, indicating that its influence on Black vaccination decision-making might be relatively limited.

Furthermore, it is important to highlight that the perpetrators of the Tulsa Massacre were predominantly White mobs, whereas the Tuskegee Study was conducted by various government entities such as the U.S. Public Health Service, a division of the Department of Health and Human Services (HHS), the National Institutes of Health (NIH), and the Centers for Disease Control and Prevention (CDC). These very agencies are currently involved in the COVID-19 vaccination program. Consequently, considering the historical

trauma inflicted by the Tuskegee Study, there is a possibility that the Black community may harbor distrust towards these government agencies, which could potentially impact their willingness to receive the COVID-19 vaccine. By contrast, the Tulsa Massacre may not have had a similar effect, making it a plausible comparison.

To do so, we modify the estimation function in Equation 1 by substituting the distance to Tulsa, Oklahoma, for the distance to Macon, Alabama, as in our baseline and maintaining the same sets of control variables. The results in Panel C of Table 5 reveal that distance to Tulsa has no impact on reducing or exacerbating the vaccination gap between high and low Black population density.

4.6 Permutation Tests

We further conduct a set of placebo regressions, substituting the baseline proximity measure (proximity to Macon County, Alabama) with the proximity to the geographic centroid of all counties (except Macon), to test if the baseline results are unique. We perform these permutation tests to see if the results hold when we examine the distance between a county and a location that has been artificially altered—the geographic centroid of each county. As noted in Alsan and Wanamaker (2018), these regressions serve as control experiments to see if we find the same vaccination effects as a function of the gradient to other U.S. locations. Another advantage of doing this is to avoid bias in selecting a particular group of locations that might be responsible for our results. In each of these tests, the distribution of the estimated values is shown in Figure 3, and the vertical line is the estimated triple coefficient from Table 1. For at least one dose (full immunization), the main estimate (proximity to Macon County) is higher than 96.8% (96.7%) of the placebo estimates.

[Insert Figure 3 Here]

4.7 Using Selection on Observables to Assess the Bias from Unobservables

Despite our attempts to control for a variety of observable factors, the estimates may still be biased by unobservable factors correlated with a location choice and subsequent vaccination behavior. In this section, we assess the likelihood that the estimates are biased by unobservables. The strategy that we use exploits the insights from Altonji et al. (2005); Nunn and Wantchekon (2011) to evaluate how much greater the influence of

unobservable factors would need to be about observable factors to fully account for the relationship between the distance to Tuskegee, the percentage of the Black population, and the vaccination rate.

We take into account two sets of restricted covariates: one with no controls and the other with a minimal set of individual controls that include the percentage of the White population and the share of the Hispanic population in 2010. We also consider several sets of full controls, including the initial set of controls from Equation 1, a second set that includes the nonlinear trend for the share of Whites and Hispanics, and a third set that further includes additional influencing factors for vaccine supply. As displayed in Table 6, our findings indicate that the impact of unobservable factors would need to be at least 20.7 times greater than observable factors for at least one dose, and 18.2 times greater for full vaccination, to fully explain our estimations. It is unlikely that unobserved heterogeneities alone can completely account for our findings.

[Insert Table 6 Here]

4.8 Testing the Migration Impact

We recognize that our arguments are based on the assumption of a static model, where both the percentage of the Black population and exposure to the Tuskegee Study are time-invariant. As a result, a potential issue arises regarding changes in demographic composition as a result of migration. Nevertheless, we have conducted two tests to mitigate such concerns. We first perform a test to determine whether population migration, which alters the demographic structure and composition of counties over time, could potentially influence our findings. This is because we calculate the proportions of different racial groups aged 18 years and above at the county level using data from the 2010 Census as the foundation for our baseline estimation. To address this issue, we substitute the percentage of the Black population in 2010 with the percentage of the Black population between 1990 and 2005 and report the results in Table A4. We find that the results remain consistent and unaffected by this adjustment.

Next, we examine migration patterns within individual states and between different states by analyzing the 2021 American Community Survey (ACS) Data. We calculate the proportion of individuals relocating from different counties for all counties. We then identify the counties that fall within the top 5 and 10 percentiles of this distribution.[22] It

[22] A mover, in this context, refers to individuals who have changed their place of residence within the past year. Figure A5 illustrates these migration patterns.

is worth mentioning that it is not feasible for us to determine the migration direction or whether a county experiences a net influx or exodus of people. To ensure that our findings are not primarily influenced by migration, we have excluded counties with the highest percentile of residents relocating from different counties. The results are displayed in Table A5. Compared to the baseline findings, excluding counties that had the highest migration rates before the pandemic had minimal impact on our results, and the point estimates are very similar to those of the baseline.

5 Mechanism

5.1 Tuskegee Impact on Blacks' Vaccine Hesitancy

Relative to other races, African American men have worse health outcomes in the United States (Hayward and Heron, 1999; Do et al., 2008; Williams and Mohammed, 2009) and the COVID-19 pandemic has only made things worse.[23] But socioeconomic factors, such as lower labor income and education attainment, weaker labor market attachment, and lack of health insurance, do not fully account for these gaps (Cutler et al., 2008; Brunello et al., 2016) and a few empirical studies indicate that mistrust of healthcare institutions with historical roots also contributes to these inequities (Alsan and Wanamaker, 2018).

Alsan and Wanamaker (2018) discuss the disclosure of the Tuskegee Study fuels mistrust in modern medicine and public health among African Americans, especially males. As noted in the theory of social identification (Bhattacharya et al., 1995; Ashforth and Mael, 1989) and studies of empathy demonstrate that individuals are more responsive to injustices perpetrated against their group and more empathetic to individuals in closer "proximity" to themselves. Therefore, when much of the truth behind Tuskegee was revealed, mistrust among African Americans toward the medical profession and public health spiked.

Public health is concerned with promoting health, preventing disease, and enhancing the quality of life at the population level. Moreover, key public health models and frameworks underscore the fact that social conditions are powerful determinants of health within a given population (Kass, 2001). As outlined in the Public Health Code of Ethics, the "effectiveness of public health policies, practices, and actions depends upon public

[23] In fact, the U.S. Department of Health and Human Services documented racial/ethnic health disparities almost 35 years ago. As noted in Arias et al. (2021), African people live fewer years, on average, than White people. They are also more likely to die from treatable conditions; more likely to die during or after pregnancy and to suffer serious pregnancy-related complications, and more likely to lose children in infancy.

trust gained through decisions based on the highest ethical, scientific, and professional standards". Regarding COVID-19 disparities, a lack of trust in public health systems demands major attention as a result of persisting racism and socioeconomic inequity (Best et al., 2021).

To evaluate whether the lower vaccination rate among Black Americans we observe is driven by medical mistrust, we investigate racial differences in vaccine hesitancy throughout the study. We use survey data from the Household Pulse Survey (HPS) regarding the individual intention to vaccinate. In particular, we use the following HPS survey question, "once a vaccine to prevent COVID-19 is available to you, would you get a vaccine?". The responses are recorded in one of the five options, "definitely get a vaccine"; "probably get a vaccine"; "unsure"; "probably not get a vaccine"; or "definitely not get a vaccine".

A dummy variable is constructed to represent an individual's reluctance to receive a vaccine, which includes their responses of "definitely not" and "probably not". The county-level hesitation rate is computed in three steps using the approach described in Beleche et al. (2021). First, we use HPS responses to estimate hesitation rates at the state level. Then, combining the Census Bureau's 2019 American Community Survey 1-year Public Use Microdata Sample, we projected values to anticipate hesitation rates in more granular areas. Lastly, we use a PUMA-to-county crosswalk from the Missouri Census Data Center to create county-level estimates. In Figure 4, we plot the measure of hesitation. The vaccine hesitancy is positively correlated with the proportion of Blacks, which depicts the geographical distribution of Black population share by county in 2010 based on data from the American Community Survey.

[Insert Figure 4 Here]

We established two measures for vaccine hesitancy, specifically categorized as "strongly hesitant" and "hesitant." We partitioned the time frame into pre and post-April for two key reasons. Firstly, beginning in April 2021, the general population in most states gained widespread access to vaccinations, potentially impacting their views on vaccine acceptance. Secondly, the Health and Politics Survey (HPS) altered the way they measured vaccine resistance in their survey.[24] The findings in Table 7 indicate that the likelihood of being both strongly reluctant and reluctant to receive vaccination increases as the percentage of the Black population grows. Nevertheless, this effect is attenuated by the proximity to

[24]For the initial three months, the HPS survey did not include an "unsure" category in the question related to vaccine probability. However, in the April survey, the HPS survey introduced a modification by adding a new "unsure" category to the scales.

Tuskegee. However, the statistical significance is only observed for the reluctance not the strong reluctance. Our results lend support to the findings in Khan et al. (2021), which noted that Black people and people of other minorities are "unlikely or very unlikely" to take the COVID-19 vaccine.

[Insert Table 7 Here]

5.2 Ruling out Supply-Side Factors

It is worth noting that the research findings may also be subject to influence from supply-side factors, such as the availability of vaccines, the presence of medical professionals, and access to vaccination services, among others. To mitigate the effects of these factors, we incorporate the number of medical doctors for a county and its interaction with the time trend, as well as the number of pharmacy sites for a county and its interaction with the time trend (both pharmacy sites and doctor/nurses are normalized by the county population) in the regression.

Table A6 shows the findings after accounting for supply-side issues. When the point estimates for individuals who received at least one dose of vaccination are compared with the baseline, the point estimates are little modified. Similarly, increasing supply-side factors have a negligible impact on the rate of being completely vaccinated.

6 Heterogeneity

Our aforementioned analysis documents a strong correlation between the Tuskegee Study and heterogeneity in vaccination behavior in regions with varying proportions of Black residents. To gain a further understanding of our findings, we examine several types of heterogeneity by varying vaccine eligibility, political ideology, education, and income levels to shed light on the relevance of these forces.

6.1 Political Ideology and Time Variation

Vaccination is an effective tool for reducing the impact of COVID-19, but systemic inequities pose a serious threat to progress. Despite the greater availability of vaccines, racial discrepancies in vaccination persist. Vaccination uptake is substantially influenced by political ideology, and Republican voter support in the 2020 presidential election has a

negative impact on vaccination take-up (Agarwal et al., 2021).

We categorize counties into Democratic (or Republican) led counties based on the Democratic (or Republican) Party's majority voter share in the 2020 presidential election and report the results in Panel A of Table 8. The Tuskegee Study's influence is greater in Republican-dominated counties than in Democratic counties because many of the former are counties in the South that are closer to Macon, Alabama. Democratic counties had a wider vaccination gap between counties with high and low proportions of the Black population, which can be attributed to the higher Black population in Democratic counties.[25]

We assume that time has a linear effect on the vaccination rate in our main specification. Separating periods based on vaccine eligibility, though, might be intriguing. As shown in Panel B of Table 6, we divide time based on the date that the White House announces that individuals 16 years of age and older are eligible for vaccinations, which is before and after April 25, 2021.[26] The findings reveal that, before vaccine availability to the general public, there was a significant vaccination gap between counties with varying proportions of Black people. However, after the vaccine became eligible, the vaccination gap was largely closed. Tuskegee, on the other hand, plays a vital role in closing the immunization gap between regions.

[Insert Table 8 Here]

6.2 Education and Income

In Table 9, we examine whether the impact of the Tuskegee Study on vaccination varies by education and income level. Panel A compares the results for counties with lower and higher-educated residents. There is a notable distinction between counties in the upper and lower 50^{th} percentiles of the education distribution. The gap in vaccination rates between a higher-educated Black community and a White community with a comparable higher-educated Black population reduces more rapidly as the distance from Tuskegee increases. The findings align with the June 2021 report from the Centers for Disease Control and Prevention, which indicated that younger adults, non-Hispanic Black adults, individuals with lower educational attainment, no health insurance, and lower household

[25] In our data, the average percentage of Black people is 19.4% in Democratic areas and only 8% in Republican areas.

[26] As demonstrated in Figure A6, the distribution and availability of the COVID-19 vaccine for adults 16 and older vary by state. To alleviate the possibility of endogeneity resulting from vaccine adoption and vaccination behavior across states, we are utilizing the announcement date rather than the state distribution of vaccines.

incomes exhibited the lowest vaccination coverage rates and expressed less willingness to get vaccinated.[27]

Khairat and his research team, as discussed in their 2022 study (Khairat et al., 2022), delved into several factors and reasons contributing to the low uptake of COVID-19 vaccines in highly hesitant communities in the United States. They identified a low vaccination rate within these communities associated with factors such as lack of a high school education, and concerns regarding vaccine availability and distribution. The primary driver of vaccine hesitancy among these communities was a lack of trust in COVID-19 vaccines, followed by apprehensions about vaccine side effects and a lack of trust in the government. Additionally, individuals with lower levels of education tend to possess limited knowledge about vaccines and their distribution, which makes them more likely to express doubts about the effectiveness of the vaccine (Borga et al., 2022). Furthermore, individuals with lower educational levels often have limited access to transportation, which can hinder their ability to get vaccinated.

However, the impact of Tuskegee on reducing the vaccination disparity does not exhibit statistical significance in counties falling within both the top and bottom 50^{th} of the income distribution. One plausible explanation for this is that the COVID-19 vaccine is available to the general public at no cost, which should make access to it equally accessible in both affluent and economically disadvantaged counties.

[Insert Table 9 Here]

7 Concluding Remarks

As of June 21, 2022, more than 1 million individuals have died due to the COVID-19 pandemic in the U.S. alone,[28] and more than 14.9 million deaths worldwide.[29] The rate of vaccination has stagnated and fallen from the high in February 2021 (Diesel et al., 2021), despite the availability of the three types of vaccines in the United States and widespread immunization programs. Furthermore, there are variances in the vaccination rate by

[27] The report can be found here: COVID-19 Vaccination Coverage and Intent Among Adults Aged 18–39 Years — United States, March–May 2021.

[28] The U.S. cases and the death toll was collected by Johns Hopkins University Coronavirus Resource Center.

[29] New estimates from the World Health Organization show that the full death toll associated directly or indirectly with the COVID-19 pandemic (described as "excess mortality") was approximately 14.9 million (range 13.3 million to 16.6 million) globally. Data source: Department of Economics and Social Affairs, United Nations.

race across regions and throughout time. Among the many elements that could lead to vaccination resistance and refusal are political affiliation, cultural norms, the perception of COVID-19's threat, trust in the vaccine itself, and faith in governmental institutions (Bagasra et al., 2021; Bian et al., 2022).

By investigating the impact of a historical event and its lasting effects on the mistrust of the healthcare system among African Americans, our study adds to the existing literature on the factors contributing to the racial disparities in the vaccination rate. Our findings demonstrate that over time, the vaccination rate for Black communities converges with that of their White counterparts in the same county. The degree of convergence varies by region. In counties closer to Tuskegee, the racial gap in vaccination rates is closing at a slower pace, likely because of the scars from the Tuskegee Study. Our study points to a spatial distribution of the intergenerational impact of historical trauma, and the need to rebuild the trust among African Americans who continue to harbor fears of the racist past of the United States.

Data availability

The author confirms that all data generated or analyzed during this study are included in this published article. Furthermore, primary and secondary sources and data supporting the findings of this study were all publicly available at the time of publication.

References

Aburto, J. M., Tilstra, A. M., Floridi, G. and Dowd, J. B. (2022), 'Significant impacts of the covid-19 pandemic on race/ethnic differences in us mortality', *Proceedings of the National Academy of Sciences* **119**(35), e2205813119.

Agarwal, R., Dugas, M., Ramaprasad, J., Luo, J., Li, G. and Gao, G. G. (2021), 'Socioeconomic privilege and political ideology are associated with racial disparity in covid-19 vaccination', *Proceedings of the National Academy of Sciences* **118**(33).

Albright, A., Cook, J. A., Feigenbaum, J. J., Kincaide, L., Long, J. and Nunn, N. (2021), After the burning: the economic effects of the 1921 tulsa race massacre, Technical report, National Bureau of Economic Research.

Algan, Y. and Cahuc, P. (2010), 'Inherited trust and growth', *American Economic Review* **100**(5), 2060–92.

Alsan, M., Garrick, O. and Graziani, G. (2019), 'Does diversity matter for health? experimental evidence from oakland', *American Economic Review* **109**(12), 4071–4111.

Alsan, M. and Wanamaker, M. (2018), 'Tuskegee and the health of black men', *The quarterly journal of economics* **133**(1), 407–455.

Altonji, J. G., Elder, T. E. and Taber, C. R. (2005), 'Selection on observed and unobserved variables: Assessing the effectiveness of catholic schools', *Journal of political economy* **113**(1), 151–184.

Arias, E., Tejada-Vera, B., Ahmad, F. and Kochanek, K. D. (2021), 'Provisional life expectancy estimates for 2020'.

Ashforth, B. E. and Mael, F. (1989), 'Social identity theory and the organization', *Academy of management review* **14**(1), 20–39.

Baack, B. N., Abad, N., Yankey, D., Kahn, K. E., Razzaghi, H., Brookmeyer, K., Kolis, J., Wilhelm, E., Nguyen, K. H. and Singleton, J. A. (2021), 'Covid-19 vaccination coverage and intent among adults aged 18–39 years—united states, march–may 2021', *Morbidity and Mortality Weekly Report* **70**(25), 928.

Bagasra, A. B., Doan, S. and Allen, C. T. (2021), 'Racial differences in institutional trust and covid-19 vaccine hesitancy and refusal', *BMC Public Health* **21**(1), 1–7.

Bazzi, S., Fiszbein, M. and Gebresilasse, M. (2020), 'Frontier culture: The roots and persistence of "rugged individualism" in the united states', *Econometrica* **88**(6), 2329–2368.

Beleche, T., Ruhter, J., Kolbe, A., Marus, J., Bush, L. and Sommers, B. (2021), 'Covid-19 vaccine hesitancy: Demographic factors, geographic patterns, and changes over time'.

Best, A. L., Fletcher, F. E., Kadono, M. and Warren, R. C. (2021), 'Institutional distrust

among african americans and building trustworthiness in the covid-19 response: implications for ethical public health practice', *Journal of health care for the poor and underserved* **32**(1), 90.

Bhattacharya, C. B., Rao, H. and Glynn, M. A. (1995), 'Understanding the bond of identification: An investigation of its correlates among art museum members', *Journal of marketing* **59**(4), 46–57.

Bian, B., Li, J., Xu, T. and Foutz, N. Z. (2022), 'Individualism during crises', *Review of Economics and Statistics* **104**(2), 368–385.

Borga, L. G., Clark, A. E., D'Ambrosio, C. and Lepinteur, A. (2022), 'Characteristics associated with covid-19 vaccine hesitancy', *Scientific Reports* **12**(1), 12435.

Brandt, A. M. (1978), 'Racism and research: the case of the tuskegee syphilis study', *Hastings center report* pp. 21–29.

Bruhin, A., Goette, L., Haenni, S. and Jiang, L. (2020), 'Spillovers of prosocial motivation: Evidence from an intervention study on blood donors', *Journal of health economics* **70**, 102244.

Brunello, G., Fort, M., Schneeweis, N. and Winter-Ebmer, R. (2016), 'The causal effect of education on health: What is the role of health behaviors?', *Health economics* **25**(3), 314–336.

Corbie-Smith, G., Thomas, S. B., Williams, M. V. and Moody-Ayers, S. (1999), 'Attitudes and beliefs of african americans toward participation in medical research', *Journal of general internal medicine* **14**(9), 537–546.

Cubbin, C., Pickle, L. W. and Fingerhut, L. (2000), 'Social context and geographic patterns of homicide among us black and white males.', *American Journal of Public Health* **90**(4), 579.

Cutler, D. M., Lleras-Muney, A. and Vogl, T. (2008), 'Socioeconomic status and health: dimensions and mechanisms', *National Bureau of Economic Research* .

Diesel, J., Sterrett, N., Dasgupta, S., Kriss, J. L., Barry, V., Esschert, K. V., Whiteman, A., Cadwell, B. L., Weller, D., Qualters, J. R. et al. (2021), 'Covid-19 vaccination coverage among adults—united states, december 14, 2020–may 22, 2021', *Morbidity and Mortality Weekly Report* **70**(25), 922.

Do, D. P., Finch, B. K., Basurto-Davila, R., Bird, C., Escarce, J. and Lurie, N. (2008), 'Does place explain racial health disparities? quantifying the contribution of residential context to the black/white health gap in the united states', *Social science & medicine* **67**(8), 1258–1268.

Dupas, P. and Miguel, E. (2017), Impacts and determinants of health levels in low-income

countries, *in* 'Handbook of economic field experiments', Vol. 2, Elsevier, pp. 3–93.

Gentzkow, M. and Shapiro, J. M. (2006), 'Media bias and reputation', *Journal of political Economy* **114**(2), 280–316.

Hayward, M. D. and Heron, M. (1999), 'Racial inequality in active life among adult americans', *Demography* **36**(1), 77–91.

Iyengar, S. and Hahn, K. S. (2009), 'Red media, blue media: Evidence of ideological selectivity in media use', *Journal of communication* **59**(1), 19–39.

Kass, N. E. (2001), 'An ethics framework for public health', *American journal of public health* **91**(11), 1776–1782.

Khairat, S., Zou, B. and Adler-Milstein, J. (2022), 'Factors and reasons associated with low covid-19 vaccine uptake among highly hesitant communities in the us', *American journal of infection control* **50**(3), 262–267.

Khan, M. S., Ali, S. A. M., Adelaine, A. and Karan, A. (2021), 'Rethinking vaccine hesitancy among minority groups', *The Lancet* **397**(10288), 1863–1865.

Knack, S. and Keefer, P. (1997), 'Does social capital have an economic payoff? a cross-country investigation', *The Quarterly journal of economics* **112**(4), 1251–1288.

Lowes, S. and Montero, E. (2021), 'The legacy of colonial medicine in central africa', *American Economic Review* **111**(4), 1284–1314.

McLaughlin, J. M., Khan, F., Pugh, S., Swerdlow, D. L. and Jodar, L. (2022), 'County-level vaccination coverage and rates of covid-19 cases and deaths in the united states: An ecological analysis', *The Lancet Regional Health-Americas* **9**, 100191.

Murthy, B. P., Sterrett, N., Weller, D., Zell, E., Reynolds, L., Toblin, R. L., Murthy, N., Kriss, J., Rose, C., Cadwell, B. et al. (2021), 'Disparities in covid-19 vaccination coverage between urban and rural counties—united states, december 14, 2020–april 10, 2021', *Morbidity and Mortality Weekly Report* **70**(20), 759.

Ndugga, N., Pham, O., Hill, L., Artiga, S. and Mengistu, S. (2021), 'Latest data on covid-19 vaccinations race/ethnicity', *Kais Family Found* .

Nguyen, L. H., Joshi, A. D., Drew, D. A., Merino, J., Ma, W., Lo, C.-H., Kwon, S., Wang, K., Graham, M. S., Polidori, L. et al. (2021), 'Racial and ethnic differences in covid-19 vaccine hesitancy and uptake', *MedRxiv* .

Nunn, N. and Wantchekon, L. (2011), 'The slave trade and the origins of mistrust in africa', *American Economic Review* **101**(7), 3221–52.

Tabellini, G. (2010), 'Culture and institutions: economic development in the regions of europe', *Journal of the European Economic association* **8**(4), 677–716.

Williams, D. R. and Mohammed, S. A. (2009), 'Discrimination and racial disparities in

health: evidence and needed research', *Journal of behavioral medicine* **32**(1), 20–47.
Yancy, C. W. (2020), 'Covid-19 and african americans', *Jama* **323**(19), 1891–1892.

Figures and Tables

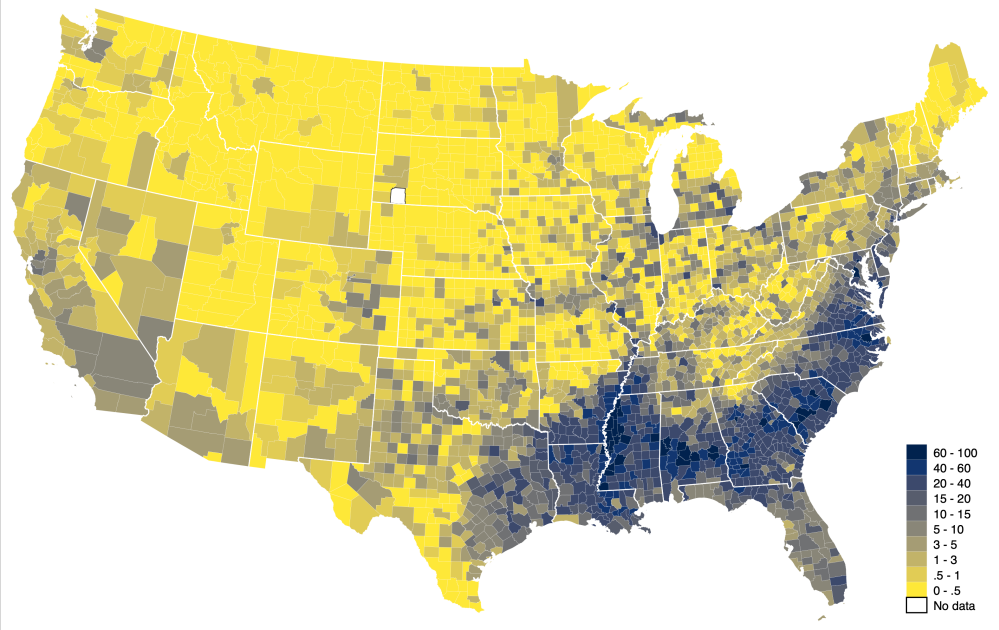
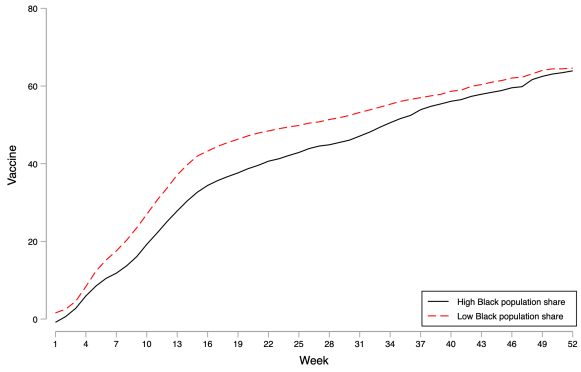
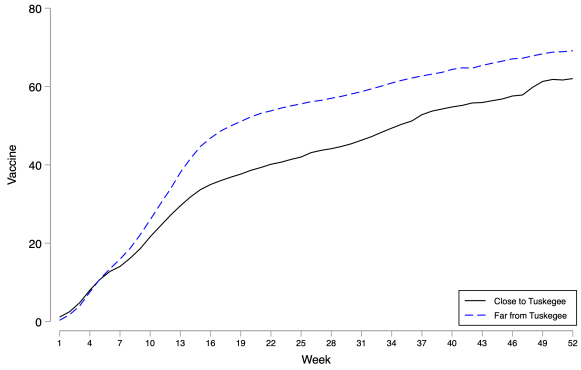


Figure 1: Share of Black population, 2010

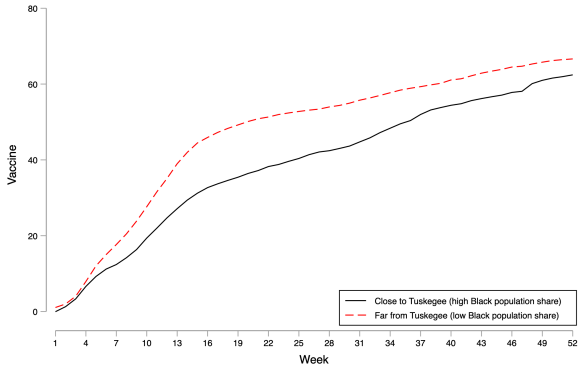
Data Source: U.S. Census Bureau. 2010 Census Redistricting Data.
Note: Darker shaded areas represent counties with higher Black population shares.



(a) Share of Black population



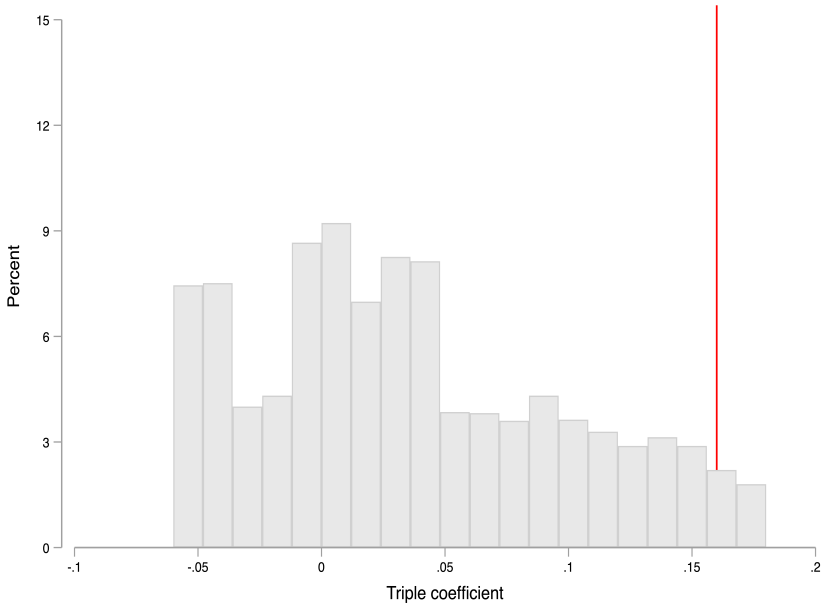
(b) Distance to Tuskegee



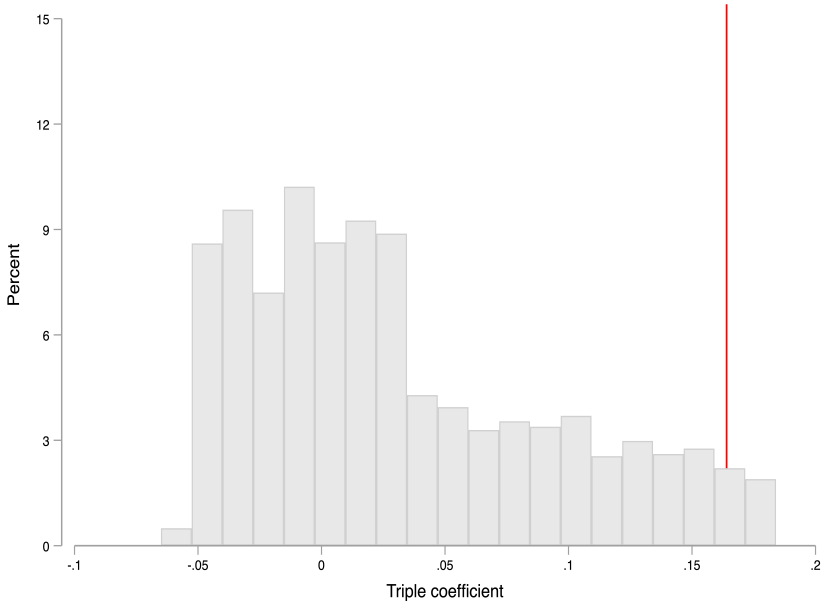
(c) Interaction of distance and Black population share

Figure 2: Vaccination coverage between week 1 and week 52, 2021

Note: The vaccination rate is calculated for the percentage of the population over 18 years old who received at least one vaccination. A low (or high) share county is one where the proportion of Black individuals is in the bottom 30^{th} (or top 30^{th}) percentiles of the distribution. If the distance is in the bottom 50^{th} percentile of the distribution, a county is considered "close"; otherwise, it is considered a "far" county.



(a) At least one dose



(b) Fully vaccinated

Figure 3: Permutation tests

Note: Frequency of the false β_1 coefficient estimated using distance from every other county in the sample and estimating Equation (1). Panel (a) displays the results for at least one dose and the results of fully vaccinated are presented in Panel (b). The vertical line denotes β_1 from baseline estimates using the true treatment distance (to Macon County) as reported in Table 1.

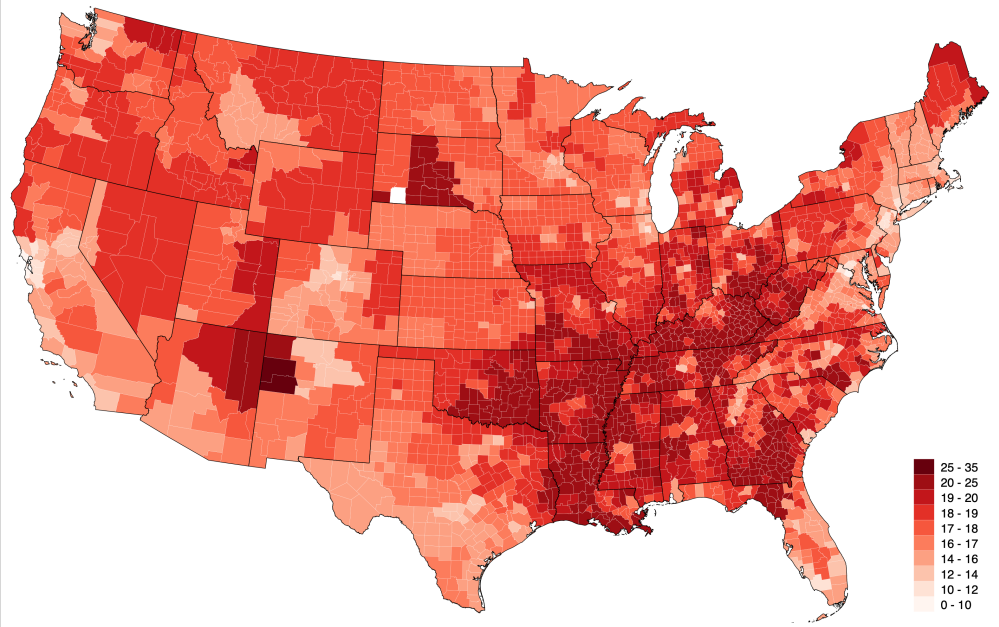


Figure 4: County-level hesitancy of COVID-19 vaccine

Data Source: Household Pulse Survey.
Note: Darker shaded areas represent a higher resistance to receiving a vaccine.

Table 1: Balance table

	Close to Tuskegee with high Black share	Far from Tuskegee with low Black share	p-value
Panel A: Baseline			
% Black aged 18 years and older	0.268	0.002	0.000
% White aged 18 years and older	0.655	0.885	0.000
% Hispanic aged 18 years and older	0.049	0.069	0.001
% Asian aged 18 years and older	0.011	0.004	0.000
% Population (25+) with no high school diploma	0.196	0.122	0.000
% Population (25+) with high school diploma	0.346	0.352	0.082
% Population (25+) with some college	0.276	0.331	0.000
% Population (25+) with college degree and above	0.182	0.194	0.001
Unemployment rate (2020)	7.168	5.571	0.000
Panel B: Supply side			
# of pharmacies per 100k population	19.42	7.71	0.000
# of primary care physicians per 100k population	43.07	43.16	0.963
# of physician assistants with NPI per 100k population	23.48	37.90	0.000
# of advanced practice registered nurses per 100k population	93.59	64.98	0.000
# of nurse practitioners per 100k population	78.25	54.37	0.000
Observations	787	718	1505

Notes: A low (or high) share county is one where the proportion of Black individuals is in the bottom 30^{th} (or top 30^{th}) percentiles of the distribution. If the distance is in the bottom 50^{th} percentile of the distribution, a county is considered "close"; otherwise, it is considered a "far" county. The p-value is for the test of the null hypothesis that the means across the samples are the same.

Table 2: Impact of Tuskegee Study on vaccination rate

	Baseline results		County-month fixed effect		County × week linear trend	
	At least one dose (1)	Fully vaccinated (2)	At least one dose (3)	Fully vaccinated (4)	At least one dose (5)	Fully vaccinated (6)
Share of Black × distance × week	0.160*	0.164**	0.158*	0.151**	0.204***	0.206***
	(0.087)	(0.077)	(0.085)	(0.067)	(0.069)	(0.055)
Share of Black × week	-1.200***	-1.442***	-0.979***	-1.386***	-0.576	-0.825***
	(0.383)	(0.398)	(0.354)	(0.342)	(0.406)	(0.301)
Distance × week	-0.006	-0.005	-0.001	-0.003	0.009**	0.009**
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	134,259	143,666	133,971	143,533	133,971	143,533
R-squared	0.966	0.970	0.999	0.999	0.999	0.999
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes	No	No
Share of Hispanic × week	Yes	Yes	Yes	Yes	No	No
Share of HS grads and above × week	Yes	Yes	Yes	Yes	No	No
County unemployment rate × week	Yes	Yes	Yes	Yes	No	No
County-month FE	No	No	Yes	Yes	Yes	Yes
County × week linear trend	No	No	No	No	Yes	Yes

Notes: The vaccination rate is calculated for the percentage of the population over 18 years old who received at least one vaccination. The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table 3: Tuskegee impact by gender and age

	18 - 24	25 - 34	35 - 44	45 - 54	55 - 64	65 and older
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Full population						
Share of Black× distance × week	0.923*	0.846**	0.918**	0.833**	1.079**	1.131**
	(0.502)	(0.390)	(0.373)	(0.387)	(0.504)	(0.471)
Share of Black × week	-8.261***	-7.796***	-8.794***	-6.785***	-7.455***	-5.108**
	(1.913)	(1.995)	(2.352)	(2.028)	(2.370)	(2.464)
Distance × week	-0.004	-0.006	-0.007	-0.004	-0.002	0.001
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	143,666	143,666	143,666	143,666	143,666	143,666
R-squared	0.970	0.971	0.971	0.970	0.970	0.970
Panel B: Males						
Share of Black × distance × week	1.817*	1.645*	1.855**	1.712*	2.298*	2.703**
	(1.046)	(0.872)	(0.850)	(0.866)	(1.175)	(1.266)
Share of Black × week	-17.403***	-15.192***	-18.547***	-15.008***	-16.153***	-13.122*
	(4.073)	(4.057)	(5.200)	(4.524)	(5.370)	(6.689)
Distance × week	-0.004	-0.005	-0.007	-0.005	-0.002	0.001
	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Observations	143,666	143,666	143,666	143,666	143,666	143,666
R-squared	0.970	0.970	0.971	0.970	0.970	0.970
Panel C: Females						
Share of Black × distance × week	1.854*	1.668**	1.755***	1.572**	2.001**	1.910**
	(0.934)	(0.671)	(0.653)	(0.683)	(0.874)	(0.743)
Share of Black × week	-14.873***	-13.736***	-15.227***	-11.906***	-13.668***	-8.383**
	(3.496)	(3.563)	(4.026)	(3.561)	(4.194)	(3.862)
Distance × week	-0.003	-0.005	-0.006	-0.003	-0.002	0.002
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Observations	143,666	143,666	143,666	143,666	143,666	143,666
R-squared	0.970	0.971	0.971	0.970	0.970	0.970
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes	Yes	Yes
Share of Hispanic × week	Yes	Yes	Yes	Yes	Yes	Yes
Share of HS grads and above × week	Yes	Yes	Yes	Yes	Yes	Yes
County unemployment rate × week	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each column presents the coefficient estimates from a separate regression. The outcome is the percentage of adults aged 18 years and older fully vaccinated. For each column, the Black share is calculated by dividing the number of Black individuals in a specific age group by the total population of the county. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table 4: Robustness checks (other races)

	White/Hispanic/Asian Americans		African Americans	
	At least one dose	Fully vaccinated	At least one dose	Fully vaccinated
	(1)	(2)	(3)	(4)
Panel A: The Tuskegee impact on Whites				
Population share by race × distance × week	-0.004	-0.005	0.160*	0.164**
	(0.014)	(0.016)	(0.087)	(0.077)
Population share by race × week	-0.440	-0.377	-1.200***	-1.442***
	(0.305)	(0.293)	(0.383)	(0.398)
Distance × week	0.007	0.010	-0.006	-0.005
	(0.012)	(0.013)	(0.004)	(0.005)
Observations	134,259	143,666	134,259	143,666
R-squared	0.965	0.969	0.966	0.970
County FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Share of White × week	No	No	Yes	Yes
Share of Black × week	Yes	Yes	No	No
Share of Hispanic × week	Yes	Yes	Yes	Yes
Share of HS grads and above × week	Yes	Yes	Yes	Yes
County unemployment rate × week	Yes	Yes	Yes	Yes
Panel B: The Tuskegee impact on Hispanics (exclude Arizona, California, Colorado, New Mexico, and Texas)				
Population share by race × distance × week	-0.030	-0.059	0.200**	0.197***
	(0.063)	(0.072)	(0.085)	(0.069)
Population share by race × week	0.787	0.941	-1.278***	-1.409***
	(0.550)	(0.634)	(0.426)	(0.511)
Distance × week	0.010	0.016*	0.000	0.003
	(0.008)	(0.009)	(0.004)	(0.004)
Observations	126,132	134,382	126,132	134,382
R-squared	0.963	0.968	0.964	0.969
County FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes
Share of Black × week	Yes	Yes	No	No
Share of Hispanic × week	No	No	Yes	Yes
Share of HS grads and above × week	Yes	Yes	Yes	Yes
County unemployment rate × week	Yes	Yes	Yes	Yes
Panel C: The Tuskegee impact on Asians (exclude California, Connecticut, Washington, D.C, Maryland, and New Jersey)				
Population share by race × distance × week	0.029	0.019	0.207**	0.214***
	(0.094)	(0.102)	(0.087)	(0.071)
Population share by race × week	0.697	1.275	-1.276***	-1.493***
	(1.317)	(1.537)	(0.447)	(0.522)
Distance × week	0.001	0.004	-0.003	-0.001
	(0.006)	(0.006)	(0.005)	(0.005)
Observations	129,527	138,934	129,527	138,934
R-squared	0.962	0.967	0.964	0.969
County FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes
Share of Black × week	Yes	Yes	No	No
Share of Hispanic × week	No	No	Yes	Yes
Share of HS grads and above × week	Yes	Yes	Yes	Yes
County unemployment rate × week	Yes	Yes	Yes	Yes

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. Hispanic and Asian Americans are highly concentrated in a few states in the U.S. In Panel B, we remove the top 5 most concentrated states for Hispanics and in Panel C, we remove the top 5 most concentrated states for Asians. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table 5: Robustness checks for other concerns

	At least one dose (1)	Fully vaccinated (2)
Panel A: Alternative measurement (newspaper exposure)		
Share of Black × Tuskegee news × week	41.983	42.415
	(54.367)	(49.439)
Share of Black × week	-0.854**	-1.244***
	(0.382)	(0.362)
Tuskegee news × week	-8.883***	-9.308***
	(1.319)	(2.421)
Observations	133,968	143,345
R-squared	0.965	0.969
Panel B: South region		
Share of Black × distance × week	0.190**	0.230**
	(0.089)	(0.082)
Share of Black × week	-0.161	-0.515
	(0.582)	(0.737)
Distance × week	0.023	0.011
	(0.022)	(0.029)
Observations	52,950	55,338
R-squared	0.963	0.964
Panel C: Tulsa impact		
Share of Black × distance × week	0.060	0.050
	(0.038)	(0.039)
Share of Black × week	-0.756	-0.963
	(0.487)	(0.611)
Distance × week	0.019***	0.020***
	(0.005)	(0.005)
Observations	134,259	143,666
R-squared	0.967	0.971
County FE	Yes	Yes
Week FE	Yes	Yes
Share of White × week	Yes	Yes
Share of Hispanic × week	Yes	Yes
Share of HS grads and above × week	Yes	Yes
County unemployment rate × week	Yes	Yes

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table 6: Using selection on observables to assess the bias from unobservables

Controls in the restricted set	Controls in the full set	At least one dose		Fully vaccinated	
		Ratio	Standard error	Ratio	Standard error
None	Full set of controls from equation (1)	21.861	2.165	21.097	1.898
None	Full set of controls from equation (1), and non-linear trend	20.704	1.989	20.677	1.865
Share of White and share of Hispanic	Full set of controls from equation (1)	40.151	7.048	29.275	3.677
Share of White and share of Hispanic	Full set of controls from equation (1) and non-linear trend	37.043	6.152	28.553	3.588
Share of White and share of Hispanic	Full set of controls from equation (1), non-linear trend, and supply-side factors	22.603	2.675	18.219	1.684

Notes: In the table, each cell represents a ratio based on the coefficients from two regressions for share of Black × distance × week. One includes a "restricted set" of covariates for the control variables and its coefficient is called β^R. The other regression includes a "full set" of controls and its coefficient is called β^F. The sample sizes are the same in both regressions and county and week fixed effects are included. Here is how we calculate the ratio: $\beta^F/(\beta^R - \beta^F)$. In Table 1, the full set of controls from Equation (1) is described.

Table 7: Impact of Tuskegee Study on vaccination hesitancy

	Strongly hesitant		Hesitant	
	Jan. - Mar. (1)	April - Sept. (2)	Jan. - Mar. (3)	April - Sept. (4)
Share of Black × distance × week	-0.009	-0.001	-0.030**	-0.013**
	(0.006)	(0.003)	(0.014)	(0.005)
Share of Black × week	1.712***	0.315***	1.551***	0.513***
	(0.332)	(0.088)	(0.451)	(0.118)
Distance × week	-0.000	-0.001***	-0.002	-0.002**
	(0.001)	(0.000)	(0.001)	(0.001)
Observations	12,428	24,856	12,428	24,856
R-squared	0.968	0.974	0.982	0.985
County FE	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes
Share of Hispanic × week	Yes	Yes	Yes	Yes
Share of Asian × week	Yes	Yes	Yes	Yes

Notes: To measure vaccine hesitancy, we use the following HPS survey question, “once a vaccine to prevent COVID-19 is available to you, would you get a vaccine?”. Five responses are recorded, including “definitely get a vaccine”; “probably get a vaccine”; “unsure”; “probably not get a vaccine”; “definitely not get a vaccine.” Strongly hesitant is defined as definitely not, and Hesitant is defined as definitely not and probably not. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table 8: Heterogeneity by political ideology and eligibility timeline

Panel A: Political ideology	At least one dose			Fully vaccinated		
	Democratic (1)	Republican (2)	Cross-Model Difference (p) (3)	Democratic (4)	Republican (5)	Cross-Model Difference (p) (6)
Share of Black × distance × week	0.167*	0.197*	0.057	0.178**	0.222**	0.004
	(0.089)	(0.114)		(0.079)	(0.101)	
Share of Black × week	-1.749***	-0.725	0.161	-2.145***	-0.609	0.404
	(0.466)	(0.503)		(0.493)	(0.494)	
Distance × week	-0.013***	-0.006	0.112	-0.014**	-0.003	0.040
	(0.005)	(0.004)		(0.006)	(0.004)	
Observations	23,012	111,143		23,909	119,653	
R-squared	0.975	0.961		0.979	0.963	
Panel B: Vaccine eligibility						
	At least one dose			Fully vaccinated		
	Before April 25 (1)	After April 25 (2)	Cross-Model Difference (p) (3)	Before April 25 (4)	After April 25 (5)	Cross-Model Difference (p) (6)
Share of Black × distance × week	0.345*	0.043	0.017	0.224	0.042	0.079
	(0.195)	(0.052)		(0.138)	(0.040)	
Share of Black × week	-3.660***	0.315	0.008	-2.551***	-0.145	0.0003
	(0.804)	(0.300)		(0.664)	(0.313)	
Distance × week	0.001	-0.007*	0.038	-0.001	-0.007*	0.035
	(0.011)	(0.004)		(0.006)	(0.004)	
Observations	41,340	92,918		42,296	101,370	
R-squared	0.960	0.974		0.953	0.976	
County FE	Yes	Yes		Yes	Yes	
Week FE	Yes	Yes		Yes	Yes	
Share of White × week	Yes	Yes		Yes	Yes	
Share of Hispanic × week	Yes	Yes		Yes	Yes	
Share of HS grads and above × week	Yes	Yes		Yes	Yes	
County unemployment rate × week	Yes	Yes		Yes	Yes	

Notes: The White House announced that all people aged 16 and older are eligible for the COVID-19 vaccine on April 19, 2021. The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table 9: Heterogeneity by education and income level

	At least one dose			Fully vaccinated		
	Bottom 50% (1)	Top 50% (2)	Cross-Model Difference (p) (3)	Bottom 50% (4)	Top 50% (5)	Cross-Model Difference (p) (6)
Panel A: Education level						
Share of Black × distance × week	0.105	0.216*	0.003	0.104	0.223**	0.001
	(0.071)	(0.110)		(0.064)	(0.103)	
Share of Black × week	-1.097**	-1.289***	0.007	-1.236**	-1.807***	0.000
	(0.466)	(0.449)		(0.542)	(0.390)	
Distance × week	-0.010**	-0.006	0.091	-0.015**	-0.004	0.014
	(0.005)	(0.004)		(0.006)	(0.004)	
Observations	65,222	69,037		68,103	75,563	
R-squared	0.966	0.968		0.972	0.973	
Panel B: Income level						
Share of Black × distance × week	0.144**	0.159	0.138	0.131**	0.173*	0.151
	(0.064)	(0.103)		(0.057)	(0.097)	
Share of Black × week	-0.579	-1.525***	0.023	-0.522	-2.041***	0.017
	(0.397)	(0.473)		(0.421)	(0.455)	
Distance × week	0.001	-0.006	0.058	0.002	-0.006	0.005
	(0.005)	(0.004)		(0.006)	(0.004)	
Observations	68,164	66,095		72,590	71,076	
R-squared	0.958	0.972		0.962	0.976	
County FE	Yes	Yes		Yes	Yes	
Week FE	Yes	Yes		Yes	Yes	
Share of White × week	Yes	Yes		Yes	Yes	
Share of Hispanic × week	Yes	Yes		Yes	Yes	
Share of HS grads and above × week	Yes	Yes		Yes	Yes	
County unemployment rate × week	Yes	Yes		Yes	Yes	

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. The level of education is determined based on the share of high school graduates and above at the county level. The median income at the county level determines the income level. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Online Appendix for

"The Lasting Impact of the Tuskegee Syphilis Study: COVID-19 Vaccination Hesitation among African Americans"

Xiaolong Hou[#*], Yang Jiao[†] Leilei Shen[‡], and Zhuo Chen[§¶]

[#]Xiaolong Hou and Yang Jiao contributed equally to this work.

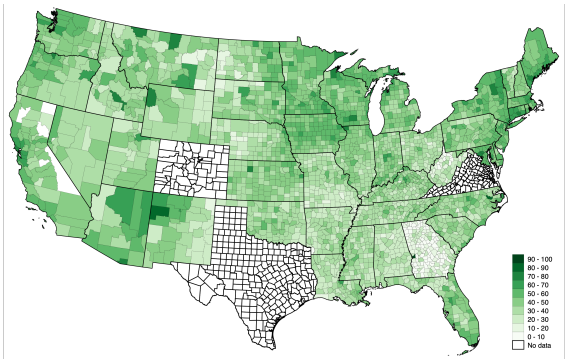
[*]John Munro Godfrey, Sr. Department of Economics, University of Georgia. Email: xh54056@uga.edu.

[†]Department of Accounting, Economics, and Finance, Texas A&M University-Texarkana. Email: yjiao0207gmail.com.

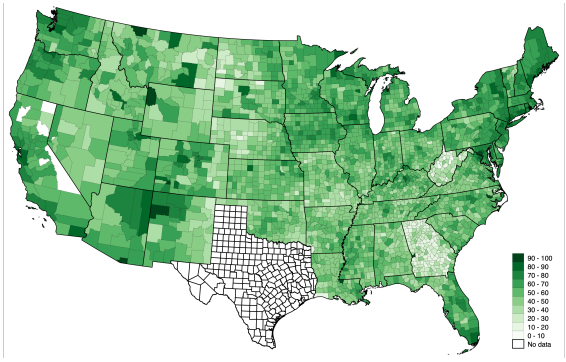
[‡]Senior Data Scientist. Airbnb. Email: Leilei.shen@airbnb.com.

[§]Department of Health Policy and Management, University of Georgia. Email: zchen1@uga.edu.

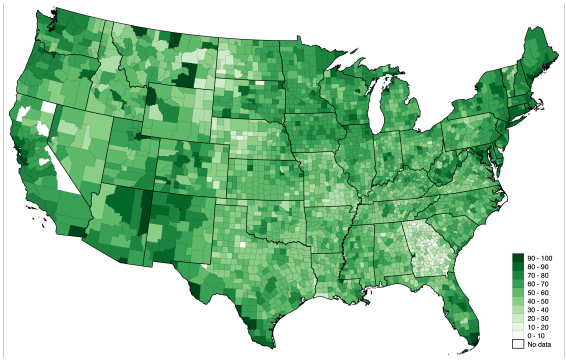
[¶]School of Economics. University of Nottingham Ningbo China.



(a) Percentage of vaccinated population, May 17–23 2021



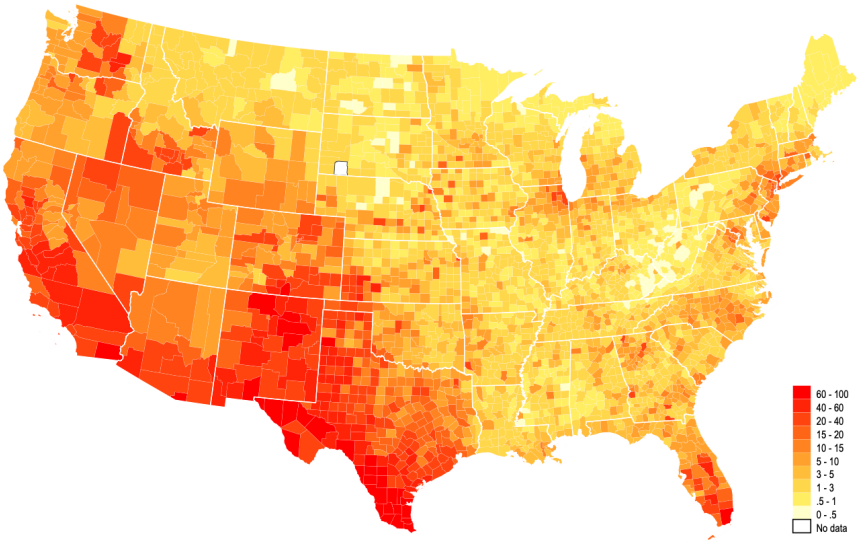
(b) Percentage of vaccinated population, Oct 4–10 2021



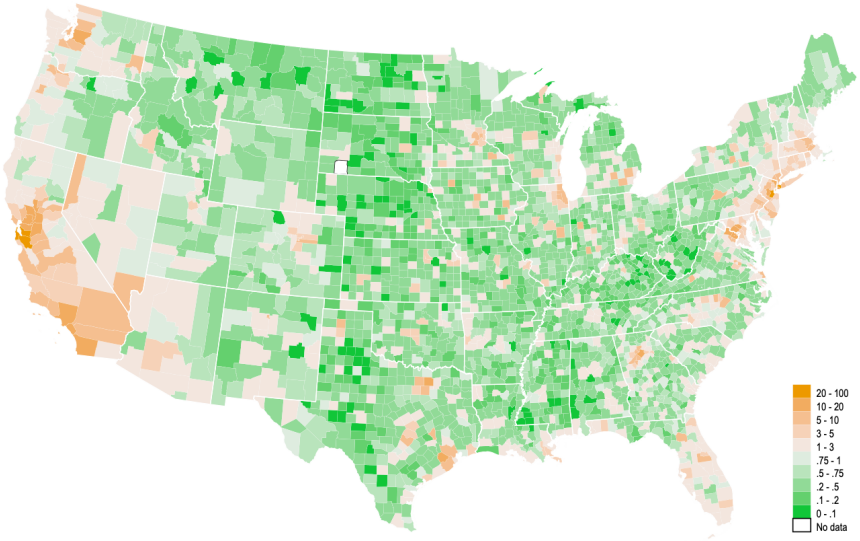
(c) Percentage of vaccinated population, Dec. 13–19 2021

Figure A1: COVID-19 vaccine rates by county and over time

Data Source: U.S. Centers for Disease Control and Prevention.
Note: Darker shaded areas represent counties with higher vaccination coverage. CDC has capped vaccination coverage estimates on the COVID Data Tracker at 95 percent.



(a) Share of Hispanic population, 2010



(b) Share of Asian population, 2010

Figure A2: Geographic distribution of Hispanic and Asian population

Data Source: U.S. Census Bureau. 2010 Census Redistricting Data.

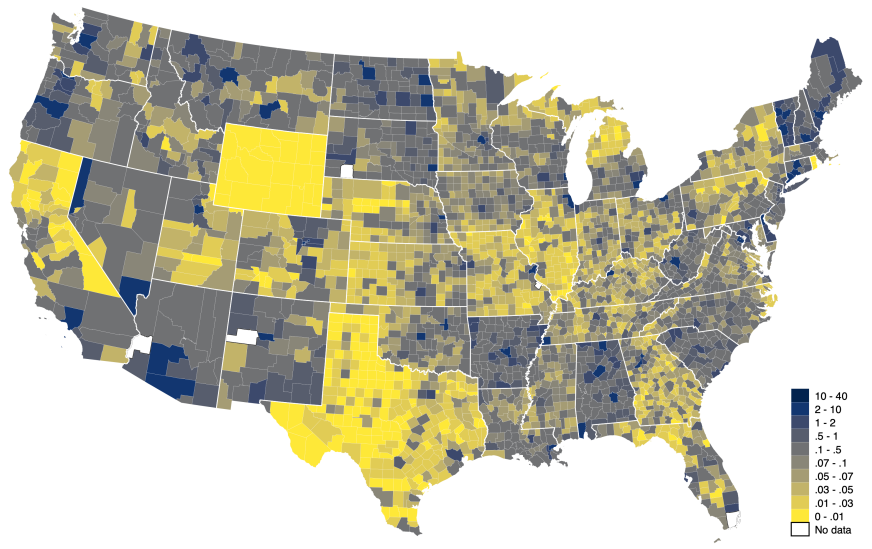


Figure A3: County exposure to Tuskegee news

Note: Exposure to Tuskegee news is computed based on Equation 2.

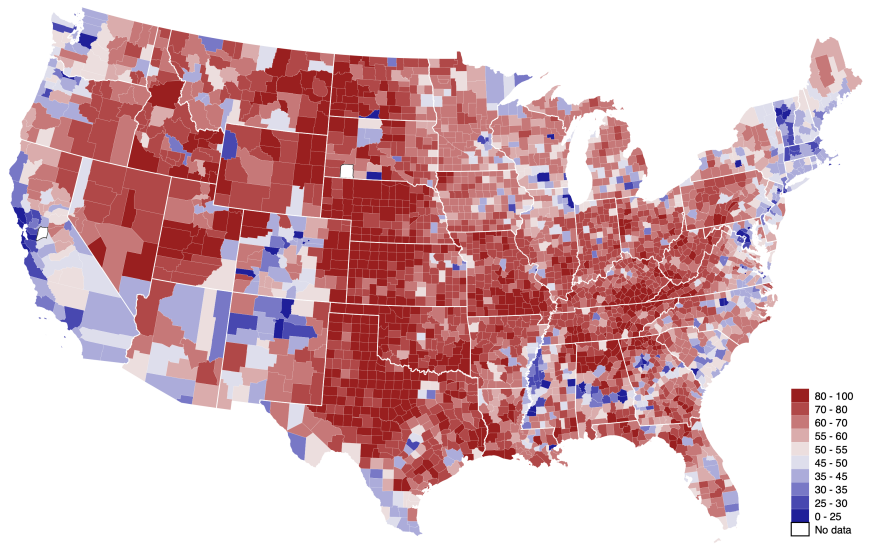
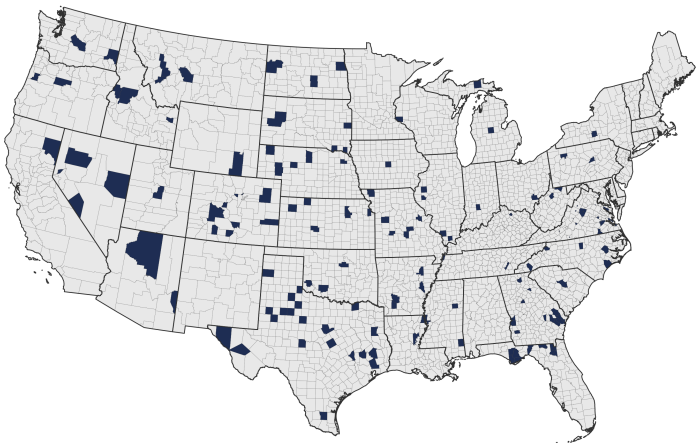
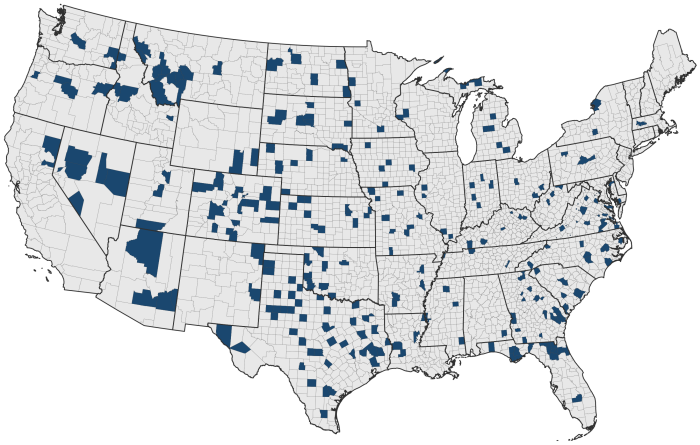


Figure A4: 2020 U.S. Presidential Election map by county & vote share

Note: Figure A4 shows the county level and vote share results of the 2020 U.S. Presidential Election. The darker the blue, the more Democratic a county voted, and the darker the red, the more Republican a county voted.



(a) Top 5% of counties by resident mobility in the past year



(b) Top 10% of counties by resident mobility in the past year

Figure A5: Geographical mobility in the past year for current residence

Data Source: 2021 American Community Survey (ACS)
Note: Darker shaded areas correspond to counties exhibiting higher levels of resident mobility in the past year.

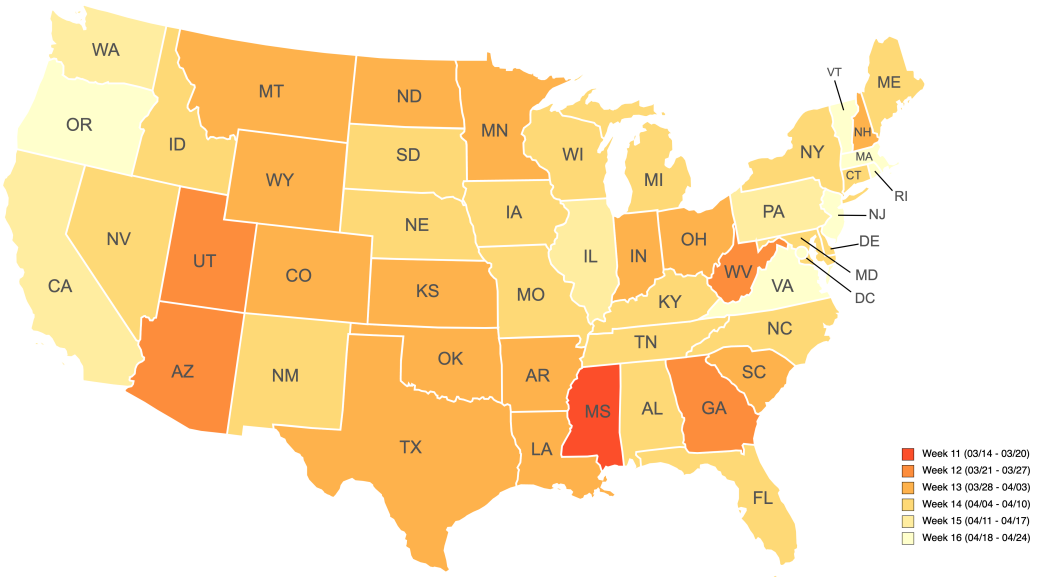


Figure A6: Earliest date of vaccine eligibility for adults age 16 and older by state

Note: Information of COVID-19 vaccination rates and distribution plans by state is collected from vaccine distribution plan.

Table A1: Summary statistics - vaccine coverage

Variable	Mean	Std. Dev.	Min.	Max.	N
Percentage of population age 18 and older receiving at least one dose					
01/30/2021(week 4)	7.69	4.86	0.00	46.80	2643
02/27/2021(week 8)	17.29	7.34	0.00	73.50	2628
03/27/2021(week 12)	30.22	10.04	0.70	95.00	2549
04/24/2021(week 16)	40.33	12.99	0.90	95.00	2437
05/29/2021(week 21)	45.72	14.73	3.20	95.00	2388
06/26/2021(week 25)	48.40	14.80	4.10	95.00	2312
07/31/2021(week 30)	51.36	15.16	5.60	95.00	2350
08/28/2021(week 34)	54.94	15.51	7.30	95.00	2437
09/25/2021(week 38)	58.06	15.21	7.70	95.00	2619
10/30/2021(week 43)	60.65	15.04	8.00	95.00	2966
11/27/2021(week 47)	62.58	15.21	11.20	95.00	2960
12/31/2021(week 52)	65.63	14.09	11.90	95.00	2900
Percentage of population age 18 and older fully vaccinated					
01/30/2021(week 4)	1.68	1.40	0.00	25.40	2643
02/27/2021(week 8)	8.53	4.28	0.00	56.50	2643
03/27/2021(week 12)	18.40	7.04	0.50	81.30	2643
04/24/2021(week 16)	30.89	10.26	0.60	95.00	2645
05/29/2021(week 21)	39.80	13.21	1.70	95.00	2645
06/26/2021(week 25)	42.85	13.94	3.00	95.00	2645
07/31/2021(week 30)	45.06	14.27	4.00	95.00	2645
08/28/2021(week 34)	47.42	14.46	4.10	95.00	2709
09/25/2021(week 38)	50.79	14.13	4.60	95.00	2842
10/30/2021(week 43)	53.54	14.03	4.70	95.00	3095
11/27/2021(week 47)	54.74	14.19	5.20	95.00	3095
12/31/2021(week 52)	56.79	13.72	10.30	95.00	3083

Notes: CDC has capped estimates of vaccination coverage at 95 percent.

Table A2: Impact of Tuskegee Study on vaccination rate (non-linear specification)

	Baseline results		County-month fixed effect		County × week linear trend	
	At least one dose	Fully vaccinated	At least one dose	Fully vaccinated	At least one dose	Fully vaccinated
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Black × distance × week	0.161*	0.167**	0.158*	0.153**	0.194***	0.192***
	(0.090)	(0.079)	(0.086)	(0.071)	(0.068)	(0.054)
Share of Black × week	-1.201***	-1.542***	-0.940**	-1.452***	-0.525	-0.750**
	(0.428)	(0.424)	(0.391)	(0.369)	(0.408)	(0.306)
Distance × week	-0.006	-0.004	-0.002	-0.002	0.011**	0.012**
	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Observations	134,259	143,666	133,971	143,533	133,971	143,533
R-squared	0.966	0.971	0.999	0.999	0.999	0.999
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes	No	No
Share of Hispanic × week	Yes	Yes	Yes	Yes	No	No
Share of White^2 × week	Yes	Yes	Yes	Yes	No	No
Share of Hispanic^2 × week	Yes	Yes	Yes	Yes	No	No
Share of HS grads and above × week	Yes	Yes	Yes	Yes	No	No
County unemployment rate × week	Yes	Yes	Yes	Yes	No	No
County-month FE	No	No	Yes	Yes	Yes	Yes
County × week linear trend	No	No	No	No	Yes	Yes
County^2 × week	No	No	No	No	Yes	Yes

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table A3: Impact of Tuskegee Study on vaccination rate (residential segregation)

	At least one dose			Fully vaccinated		
	Bottom 25% (1)	Top 25% (2)	Cross-Model Difference (p) (3)	Bottom 25% (4)	Top 25% (5)	Cross-Model Difference (p) (6)
Share of Black × distance × week	0.290**	0.257***	0.530	0.252**	0.246***	0.427
	(0.122)	(0.071)		(0.106)	(0.062)	
Share of Black × week	-0.241	-2.037***	0.013	-0.045	-2.210***	0.017
	(0.581)	(0.453)		(0.534)	(0.486)	
Distance × week	0.000	-0.004	0.647	0.006	-0.006	0.274
	(0.013)	(0.004)		(0.013)	(0.004)	
Observations	21,879	23,310		22,935	24,388	
R-squared	0.958	0.973		0.955	0.978	
County FE	Yes	Yes		Yes	Yes	
Week FE	Yes	Yes		Yes	Yes	
Share of White × week	Yes	Yes		Yes	Yes	
Share of Hispanic × week	Yes	Yes		Yes	Yes	
Share of HS grads and above × week	Yes	Yes		Yes	Yes	
County unemployment rate × week	Yes	Yes		Yes	Yes	

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. The segregation index between Black and White county residents with higher values indicates more residential segregation. Columns 1 and 4 show those using the sub-sample with counties on the bottom 25% of the distribution of residential segregation. Columns 2 and 5 show those using the sub-sample with counties in the top 25% of the distribution of residential segregation. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table A4: Robustness checks: constructing differential Black share

	Baseline (1)	1990 (2)	1995 (3)	2000 (4)	2005 (5)
Panel A: At least one dose					
Share of Black × distance × week	0.160*	0.136*	0.140*	0.144*	0.154*
	(0.087)	(0.074)	(0.076)	(0.079)	(0.081)
Share of Black × week	-1.200***	-0.563**	-0.656**	-0.802***	-0.979***
	(0.383)	(0.244)	(0.249)	(0.261)	(0.312)
Distance × week	-0.006	-0.002	-0.002	-0.003	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Observations	134,259	134,237	134,237	134,259	134,259
R-squared	0.966	0.966	0.966	0.966	0.966
Panel B: Fully vaccinated					
Share of Black × distance × week	0.164**	0.143**	0.147**	0.148**	0.160**
	(0.077)	(0.066)	(0.068)	(0.071)	(0.072)
Share of Black × week	-1.442***	-0.801**	-0.886***	-1.022***	-1.206***
	(0.398)	(0.318)	(0.321)	(0.322)	(0.353)
Distance × week	-0.005	-0.001	-0.002	-0.002	-0.004
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	143,666	143,644	143,644	143,666	143,666
R-squared	0.970	0.970	0.970	0.970	0.970
County FE	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes	Yes
Share of Hispanic × week	Yes	Yes	Yes	Yes	Yes
Share of HS grads and above × week	Yes	Yes	Yes	Yes	Yes
County unemployment rate × week	Yes	Yes	Yes	Yes	Yes

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table A5: Robustness checks: impact of migration

	Baseline		Excluding top 5 percentile		Excluding top 10 percentile	
	At least one dose	Fully vaccinated	At least one dose	Fully vaccinated	At least one dose	Fully vaccinated
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Black × distance × week	0.160*	0.164**	0.159*	0.165**	0.158*	0.163**
	(0.087)	(0.077)	(0.087)	(0.076)	(0.087)	(0.077)
Share of Black × week	-1.200***	-1.442***	-1.207***	-1.453***	-1.201***	-1.442***
	(0.383)	(0.398)	(0.387)	(0.401)	(0.389)	(0.404)
Distance × week	-0.006	-0.005	-0.006	-0.005	-0.006	-0.005
	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Observations	134,259	143,666	128,755	137,733	122,996	131,566
R-squared	0.966	0.970	0.967	0.971	0.967	0.971
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes	Yes	Yes
Share of Hispanic × week	Yes	Yes	Yes	Yes	Yes	Yes
Share of HS grads and above × week	Yes	Yes	Yes	Yes	Yes	Yes
County unemployment rate × week	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.

Table A6: Ruling out supply side confounding effects

	At least one dose			Fully vaccinated		
	(1)	(2)	(3)	(4)	(5)	(6)
Share of Black × distance × week	0.144*	0.139*	0.139*	0.148**	0.143**	0.144**
	(0.081)	(0.079)	(0.079)	(0.071)	(0.069)	(0.069)
Share of Black × week	-1.013***	-0.905***	-0.896***	-1.259***	-1.131***	-1.081***
	(0.324)	(0.323)	(0.333)	(0.329)	(0.331)	(0.345)
Distance × week	-0.001	-0.002	-0.002	-0.000	-0.001	-0.002
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
Observations	134,251	134,251	134,251	143,658	143,658	143,658
R-squared	0.967	0.968	0.968	0.971	0.972	0.972
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Share of White × week	Yes	Yes	Yes	Yes	Yes	Yes
Share of Hispanic × week	Yes	Yes	Yes	Yes	Yes	Yes
Share of HS grads and above × week	Yes	Yes	Yes	Yes	Yes	Yes
County unemployment rate × week	Yes	Yes	Yes	Yes	Yes	Yes
Pharmacies × week	Yes	Yes	Yes	Yes	Yes	Yes
Primary physicians × week	No	Yes	Yes	No	Yes	Yes
Nurses × week	No	No	Yes	No	No	Yes

Notes: The distance is scaled by dividing the raw distance by 100 miles for interpretation purposes. The number of pharmaceutical sites, primary physicians, and nurses are normalized by the county population. All regressions are weighted by total population age 18 and older at the county level and standard errors in parentheses are clustered at the state level. * $p < 0.01$; ** $p < 0.05$; *** $p < 0.01$.