Tuskegee and the Health of Black Men

Marcella Alsan and Marianne Wanamaker*

April 2016
PRELIMINARY. COMMENTS WELCOME.

Abstract

JEL Codes: I25, O15 For forty years, the Tuskegee Study of Untreated Syphilis in the Negro Male passively monitored hundreds of adult black males with syphilis despite the availability of effective treatment. The study's methods have become synonymous with exploitation and mistreatment by the medical community. We find that the historical disclosure of the study in 1972 is correlated with increases in medical mistrust and mortality and decreases in outpatient physician interactions for black men. Blacks possessing prior experience with the medical community, including veterans and women, appear to have been less affected by the disclosure. Our findings relate to a broader literature on how beliefs are formed and the importance of trust for economic exchanges involving asymmetric information.

*We are grateful to William Collins, Joe Ferrie, Nathan Nunn, John Parman, Achyuta Adhvaryu, Arun Chandrasekhar, Martha Bailey, Rebecca Diamond, Claudia Goldin, Melanie Morten, Mark Duggan, Mark Cullen, Melissa Dell, Nancy Qian, Ran Abramitzky, Pascaline Dupas, Rema Hanna, Grant Miller and participants at NBER DAE, University of Tennessee, Vanderbilt Health Policy, Carnegie Mellon Applied Microeconomics, University of Copenhagen Economics, University of Pennsylvania Health Policy, ASSA 2016 Berkeley Population Center, University of Chicago Harris and Stanford Health Policy for constructive comments. We thank the CDC for providing access and to the administrators at the Atlanta and Stanford Census Research Data Centers for their help in navigating the restricted data. We thank Michael Sinkinson, Martha Bailey, Andrew Goodman-Bacon and Walker Hanlon for sharing data and methods. Mario Javier Carrillo, Anlu Xing and Afia Khan provided excellent research assistance.

1 Introduction

> The Tuskegee study became a symbol of their mistreatment by the medical establishment, a metaphor for deceit, conspiracy, malpractice, and neglect, if not outright genocide.
>
> Corbie-Smith *et al* (1999)

The Tuskegee Study of Untreated Syphilis in the Negro Male (TSUS) is frequently cited as an important factor contributing to the alienation of black Americans from the healthcare system. For 40 years, between 1932 and 1972, the U.S. Public Health Service (PHS) followed hundreds of black men in Tuskegee Alabama, many of whom had syphilis, for the stated purpose of understanding the natural history of the disease. The men were denied highly effective treatment for their condition (most egregiously, penicillin, which became standard of care by the mid-1940s) and were actively discouraged from seeking medical advice from practitioners outside the study (**?**). Participants were subjected to blood draws, spinal taps, and, eventually, autopsies. Survivors later reported that study doctors diagnosed them with "bad blood" for which they believed they were being treated. Compensation for participation included hot meals, the guise of treatment, and burial payments.

News of the Tuskegee study became public in 1972 in an exposé by Jean Heller of the Associated Press, and detailed narratives of the deception and its relationship to the white medical establishment were widespread.[1] Anecdotally, TSUS is often invoked as a reason black Americans, particularly black men, mistrust the medical establishment, contributing to delays and avoidance in care seeking, wariness of public health campaigns, low participation in medical trials, and overall worse health outcomes.[2,3] The stubborn persistence of racial disparities and the lagging progress of black men's health, even conditional on socioeconomic status, adds to the anecdotal evidence.[4]

In this paper, we investigate whether the Tuskegee disclosure contributed to racial health disparities

[1] Contemporary coverage included *The New York Times* and *The Chicago Tribune*, as well as newspapers and magazines with a predominantly black readership.

[2] An October 2015 episode of ABC sitcom "Black-*ish*", a sitcom which centers on an African-American family, begins with a discussion of the Tuskegee experiment and the effect it has on the health-seeking behavior of the protagonist's aging father. The results of this paper also accord with recent coverage of a severe tuberculosis outbreak in Marion, Alabama reported in *The New York Times.* The article references the Tuskegee study as a cause for medical mistrust and a reason why tuberculosis was left unchecked (**?**).

[3] Fear and mistrust of the medical system has also been documented among Jewish survivors of the Holocaust and linked to the prominent role Nazi physicians played in directing the genocide (**?**).

[4] For a comprehensive review of racial inequalities in US medical care, see the **?** *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care.* Black men have the lowest life expectancy of any demographic group in the United States. Although gaps are closing, as of 2010 black men can expect to live 5 years less than their white male peers and 6.3 fewer years than black females (**?**, p. 8).

in the years following 1972. To do so, we rely on a variety of survey and administrative data, including measures of trust in doctors in the General Social Survey (GSS) (**?**), health seeking behavior reported in the National Health Interview Survey (NHIS) (**?**), and detailed annual mortality data available by race, age group, gender and cause from the Centers for Disease Control and Prevention (CDC) (**?**). We focus our attention on these measures for older (45-74) individuals since the mortality and health-seeking behavior of younger individuals is driven by acute conditions such as childbirth or trauma where the needs for care are urgent and the benefits immediate.[5] We posit that mistrust is more likely to dissuade medical care for the management of chronic conditions or for preventive care, and our preferred specifications for mortality are restricted to causes of death related to chronic disease.

We test the hypothesis that the Tuskegee study disclosure had a stronger impact on the behavior of individuals who more readily identified with the study's impoverished black male subjects. In particular, we hypothesize that black men would have been more affected than either black females or white males. Our approach is supported by both survey data specific to TSUS highlighting racial and gender differences in related mistrust (**???????**) as well as research on the neurological basis of the empathy response.[6] In addition to race and gender, we proxy for proximity to the study population using two different measures: distance from Tuskegee, Alabama and the prevalence of black migrant inflows from the state of Alabama.

The effects of the Tuskegee disclosure are interpreted within a multi-period model of health belief formation and healthcare utilization with heterogeneous agents. The agents vary along two dimensions: their medical experience as young adults and their proximity to Tuskegee. Proximity for these agents, which can be based in geography, culture, or both, is interpreted as a factor facilitating information transmission regarding the Tuskegee experiment but also affecting the weight individuals place on this information. Agents in this model move through life stages updating their beliefs about whether doctors can be trusted, incorporating both positive and negative signals via Bayes' Rule. The model predicts that subgroups who are exposed to routine medical care in young adulthood, such as women during pregnancy or military enlistees, will have a more muted behavioral response to the Tuskegee disclosure relative to individuals who lack such prior exposure, in particular non-veteran black men. The model further predicts that individuals in closer proximity to the study's subjects will exhibit a heightened response to disclosure.[7]

[5] Injuries and violence are the leading cause of death for US children and young adults between ages 1 and 44 (**?**).

[6] The physiological basis of these findings has recently been elucidated by neurocognitive research on how empathetic responses are modulated (**?**) and how such responses are stronger for a member of one's own group (**?**).

[7] The study distinguishes between the non-veteran and veteran population of men. The adjective veteran is used if the sample consists of veterans. Black veterans were treated differently from civilians and the men in the study. Perhaps the most egregious

These predictions are tested using a Difference-in-Difference-in-Difference (DDD) framework that compares healthcare utilization and outcomes across demographic groups (e.g. black men versus white men, or black men versus black women, the first difference) before and after 1972 (second difference) for those in varying proximity to the event (third difference). Whenever feasible, we condition on a rich set of control variables found to be correlated with health seeking behavior, including education (**?????**), income (**??**), marital status (**??**), and urbanization. The finest level of geography observable in the utilization regressions is a respondent's state, and our results contain state-year fixed effects to capture the state-specific implementation and diffusion of healthcare programs and policies as well as group-by-year effects for black men to capture any time-varying factors that affected all black men nationwide. For mortality outcomes, which are measured at the level of state economic areas (SEAs), we replace state-year fixed effects with SEA-year fixed effects and retain group-by-year fixed effects to again capture time-varying mortality changes specific to locations or demographic groups. Our estimates, then, represent the post-1972 difference in healthcare utilization or mortality for black men relative to black women or white men as a function of proximity to the study's subjects, net of pre-1972 differences by race, sex, and proximity.

We focus on the short- and medium-term impact of the disclosure on health behaviors and outcomes. Given the data limitations, persistent causal impacts are more difficult to identify. Yet, the effects we describe may shed light on factors influencing healthcare demand and deepen our understanding of how this watershed event affected the relationship between black Americans and the US healthcare system.[8]

We estimate that the Tuskegee disclosure reduced utilization of routine care (outpatient physician contacts) among older black males by 1.9 interactions per year per 1000 kilometers of proximity to Tuskegee when compared to their white peers and 1.4 interactions per year relative to black women. The majority of black men lived in close proximity to Tuskegee, and these estimates correspond to average treatment effects of XX and YY%. **(Exact numbers awaiting disclosure review.)** Similar results obtain for alternative measures of proximity. Our utilization results are driven by the behavior of low-income black men and black men with lower levels of education, a population in closer socioeconomic proximity to the victims of TSUS

example of this is that men in the Tuskegee study who were drafted into the military were not given penicillin, by the order of the US Public Health Service, despite the fact the Draft board was screening and treating all men at the time (**??**).

[8] James H. Jones reflected on what motivated him to write *Bad Blood: The Tuskegee Syphilis Experiment*: "First and foremost, I wanted to examine the role of race in medicine. Specifically, I sought to learn how racial attitudes affected the perception of disease that white physicians brought to their African American patients, and having done so, I wanted to learn how those attitudes altered the ways in which white physicians responded to disease in the black community. Scholars had taught us a great deal about race and politics, race and social structure and race and the economy. But we knew very little about the relationship between race and medicine. The Tuskegee Syphilis Study, I was convinced, was a critical case that could help to fill this lacuna (**?**, p.xii)."

and in line with our hypotheses regarding connectedness to the study population.[9] Our mortality results imply an annual increase in black male mortality of between 4% and 9% per thousand kilometers of proximity to Tuskegee. These estimates translate into a one year loss of lifespan for older black men.[10]

An alternative explanation for our findings is that access to healthcare was lower for blacks than for whites, particularly in the South, and this might have precluded black men from obtaining routine medical care regardless of their underlying demand. But this explanation is inconsistent with other economic research concluding that access to health services improved in the years following the Civil Rights Act for black Americans, resulting in an overall reduction in racial health disparities during this period. There were various forces working towards convergence at this time, including the implementation and expansion of Medicaid and Medicare (**?**), the desegregation of hospitals (**??**), and political enfranchisement, which led to an increase in the flow of public goods to black communities (**?**).[11] Our estimates imply that the Tuskegee revelation stalled the overall pattern of convergence for older black men, producing a drag on healthcare utilization growth for this demographic group, and contributing to an abrupt divergence in mortality relative to their white peers. Results comparing black men to black women directly address the concern that our strategy is picking up differential access or rollout of programs after 1972 for blacks living in the South. Our findings on health-seeking behavior and mortality are robust to estimation on a within-South sample and to various transformations of the distance variable. Furthermore, the addition of high frequency county-level data on transfer payments related to healthcare and disability (Medicare, Medicaid, Social Security) does not materially affect results and we find no evidence that expenditure levels for these social programs exhibited a temporal or geographic pattern mirroring our main empirical results.[12] Alternatively, there may have been other factors, such as high unemployment of unskilled labor or mass incarceration, correlated with proximity to Tuskegee that disproportionately affected the mortality and healthcare utilization of black men. Again, however, conditioning on a rich set of individual and location-specific socioeconomic characteristics, including unemployment, does not affect our estimates.[13,14]

[9] We also find a reduction in the probability that black men were admitted to the hospital, though conditional on admission, their length of stay was longer–consistent with the notion that black men had more advanced disease upon presentation to the hospital.

[10] See Appendix for details of the life table calculation.

[11] For an historical account of US black-white health disparities, see **?**.

[12] See Appendix Tables 5 and 8.

[13] These two robustness checks (in Appendix Table 4) cannot be performed in the baseline specification, which include location-year fixed effects. But when we revert to location and year fixed effects and then add these controls, we see little change in estimates.

[14] We thank Jacob Vigdor for this comment. State-by-year and SEA-by-year fixed effects control for any year-specific shock that affects all men or all blacks, but cannot account for shocks that are race- or gender-specific and correlated with geography. Unfortunately, the Current Population Survey (CPS) only has longitudinal data for 11 states and approximately 19 MSAs dating

We pursue a number of strategies to provide further evidence that the correlations we find are indeed causal. To allay concerns that our treatment is correlated with geographical features that affect health attitudes and behaviors more broadly (**??**), we demonstrate that our findings are absent in comparisons between white men and white women, and in comparisons between black women and white women. In addition, we show that the coefficient on a triple interaction with proximity to the SEA housing Tuskegee or to the state of Alabama is larger than 93% to 98% of placebo tests using proximity from all other SEA's or states in the dataset. Thus the effect we measure is specific to geographic proximity to Tuskegee and not to a post-1972 condition affecting the South overall.

To test whether the behavioral responses we observe are driven by medical mistrust, we utilize survey data from the 1998 wave of the GSS on whether individuals trust a doctor's judgment and whether they suspect that the medical establishment will deny them necessary treatment or services. We show that the level of mistrust reflected in responses to each of those questions is greater for black men than for black women, and for black men than white men. When we interact race and gender indicators with a measure of the distance of an individual from Macon County, Alabama *in 1972*, we find the same geographic gradient apparent in our baseline utilization and mortality results. This increased mistrust by geographic proximity to Tuskegee is present even when conditioning on the current state of residence (in 1998) and on the overall level of mistrust revealed in the respondent's answers to other survey questions.[15]

This paper builds on and contributes to several literatures in economics. First, our study is motivated by the theoretical contributions of **?** regarding the transmission and updating of beliefs, and empirical work by **?** regarding the role of history (namely, the export slave trade) in shaping interpersonal mistrust in Africa. Our findings also connect to rich empirical evidence on the importance of trust for economic development (**??**) and firm management (**?**). Finally, the research sheds light on questions in development economics regarding low demand or uptake for products that have been shown to improve health (see review by **?** and Chapter 3 of **?**). The findings presented herein suggest that historical exploitation and its enduring impact on beliefs may explain some of the uptake paradox.[16]

back to the early 1970s. We therefore use race- and gender-specific unemployment from the 1970 census by county interacted with a time-trend to control for macroeconomic conditions for each group. See Appendix Table 4.

[15] Although attrition is an issue in such a sample, given the findings on mortality described above, these estimates likely represent a lower bound.

[16] This is also provided as a potential explanation for aforementioned low demand in **?**, p.58: "Faith or its secular equivalents, beliefs and theories, is clearly a very important part of how we all navigate the health system."

2 Literature Review and Historical Context

The Tuskegee Study was designed to trace the course of untreated syphilis. The organism that causes the disease is closely related to that causing Lyme disease, and both bacterial diseases manifest themselves in stages. The first stage of sexually acquired syphilis is often an ulcer, followed by a full body rash that includes the palms and soles. However, it is the tertiary (or late-stage) syphilis that inflicts the most damage. The third stage is characterized by gummas (syphilitic tumors teaming with the bacteria) which coalesce and eat away at bone (frequently the nasal bridge) as well as other organs, and show a predilection for the arch of the aorta. Neurosyphilis (an attack on the nervous systems) presents in late-stage syphilis with paresis, gait disturbance, blindness and dementia (**?**).[17] According to **?**, much of the natural history of syphilis outlined above was known at the time the study commenced : "The germ that causes syphilis...and the complications that can result from untreated syphilis, were all known to medical science in 1932–the year the Tuskegee Study began. Since the effects of the disease are so serious, reporters in 1972 wondered why the men agreed to cooperate. The press quickly established that the subjects were mostly poor and illiterate, and the PHS had offered them incentives to participate."[18] These incentives included physical exams, hot meals, and burial stipends that would be paid to their survivors. Most of the men also believed they were receiving some form of treatment. Approximately 600 black men (399 with syphilis) were recruited to the study using these techniques and followed passively for forty years while the disease took its toll.

In 1972 news of the Tuskegee study was leaked to the press and quickly spread throughout the black community. By the end of the study that year, only 74 of the test subjects were alive. Of the original 399 men with syphilis, 28 had died of syphilis, 100 were dead of related complications, 40 of their wives had been infected, and 19 of their children were born with congenital syphilis (**?**). Medical historian Allan **?** summarizes the study as follows: "In retrospect the Tuskegee Study revealed more about the pathology of racism than the pathology of syphilis; more about the nature of scientific inquiry than the nature of the disease process....The degree of deception and the damages have been severely underestimated."

Those damages may include a legacy of medical mistrust among black Americans. The Final Report of the **?** noted that, "The Study continues to cast a long shadow over the relationship between African Americans and the biomedical professions," and was "a significant factor in the low participation of African

[17] Syphilis can also be transmitted mother to child and cause severe congenital problems including stillbirth.

[18] According to **?**, p. 1500, "the fact that Whites ruled Blacks in Macon County, coupled with the Black men's extreme poverty and almost total lack of access to health care, made the men willing subjects."

Americans in clinical trials, organ donation efforts, and routine preventive care." Further evidence of the role of mistrust in the health of black Americans comes from clinical psychology. **?** recruited 600 African-American men from around the country from local barber shops. Their purpose was to assess predictors of self-reported delays in medical care. The most important factor was medical mistrust–which doubled or tripled the odds ratio of delays in blood pressure and cholesterol screening. The authors were motivated to study medical mistrust as a potential predictor of preventive health seeking behavior because it "is higher among African-Americans" and "is linked to visible incidents of race-based malice towards this group (e.g. The Tuskegee Study of Untreated Syphilis in the Negro Male)."

Despite this, the empirical evidence linking Tuskegee to medical mistrust and to health behaviors and outcomes is much more limited and does not attempt to establish causality.[19] **?**, p. 541 used a discussion group setting to understand the reluctance of black men to participate in medical research. In the group, "participants made [unprompted] reference to the Tuskegee Syphilis Study." Similarly, participants in the same study used the Tuskegee event to justify "their belief that 'doctors value your life less than their own'" and that "African Americans still need to be suspicious when dealing with the medical establishment." To our knowledge, we are the first to relate racial disparities in healthcare utilization and health outcomes to this historical event in a quasi-experimental framework.

3 Data on Health Seeking Behavior and Health Outcomes

We measure health-seeking behavior using individual-level data on health practices and conditions from the NHIS. The survey began in 1963 with a relatively small set of questions related to health and health-seeking behavior.[20] The set of questions expanded over time, and the wording of questions and responses also changed. We rely on the harmonization provided by the Integrated Health Interview Series (IHIS), which is based on the NHIS public-use data, and note below where this harmonization matters for interpreting results. The public-use samples are stripped of geographic identifiers necessitating the use of restricted access for these data from the National Center on Health Statistics (NCHS) in order to incorporate location-based

[19]The most recent (and perhaps only) attempt to calculate the observational correlation between Tuskegee knowledge and medical mistrust (**?**) has been roundly criticized for methodological failures (**?**).

[20]The US Public Health Service of the United States — the same entity that carried out the Tuskegee study — conducted the survey. Thus it is reasonable to suppose we are estimating a lower bound since mistrustful individuals might have refused to answer the survey.

proximity measures described above and in Section 6 below.[21]

Each individual in the survey self-reported the interval since their last doctor's visit or physical, the number of physician visits and other physician interactions (including phone calls) in the last 12 months, as well as the frequency and duration of hospital admission.[22] The data also contain detailed demographic information, including family structure, income, education, and veteran status. From the surveys we have more than 200,000 respondents between 1969 and 1977.[23] The NHIS data lack morbidity measures, and we rely on separate mortality data to assess whether the changes in beliefs and behaviors which we document below translate into an effect on longevity.[24]

In doing so, we follow a large literature that uses mortality to assess the efficacy of health interventions.[25] County-level mortality statistics by age, race and gender are available from the U.S. Department of Health, Education, and Welfare for each year between 1968 and 1988. We merge population data for the same period from the CDC to calculate mortality rates by age group, race and gender (**?**). In these data, there are counties in the United States with few blacks so the number of deaths in a particular year is very low (and sometimes zero). Mortality rates based on a sparse number of events exhibit a large amount of random variation (**?**, p.3), and the NCHS suggests aggregating over space or time if small numbers are encountered. For this reason, we aggregate our annual county-level values to the state enumeration area (SEA) level and measure biennial mortality rates.[26]

Our baseline results are for chronic mortality, a measure that encompasses deaths from cardiovascular disease, cancer (all forms) smoking-related respiratory disease, gastrointestinal disorders (including ulcers and cirrhosis) and diabetes. These categories reflect the leading causes of death for older Americans during the time period, and represent diagnoses for which therapies and interventions by health professionals could

[21] We accessed the IHIS harmonized data with location-based relatedness proxies attached, inside Census Research Data Centers (RDCs) at Atlanta and Stanford. In later years, the NHIS data have been linked to mortality files, but these linked data are not available for our study period.

[22] Proxy respondents were only used for children and adults who were not at home at the time of the interview. Because phone calls to medical providers were counted as physician interactions, all NHIS regressions control for whether the household has a telephone.

[23] In 1978 the race categorization changed, with mutiple categories included. We have extended our analysis to 1981 (the year the question on doctor visits changes) and obtain similar (though not disclosed) results.

[24] Self-reported health status and other plausible metrics of morbidity in the NHIS are unavailable prior to 1973.

[25] Recent examples include **????** and **?**.

[26] County-level and annual mortality rates are in Appendix Table **6**. Using data from the mortality NCHS' Multiple Cause of Death file for ealier years is complicated as noted on the NBER website: "The chief of the NCHS mortality branch has said that while the 1959-1967 files are generally ok, they have not been rigorously verified. 'Counts by selected causes and demographic groups seem to match up with VSUS, but because in some cases these files had to be reconstructed and pieced together from different sources-some were damaged or lost-we cannot at this time be certain as to their accuracy.'"

plausibly prevent or reduce mortality.[27] We also utilize age-adjusted mortality rates that are not restricted to chronic diseases.[28] Since the impact of the Tuskegee revelation might affect mortality with a lag and last longer than the impact of utilization, we include mortality rates from 1968 to 1987 in our main analysis. Note that we do not have place of birth in the publicly available compressed files. But because we are interested in exposure as a function of 1972 residence for a group of older men who were geographically mobile earlier in the century, we propose that place of death is a better proxy for where they lived at the time of the disclosure than the place of birth. To the extent our current proximity measures capture information spread via social networks, then having data on *individual* migration histories, including time spent in an area before moving or detailed network data would be useful. Unfortunately, these data are not available for our time period of interest.

4 1960s and 70s Patterns of Healthcare Utilization and Mortality

The 1960's marked an era of rapid convergence in healthcare access and utilization for black Americans relative to their white peers. Hospitals, initially racially separated or segregated according to designated white and black beds, gradually integrated over the 1950s and early 1960s, a process that culminated not with the Civil Rights Act of 1964, but with Lyndon Johnson's insistence that any hospital receiving Medicare funding fully desegregate by July 1966. The process was quicker in the North than in the South; separate hospital wings were present in 75% of hospitals in the South as late as April of 1966, and Mississippi was a notable laggard. Still, full compliance appears to have been achieved by July of the same year (**??**). These decades also witnessed expansions on public health insurance coverage and services, some of which favored black Americans. The list of public programs includes the Kerr-Mills Act, expansions in Medicare and Medicaid, local health center construction, and the expansion and funding of public hospitals.

By 1970, black Americans had gained substantial ground in terms of healthcare utilization, even if the quality of the services they acquired lagged their white peers.[29] Using data from The National Health

[27] For instance, blood pressure and glucose control for the management and prevention of cardiovascular and diabetic complications such as heart attack, stroke and kidney failure as well as counseling on prevention, early detection and treatment of cancers. See Appendix Section 5 for a discussion of medical and public health innovations over this time period.

[28] We follow the demographic literature and use the standard 1940 population for reference. See the Data Appendix for further details.

[29] Blacks were more likely than whites to utilize public and community-based healthcare resources, including charity clinics, public health departments, community health centers, and city and county hospitals. Whites were more likely to make use of private and non-profit hospitals. Some of this difference was spatial, the product of rapid suburbanization of the white population (**?**, p. 391).

Interview Survey (NHIS), among individuals ages 45 to 74, the percentage of black women who had contact with a medical professional within the past year rose between 1963-64 and 1967-68 from 64% to 73%. The gains for black men were similarly impressive; the same metric rose from 57% to 65%. These increases of around 9 percentage points were over twice as large the increases experienced by white males and females over the same time period.[30]

The convergence of black and white healthcare utilization rates in the years prior to Tuskegee are closely mirrored by convergence in mortality rates. Using CDC-provided annual county-level mortality statistics by age, race and gender, as described above, we plot the racial difference in age-specific mortality rates (ASMR) (black minus white) for both men (solid line) and women (dashed line) in Figure 1. Panels (A) and (B) demonstrate that, for infants and children 1 to 4 years of age, there was a marked reduction in racial health disparities that continued uninterrupted after the 1972 disclosure. But the pattern for adult mortality is starkly different. Panels C and D plot the same data for adults aged 55-64 and 65-74. At these ages, we observe a striking divergence in the mortality rates for black men relative to white men beginning in the early 1970s, a pattern which is not reflected in the difference between black and white women. The mortality pattern for young adults (contained in Appendix Figure 2) mimics that for infants and children.[31] Unlike younger ages, where parental behavior (for early childhood) and unintentional injuries (for young adults) play key roles in determining mortality, premature and preventable mortality in older adults is more frequently driven by health behaviors, including timely diagnosis and management of chronic illness.[32] To facilitate comparison across age groups, the mortality rates represented by Figure 1 are those from all deaths.

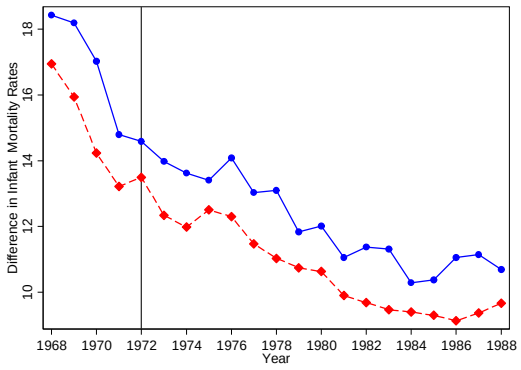
[30]The level of utilization for black (and white) men is lower than for women in every period; the gap is one-third larger for those who have never served in the military. Note that we do not have a consistent, continuous series of utilization data in the pre-1969 period.

[31]The appendix also contains figures for Southerners alone (Appendix Figure 1) and all other age groups (Appendix Figure 2). These latter figures indicate that the 1960's convergence was particular to older black men.

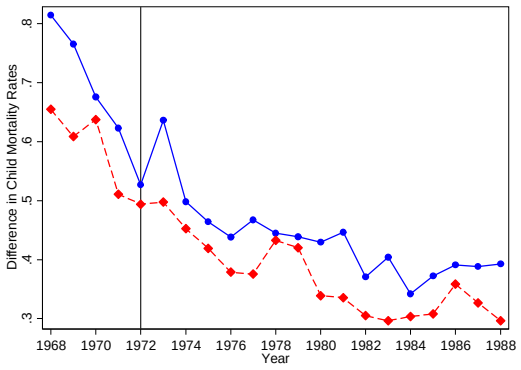
[32]An alternative way to have plotted these data would be to plot the racial differences in ASMR against age at the time of the Tuskegee disclosure (rather than year). From that perspective, in Panel D, year 1972 corresponds to those roughly 60 years old at the time of the disclosure and moving out in time captures the impact on younger cohorts. Since care of chronic diseases becomes particularly important in middle age, one way to interpret these graphs is that the health impact was larger for individuals exposed at a time when their actions could yet play an important role in determining their longevity. Results for chronic mortality in older adults are not remarkably different from the results shown in Panels C and D.

FIGURE 1: BLACK-WHITE MORTALITY DIFFERENCES BY AGE AND SEX

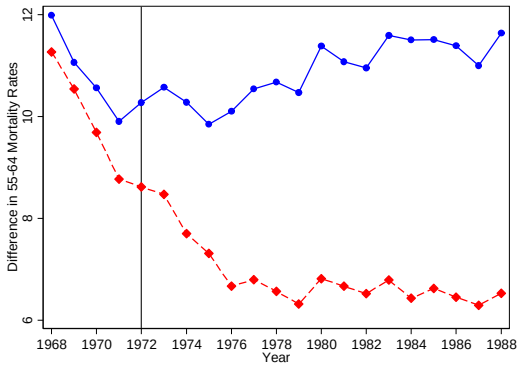
Panel A. Infant Mortality Rate



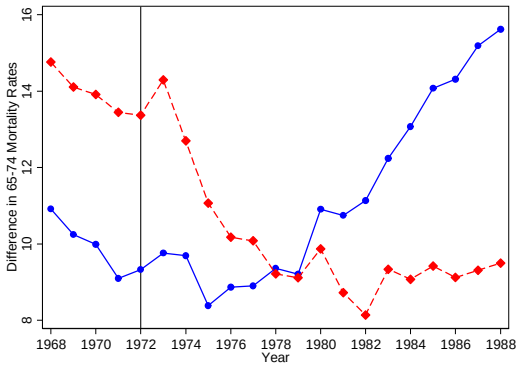
Panel B. Child Mortality Rate



Panel C. 55-64 Mortality Rate



Panel D. 65-74 Mortality Rate



Notes: The data are from the CDC compressed mortality files and represent the black-white difference in age-specific mortality rates. Each mortality rate is calculated by dividing the number of deaths in the relevant population by the at-risk population. The solid (blue) line represents the difference for males, and the dotted (red) line represents the difference for females. The vertical line represents the year "The Tuskegee Study of Untreated Syphilis in the Negro Male" was disclosed. For additional figures, including plots of all other age-specific mortality rates and South only, see the Appendix.

The significance of these trends for older black men is apparent in a simple double-difference (DD) regression comparing the post-1972 outcomes of these men (ages 45-74) relative to their white or female peers. For this exercise, we use chronic mortality data and estimate the post-1972 divergence of black men as the coefficient θ_1 in the following equation:

$$\ln(Y_{gat}) = \alpha + \theta_1(post_t * blackmale_g) + \theta_2(blackmale_g) + \pi_t + \varepsilon_{gat}.$$

Mortality rates for older adults (ages 45-74) are measured at the level of each SEA, denoted by a, for each of four demographic groups (denoted g): black males, black females, white males, white females. The

variable $blackmale_g$ is equal to one for each SEA-year observation for black males, and zero otherwise.

We begin by estimating the above equation on an all-male sample, comparing the experiences of black men to those of white men before and after 1972. As expected given the results in Figure 1, θ_1 is positive and significant, indicating post-1972 divergence for older black men relative to their white peers of approximately 6 log points. The estimated value of θ_1 and the associated 95% confidence interval is given in the heading for Figure 2, panel A.

Importantly, the estimated value of θ_1, essentially an average "treatment" effect of being a black male after 1972, masks significant heterogeneity. We find a strong geographic gradient along a measure of distance from Macon County, Alabama, the county that houses Tuskegee. When we estimate a θ_1 coefficient for every SEA in the sample, utilizing multiple pre- and post- observations within each SEA for identification, the resulting data indicate a strong negative correlation between distance from Macon County and the value of these coefficients. In Figure 2, Panel A, we plot the value of these coefficients averaged across SEAs in distance groups of approximately 180km, as well as a linear line of best fit for the underlying disaggregated data (in blue).[33] The size of the data bubbles are indicative of the number of older black men (aged 45-74) living in each group of SEAs. The difference in black male chronic mortality relative to white male mortality increases by 6.1 log points following 1972, but the geographic gradient of this result is stark. Black men in close proximity to Tuskegee saw an increase in relative chronic mortality in excess of 20 log points, and the effect declines linearly with distance.

As further evidence of the post-1972 divergence of black male mortality, we calculate θ_1 for an empirical specification restricted to black male and black female mortality observations. In this case, θ_1 represents the post-1972 mortality change for black men relative to their female peers. As indicated in the figure header for Panel B below, black male chronic mortality increased relative to black female mortality by a statistically significant 5.6 log points after 1972, but again a geographic gradient is apparent. The value is upwards of 10 log points in locations closest to Tuskegee and declines linearly in distance as we move away from Tuskegee.

These temporal and geographic mortality patterns for black men are not mirrored in other sub-groups. In Panel C, we estimate a similar regression, replacing $blackmale_g$ with $blackfemale_g$ and restricting the sample to be all female. In this case, θ_1 represents the change in chronic mortality of black women relative

[33] For graphical exposition, we collapse SEAs in the data into 20 buckets representing 20 equal distance intervals. Because fewer SEAs in the North and West housed black residents, buckets further than 1500 kilometers from Tuskegee (10 buckets) were collapsed again into 5 buckets representing equal distance intervals.

to white women in the post-1972 era. Consistent with our Figure 1 results, we find that black women continue to converge on their white peers after 1972, and the estimated value of θ_1 is -7.3 log points. But the geographic gradient in these results is slight and upward-sloping, indicating that black women in closer proximity to Macon County were, if anything, converging more quickly than their peers further afield. (Note the change in scale relative to panels A and B.) A similar exercise in Panel D examines the experience of white men with respect to their female peers. In this case, we replace $blackmale_g$ with $whitemale_g$ in the estimating equation and restrict the sample to all white chronic mortality observations. Overall, white men closed the gap in chronic mortality relative to white women by 7.6 log points, perhaps as they became beneficiaries of medical technologies and advice that spread over this time period. But, again, we see little evidence that proximity to Macon County was detrimental for the health of white men in the post period. Although the blue line of best fit is downward sloping, this result is driven by the negative coefficients from distances beyond 2000 kilometers from Alabama. The estimated values of θ_1 within the South show no geographic gradient.

As a placebo test of the conclusion that a geographic gradient in health outcomes for black men emerged in the post-1972 period, we look for the same patterns in the years prior to the Tuskegee disclosure by instituting a false definition of $post$ within the set of mortality observations observed prior to 1972. We plot estimates of θ_1 for the same set of distance buckets in panels E and F of Figure 2, defining $post$ to be one for all observations after the mid-point of the pre period. These falsification tests indicate no geographic gradient in the estimates of θ_1, and the slope of the linear line of best fit (in blue) is flat for both the all-male and all-black specifications. This is in contrast to the negative and highly statistically significant slope coefficient in panels A and B.[34]

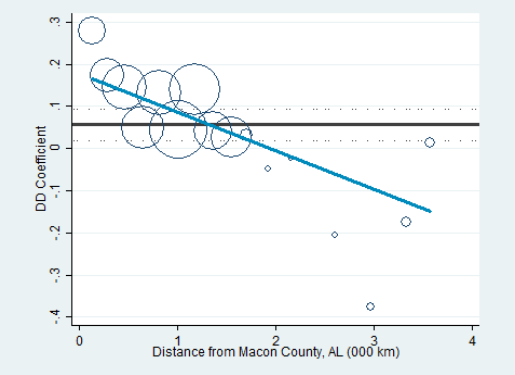
[34]For confidentiality reasons, we cannot generate the same figures above for the NHIS healthcare utilization data.

FIGURE 2: DIFFERENCE-IN-DIFFERENCE COEFFICIENTS AND GEOGRAPHIC GRADIENTS

Panel A. θ_1 - All Male Sample

DD Coefficient: 0.061

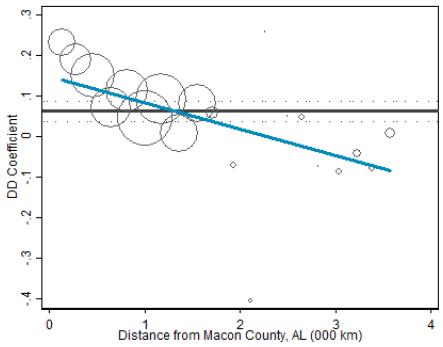
95% CI: [0.035, 0.086]



Panel B. θ_1 - All Black Sample

DD Coefficient: 0.056

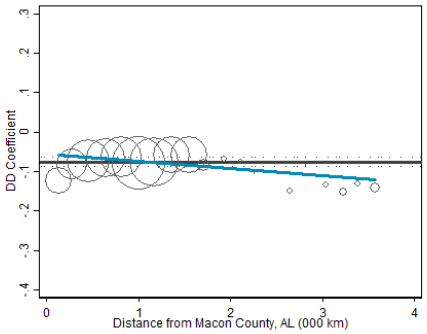
95% CI: [0.019, 0.092]



Panel C. θ_1 - All Female Sample

DD Coefficient: -0.073

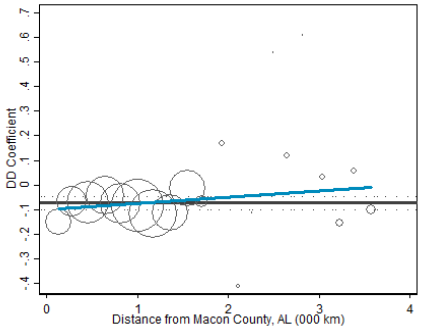
95% CI: [-0.101, -0.045]



Panel D. θ_1 - All White Sample

DD Coefficient: -0.076

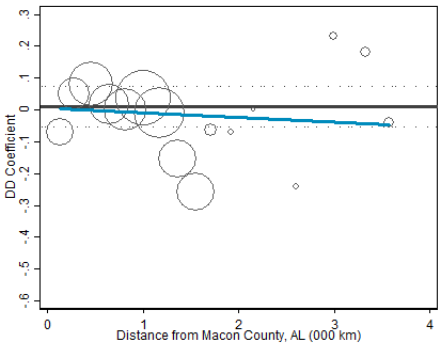
95% CI: [-0.088, -0.064]



Panel E. θ_1 - All Male Sample with False Post

DD Coefficient: -0.014

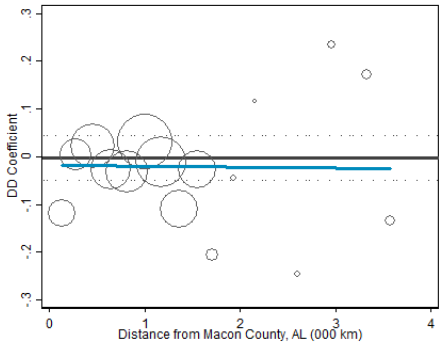
95% CI: [-0.067, 0.039]



Panel F. θ_1 - All Black Sample with False Post

DD Coefficient: 0.010

95% CI: [-0.054, 0.074]



Notes: These figures represent the coefficient in each SEA-specific regression for log chronic mortality averaged across all SEAs in intervals of approximately 180 kilometers of distance from Macon County, Alabama. (The maximum observed distance from Tuskegee is divided into 20 intervals of equal distance width.) Thus, the point estimate at interval 1 in Panel A is the averaged coefficient on blackmale*post for all SEAs between 0 and approximately 180 kilometers from Macon County. Fitted solid lines (blue) represent the underlying linear relationship between distance and the coefficient on blackmale*post across all SEAs. The figure also includes the estimated coefficient on blackmale*post in a regression specification that pools all SEAs (black line) and the 95% confidence intervals for that estimate (dashed lines). The size of the circle is a population weight: the number of older adult (45-74) black men in the SEAs in that distance interval. Panels E and F represent the double difference plots using data from the period prior to the Tuskegee disclosure and using a placebo disclosure date of 1970. Panel A and B estimate the same specification and utilize the actual timing the Tuskegee disclosure. Panel C represents the coefficients on blackfemale*post and Panel D represents coefficients on whitemale*post for an all-female and all-white specification,

The geographic gradients documented in Figure 2, both across all regions and within the South, serve to inform both our theoretical framework and the empirical specifications to follow.

5 Theoretical Framework

As a framework for our empirical analysis, we provide a model of how prior beliefs, experience with the medical profession and proximity to the Tuskegee study subjects interact to determine health-seeking behavior. In this regard, our model relates both to work by **?** and **?** on how prior experience can influence beliefs and behavior and to belief formation models by **?** and **?**.[35] The model's equilibrium provides testable predictions which will inform our empirical analysis.

We consider a multi-period model of belief formation where individuals engage in Bayesian updating regarding whether doctors can be trusted to treat them appropriately.[36] Agents live for three periods (childhood, young adulthood, and old adulthood). They update their beliefs over two possible states of the world: the good state $(\theta = 1)$ where fraction $q > \frac{1}{2}$ of the doctors are trustworthy, and the bad state $(\theta = 0)$ where the fraction is only $1 - q'$, $q' > \frac{1}{2}$. To simplify the exposition, we set $q = q'$. At $t = 0$, agents are born with a prior belief over the probability that the state is good: π_i^{prior}.

At $t = 1$, some young adult agents are induced to visit the doctor. This inducement might come either from gynecological or obstetric care associated with being a young woman in her reproductive prime or

[35] Our study also relates to theoretical work on identity (**?**); however, we chose not to assume individuals have different preferences based on their identity, but rather that individuals' beliefs are affected to a greater extent by news that concerns the group with which they most nearly identify.

[36] **?**, p. 1183 defines trust as "an optimistic expectation or belief regarding other agents' behavior". Fafchamps further discusses the origins of trust and uses this to distinguish two types: personalized trust, which forms from "repeated interpersonal interactions", and generalized trust, which has its origins in "general knowledge about the population of agents." It is the latter we seek to model. Although many different updating processes could explain the empirical results, we prefer to use a rational model based on Bayesian updating. We thank Marcel Fafchamps for the comment.

from military enlistment for a young man. We assume such individuals, via this experience, receive one of two signals based on their experience: $S = 1$ indicating the state of the world is good or $S = 0$ indicating the state of the world is bad. In accordance with the actual underlying trustworthiness of medical providers, when the true state of the world is good (bad), agents will receive the signal $S = 1$ $(S = 0)$ with probability q $(1 - q)$ and signal $S = 0$ with probability $1 - q$ (q). These experienced agents then rationally update their beliefs regarding the state of the world using Bayes' rule.[37] If agents share a common prior, π^{prior}, the resulting beliefs at the end of young adulthood ($t = 1$) can be ordered as follows:

$$\Pr(\theta = 1|S = 1, \pi^{prior}) > \pi^{prior} > \Pr(\theta = 1|S = 0, \pi^{prior}).^{38}$$

In other words, individuals who encounter a trustworthy provider hold a higher posterior than inexperienced agents who, in turn, hold a higher posterior than experienced agents who encounter an untrustworthy provider.

Our focus is on the behavior of older men, and we continue by deriving predictions over the probability of seeking medical care in the last period, when the details of the Tuskegee experiment are revealed. In addition to heterogeneity regarding prior experience with the medical system, agents differ regarding their proximity to the victims of the Tuskegee experiment. This proximity reflects both connectedness that facilitates the transfer of detailed information regarding the experiment, as well as cultural connectedness affecting the degree to which individuals update their prior about how doctors will treat them based on the same. At $t = 2$, both experienced and inexperienced agents enter middle age, at which time Tuskegee is disclosed via a bad signal, $S = 0$, indicating the state is $\theta = 0$. Those in close proximity to Tuskegee receive the signal and update accordingly, accepting that the experience of the Tuskegee victims is relevant for their own experience. Agents more distant do not update in accordance with the signal, either because they do not receive it or because they receive it but do not believe it has relevance for their experience. Therefore, we can depict the posterior information structure as a 2x2 matrix based on proximity to Tuskegee and prior experience with the medical system:

[37] Such individuals cannot convincingly share their information with inexperienced agents. For the experienced group, their doctors visits are induced and therefore the fact that they visit is uninformative to inexperienced agents. We can relax this assumption and derive the same predictions, provided individuals weigh the signal from others' experience less than the signal from their own experience.

	"Near" to Event	"Far" from Event
Experienced	$\pi_i^{post} = \Pr(\theta = 1 \mid S = 1 \text{or } S = 0, S = 0, \pi_i^{prior})$	$\pi_i^{post} = \Pr(\theta = 1 \mid S = 1 \text{ or } S = 0, \pi_i^{prior})$
Unexperienced	$\pi_i^{post} = \Pr(\theta = 1 \mid S = 0, \pi_i^{prior})$	$\pi_i^{post} = \pi_i^{prior}$

By a similar logic as above, and again assuming a common prior, the ordering of posterior beliefs in the matrix above is as follows:

$$\begin{aligned} \Pr(\theta &= 1|S = 1, \pi^{prior}) > \Pr(\theta = 1|S = 1, S = 0, \pi^{prior}) > \Pr(\theta = 1|S = 0, \pi^{prior}) \\ &> \Pr(\theta = 1|S = 0, S = 0, \pi^{prior}). \end{aligned}$$

π^{prior} cannot be ordered without further assumptions on θ.[39]

Older agents must now decide whether to seek out medical care. The gross utility from medical care, u_i, is a function of their choice of action, a_i, and the state of the world, while the cost, $c(a_i)$, is dependent only on their actions:

$$u_i = u(a_i, \theta) = \begin{array}{c} 1 \text{ if } a_i\text{=1 \& } \theta = 1 \\ 0 \text{ if } a_i\text{=1 \& } \theta = 0 \end{array} \text{ and 0 if } a_i = 0.$$

$$c = c(a_i) = \begin{array}{c} c \text{ if } a_i\text{=1 where } c < q \\ 0 \text{ if } a_i\text{=0} \end{array}.$$

where $a_i = 1$ ($a_i = 0$) denotes visiting (not visiting) a doctor. A rational agent seeks out medical care in old age if and only if $EU(a_i = 1) > EU(a_i = 0) \equiv 0$ where EU denotes expected utility. This inequality hinges on a threshold rule involving the posterior belief, and the agent seeks care if and only if:

$$\pi_i^{post} > \Omega.$$

where $\Omega \equiv \frac{q+c-1}{(2q-1)}$.

Because these posterior beliefs map directly into actions, we can use the ordering of posterior beliefs to make empirical predictions about the behavior of individuals in older age.[40] *In the aftermath of the Tuskegee disclosure*, older adults will exhibit health-seeking behavior such that:

[39] See Theorem 2 in the Appendix.

[40] One can also show that the threshold posterior above which agents visit the doctor, π^*, is declining in the benefits doctors provide and increasing in the cost of the visit. Generalizing the utility of visiting a good doctor from $u_i = 1$ to $u_i = B$: $\frac{\partial \pi^*}{\partial B} = \frac{-c}{2q-1}\frac{1}{B^2} < 0$ and $\frac{\partial \pi^*}{\partial c} = \frac{1}{B(2q-1)} > 0$ for $q > \frac{1}{2}$. Also see the Appendix for details.

1. Conditional on experience in young adulthood, agents closer in proximity to the Tuskegee event will exhibit lower health-seeking behavior rates on average than agents farther away.[41]

2. If the true state of the world is $\theta = 1$, inexperienced agents closer in proximity to the Tuskegee event will exhibit lower health-seeking behavior rates than experienced agents in the same locations.[42]

If we assume black agents are in closer proximity to Tuskegee than their white peers and that black men are in closer proximity than black women, the first implication above implies a testable prediction that black men should have been more affected by the disclosure than both white men and black women.[43] Additional measures of proximity and their testable implications are proposed below.

6 Proximity Measures

As described in the model above, as well as in other models of trust formation (**??**) and in a vast psychological literature, individuals tend to be more affected by news if they can identify with the subject (**??**).[44] Tabellini models agents around a circle, then randomly couples them to play the prisoner's dilemma. Agents are modeled as enjoying noneconomic benefits of cooperation which decline in distance. Distance, as described by Tabellini, "could refer to geography, but also to social or economic dimensions such as religion, ethnicity and class." We build on Tabellini's description of distance–combining geography with age, race and gender to identify the effects of the disclosure. We begin with the assumption that race and gender will determine proximity such that black men are more affected by the news of Tuskegee than either black women or white men. We then add one of three additional geography-based measures of proximity to complete the DDD framework, which also relies on a pre- and post-1972 difference. In Section 11, we discuss heterogeneous effects by education and income measures, additional measures of proximity to the study's subjects.

[41] See Theorem 2 in the Appendix.

[42] See Theorem 3 in the Appendix.

[43] Data from the Survey of Black Americans supports the notion that black men from the South identified more with the men from the study. When asked, "How close does the respondent feel to black people who are poor" 78% of black men born in the South answered they felt "very close" compared to 60% of men born elsewhere. Similar numbers for black females are 74 and 63%. We thank Trevon Logan for pointing out this source.

[44] As evidence of this, although the Rodney King beating, which took place in March 1991, was widely publicized on many media markets, opinions on the police force shifted most markedly for black men. Though we lack geographic identifiers in the survey data, comparing 1989 and 1992 polling data, the percentage of people who disagree with the statement: "These days police in most cities treat Blacks as fairly as they treat Whites" jumped from 63.9 to 81.5% among black men, 77.5 to 85.4% among black women, 67.3 to 44.9 among white men and 67.4 to 34.5 among white women (**??**).

Our empirical results rely on two different proxies for the proximity of individuals to the Tuskegee study's subjects and, thus, their subsequent belief updating process following the study's disclosure. Our primary proxy is a linear measure of geographic proximity, inspired by the linear relationship between post-1972 mortality and distance from Macon County, Alabama, as displayed in Figure 2. This measure of geographic proximity is mapped in Panel A of Figure 3. Linear distance creates concentric circles of equal proximity around the impact point of Tuskegee, Alabama. We wish to remain agnostic regarding the precise mechanism through which geographic proximity affects behavior, but note that geography may capture information spread through formal and informal networks or cultural similarity as discussed above.[45] This proximity may also reflect the information content of the signal — if, for example, individuals at greater geographic distances do not believe the details of the Tuskegee experiment provide relevant information regarding the trustworthiness of their local medical professionals.

As an alternative measure that captures both informal information networks as well as cultural similarity, we use the fraction of black migrants to a particular state or SEA from the state of Alabama (Panel B). (We drop Alabama observations from the sample in this analysis.)[46] The Tuskegee disclosure followed several decades of large-scale migration, particularly among African-American males, and we hypothesize that locations with a large number of incoming migrants from Alabama will also hold a high cultural proximity to the event and a more direct path over which information about the event may have traveled. We take advantage of the complete count version of the 1940 U.S. Census, which contains a question about 5-year migration patterns with detailed geography measures, to calculate this statistic.[47] Migration rates are highest in young adulthood; thus, the migration patterns from 1935-1940 will be reflective of cultural proximity in individuals aged 45-74 in 1972, provided individuals remain connected to their locations of origin 30 years hence. Further, the 1940 Alabama migration rates are also relevant as migration patterns are roughly stable over time, and new migrants contribute to the cultural proximity and information flows of the broader

[45] We exclude the SEAs containing Los Angeles and San Francisco from the analytic sample when we use distance as a proxy for treatment intensity. These two cities hold most of the black population in the West region, and our distance-based proxy for treatment intensity does not appear to work well for them. We do, however, include these two cities when using the migration instrument.

[46] We normalize by black migrants for two reasons. First, because it is informative regarding what percentage of new information is coming from Alabama. (If the denominator was the black population, for example, the new information measure would be diluted.) Second, and more importantly, we only observe the migration variables in the five year period of 1935 to 1940. But the migration of this generation of men extended from 1918 up through 1960. The patterns of migration were persistent over time, so even though the 1935-1940 measure is not an accurate measure of the absolute number of migrants, the relative measure (those from Alabama divided by all migrants) is a good proxy for how connected the stock of black residents in 1972 would have been to Alabama, given about 40 years of migration that preceded 1972.

[47] Using later census years would require calculating the statistic from a much smaller sample of black Americans.

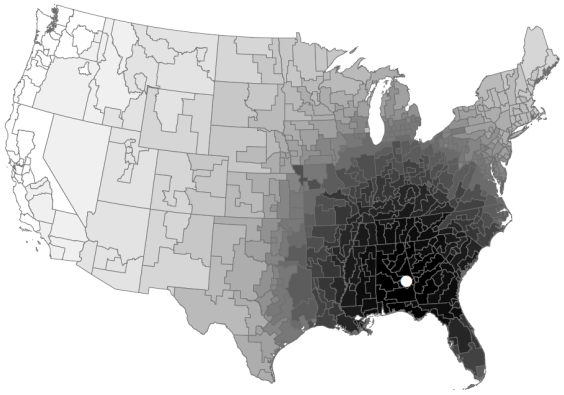
population in their destination location (**?**). We also considered whether formal information networks, such as media exposure, could serve as an alternative treatment, but prefer our proximity and migrant measures for two main reasons. First, defining media exposure is complicated by a dearth of circulation statistics for local black newspapers and the absence of local television and radio archives suitable for content analysis.[48] Second, in our results (discussed below) we find that less educated black men appear to be those most affected by the disclosure and are driving our results. This finding, coupled with the wide dissemination of the Tuskegee study by the Associated Press, suggest to us that information about the study was necessary but not sufficient to modify health-seeking behavior; individuals also had to empathize with the study subjects enough to believe that doctors could treat them in a similar fashion. As such, we prefer to use the distribution of local black newspapers from **?**, as well as the distribution of televisions as a control variable rather than a treatment, specifically to test the hypothesis that the same informational content is more salient for black men than for other demographic groups.[49]

[48] Newspaper circulation figures that are available are rarely audited (this stands in contrast to the more mainstream local daily newspapers from **?**). Our results are robust to including controls for the prevalence of black newspapers (quantified on an extensive margin–see Appendix Table 5). Appendix Section 4 contains more discussion on media coverage of the Tuskegee study.

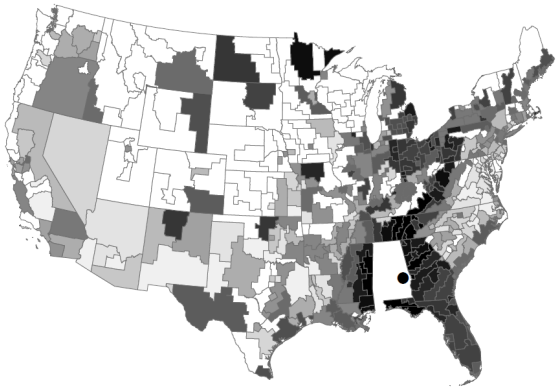
[49] We searched Newspaper Archive as well as other sources for published articles as an alternative measure of the intensive margin. Of the 876 articles about Tuskegee in 1972, 94.5% came from the Associated Press (AP). Thus, access to information about Tuskegee was related to the presence of black publications, not related to independent reporting as essentially none of the coverage was from independent sources. Black publications can therefore be viewed as a proxy for news coverage.

FIGURE 3: PROXIMITY MEASURES

Panel A: Distance to Tuskegee



Panel B: Fraction Black Migrants from Alabama



Notes: Figure 3 plots proximity measures. Panel A represents distance from Macon County for each SEA calculated using ArcGIS proximity tools. Panel B represents the fraction of black migrants from Alabama by SEA calculated using the 1940 census. Darker tones reflect closer to Tuskegee or a higher fraction of Alabama migrants.

7 Empirical Strategy

The goal of our empirical work is to identify the effect of Tuskegee on health-related beliefs, behavior and outcomes. Our hypothesis is that revelation of the Tuskegee syphilis study increased mistrust of the medical system for black Americans who identified most closely with the subjects and consequently reduced their propensity to engage with medical professionals. As predicted by our model, this impact is expected to be greater for groups that had limited exposure to the medical field prior to the disclosure (e.g. black non-veteran males) and groups in closer proximity to the study's subjects.[50]

To test our predictions, we employ a Differences-in-Differences-in-Differences (DDD) estimator with differences taken over time, race or gender, and a measure of proximity. Each of the proximity measures discussed above is based on geography, and we denote these proxies by P_k where $k \in s, a$ represents measurement at the state or SEA level, respectively. Our data contain healthcare utilization and mortality outcomes for individuals (in the utilization data) and demographic groups (in the mortality data). The relevant estimating equation for healthcare utilization for individual i of demographic group g measured in

[50]CITE STATS on how many black women use reproductive services in 1970 and the draft lottery being in place up until the Vietnam War. We are unable to observe veteran status in the mortality data or in the GSS survey data.

state s at time t (Y_{igst}) is:

$$Y_{igst} = \alpha + \beta(P_s * post_t * blackmale_g) + \mu\,(P_s * blackmale_g) + \theta_{gt} + \gamma_{st} + X_{igst}\delta + \epsilon_{igst}. \quad (1)$$

while the estimating equation for the mortality rate of demographic group g measured in SEA a at time t is:

$$Y_{gat} = \alpha + \beta(P_a * post_t * blackmale_g) + \mu\,(P_a * blackmale_g) + \theta_{gt} + \gamma_{at} + \epsilon_{gat}. \quad (2)$$

In both cases, θ_{gt} represents group-year fixed effects and γ_{st} and γ_{at} represent state-year and SEA-year fixed effects, respectively. In regressions at the individual level, X includes controls for income, education, age, whether or not the individual owns a telephone and marital status as well as an indicator for rural-urban status. With these controls, β measures the differential impact of the Tuskegee disclosure on group g, controlling for any differences in the interaction between P and group g in the years prior to 1972, for any geography-specific time effects and for any group-specific time effects.[51]

7.1 Baseline Specifications

We begin with an all-male sample and estimate the impact of TSUS on black males relative to white males in the years following 1972. We use the distance from Macon County, Alabama as the proximity measure (Panel A of Figure 3 above) in the baseline specification, exploring alternatives later in the paper. Our identifying assumption requires there be no other systematic shocks to black men that affect the health outcomes of interest correlated with proximity to Macon County, but not due to the timing of the study's disclosure. As discussed in Section 2.2 on patterns of healthcare utilization, most policy changes coinciding with the timing of the disclosure served to increase access for disadvantaged populations, particularly in the South, and would therefore bias our estimates towards the null. Although access to care and insurance coverage for black and white individuals differed in this period, any time-invariant geographic differences in these factors are absorbed by the $P_s * blackmale_g$ interaction term and any geography-invariant time effects for black men are absorbed by group-year (in this case, race-year) fixed effects. This includes any national policy changes that would have affected black Americans. Still, we note that there is no major expansion

[51] Note that, because the geography fixed effects are at the same level as the distance measure, the triple difference coefficient identifies the differential effect of the health of black versus white men who are at the same distance from Tuskegee before versus after the news, absorbing the main effect of differences between places near versus far from Tuskegee, which is allowed to vary over time.

or contraction of public insurance in or around 1972 that would serve to confound interpretation of our results. Further, **?** estimates minority adult eligibility for Medicaid at 4.5%, a fraction unlikely to drive the results below. Our inclusion of location-year fixed effects nonparametically controls for geography-specific shocks affecting all male residents in a given year, such as the rollout of Medicare and Medicaid and their various expansions. In the Appendix, we also test for the importance of social program growth by reverting to a specification with geography and year fixed effects (instead of geography-year fixed effects) and then measuring how much our coefficients of interest attenuate when we include location-specific expenditures on Medicare, Medicaid and Social Security (Appendix Table 5). Our results are not materially affected by the addition of these controls.[52]

We follow these all-male estimates with a set of all-black estimates, estimating the equations above on a sample of black individuals only. In this case, group-year fixed effects are gender-year effects and control for any time-varying conditions affecting the health of all black men nationwide. In this setting, location-year effects net out any time-varying health advantages or disadvantages associated with geography that affected all black Americans, reducing the scope for unobserved factors to those affecting only black men, but not black women, in closer proximity to Macon County, Alabama.

Our analytical sample is limited to individuals 45 to 74 in order to facilitate comparisons between black men and women, the latter being more likely to seek healthcare in their reproductive years.[53] This stipulation also avoids health-seeking behavior and mortality associated with violence in youth. We limit our analysis on health to a window around the disclosure of 1969-1977 for utilization measures and 1968-1987 for mortality.[54] We do not attempt to model the carry-on effects of TSUS to future cohorts of black men given that our location-based identifying assumption becomes more tenuous as one moves out in time.[55]

To facilitate interpretation of regression coefficients, geographic distance is transformed to proximity to Macon County, Alabama in thousands of kilometers, and coefficient estimates reflect the differential impact of the Tuskegee disclosure on black men relative to white men (or black women) in the years following 1972, per thousand kilometers of proximity to Macon County, relative to the difference between the two groups

[52] Appendix Table 8 also demonstrates that our treatment variables do not predict Medicare, Medicaid or Social Security expenditures.

[53] Appendix Table 1 contains summary statistics for the main variables in our analysis. Individuals had contact with a physician an average of 4.4 times per year. The average mortality rate in our analytic sample is 17.8 persons per thousand per two-year interval, or an annual rate of 8.9 per thousand.

[54] 1968 is the first year mortality data are recorded and NHIS data do not include utilization measures prior to 1969. The slightly longer window for mortality accounts for the lag between health behaviors and death.

[55] In addition, for future cohorts of black men, the AIDS epidemic may confound interpretation if the incidence of HIV is correlated with our treatment intensity variables.

in the years prior for each of the study outcomes.[56] In the mortality data, distances are measured from the geographic centroid of each SEA to the centroid of Macon County, Alabama. For the NHIS outcomes, which are available at the state level, we use county-level black adult population statistics and geographic county centroids to create population-weighted state centroids which are used to construct proximity as above.[57]

Our baseline mortality results contain two other important sample adjustments. First, we restrict the sample to observations where we observe a non-missing mortality rate for both analytic groups in that year. For example, in the all-male estimation sample, we do not include any mortality observations for white male-years if the corresponding black male mortality rate for that two-year period and county is unavailable. Practically, this implies eliminating all-white SEAs from the analytic sample as the primary cause of missing observations is an zero population. We impose this rule consistently throughout the mortality estimation, but it has little impact on the within-black estimates as there are few SEAs with a black male (female) mortality observation without a corresponding observation for black females (males). Second, our preferred specification uses chronic mortality as the outcome of interest as we are interested in deaths from diseases that medical care or advice could have prevented or forestalled. But we also report results for mortality from all causes, and we follow standard practice by age-adjusting these mortality rates to control for differences in the age structure across demographic groups.[58] More details are contained in the appendix.

NEED TO UPDATE THIS STATEMENT. Finally, our baseline mortality estimates are not population weighted, although we report results with population weighting in robustness tests. Because the black male population is highly correlated with our preferred proximity measure (geographic proximity to Tuskegee), weighting by the size of the black population serves to give greater weight to observations closer to Tuskegee and complicates the interpretation of the coefficients of interest.

8 Results

8.1 Event Study Estimates

We first present the results of event study specifications for 1 and 2 in which we generate a coefficient on $P * blackmale$ as well as a coefficient on $P * blackmale$ interacted with each year (except 1972) in lieu of

[56] Specifically, we use $\max(\text{distance}) - (\text{distance}_i)$.

[57] These centroids represent the average latitude and longitude of black individuals in each state based on the black population of counties (**?**).

[58] We use the standard 1940 population as the basis for adjustment. We do not find effects for violence-related mortality, reported in Appendix Table 7.

estimating β. Specifically, we estimate:

$$Y_{igst} = \alpha + \sum_{n \neq 1972} \beta^n (P_s * I_t^n * blackmale_g) + \mu (P_s * blackmale_g) + \theta_{gt} + \gamma_{st} + \epsilon_{igst}, \qquad (3)$$

and a similar equation for mortality. Figure 4, Panel A plots estimates of the difference in the proximity to Tuskegee gradient in health care utilization for treated (black men) versus control (white men) groups *relative to* the difference in 1972 along with 95% confidence intervals. Panel B depicts a similar comparison where the treated group (black men) is compared to a different control group (black women), underscoring that the pattern is not driven by healthcare related supply-side factors that differentially affected all blacks at a given distance from Tuskegee. The pre-Tuskegee disclosure estimates are statistically indistinguishable from zero except for the first year (1969), but there is a statistically significant and sustained change beginning in 1972.

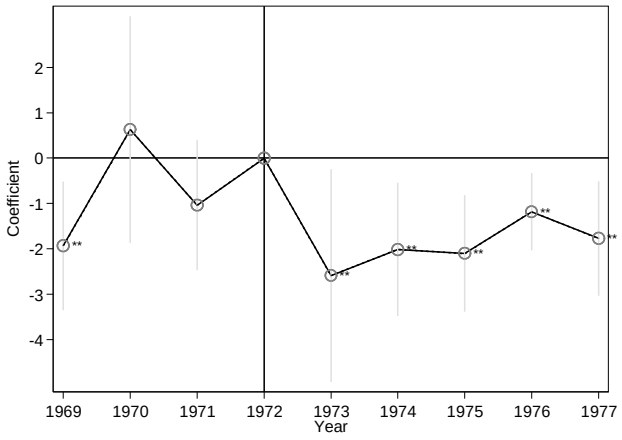
In the next set of figures (Figure 5), we ascertain whether the differences in health-seeking behavior prompted by Tuskegee potentially translated into a widening racial gradient in chronic disease mortality measured in logs. (Recall that the mortality data have been converted to biennial measures). Just as in Figure 4, the plotted β coefficients indicate no discernible difference in the coefficient on $P * blackmale$ in the years leading up to 1972 relative to the same measured in 1972/1973. After the disclosure of Tuskegee, however, the mortality rate for black men develops a steeper geographic gradient as indicated by the β coefficients for years after 1973. The biennial estimates are frequently statistically significant.

FIGURE 4: EVENT STUDY ON UTILIZATION

Panel A: Male Sample



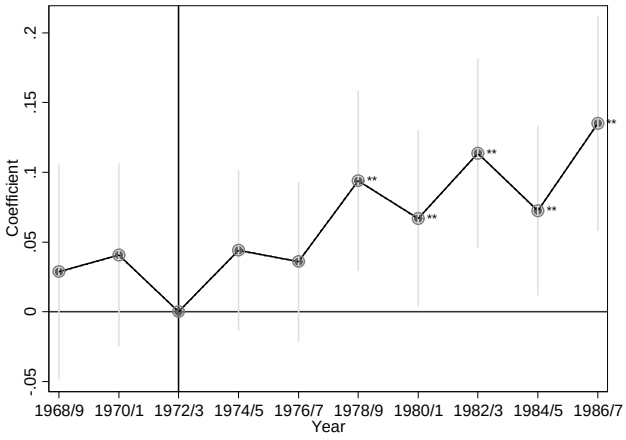
Panel B: Black Sample



Notes: Event study of the coefficient on the interaction of proximity and a black male indicator for each year. The dependent variable is the number of outpatient visits within the last 12 months. Data are from the harmonized version of the National Health Interview Survey (NHIS) available from IPUMS and merged with restricted identifiers for use in the RDC. The sample includes non-veteran males ages 45-74 and black women. Panel A is estimated on an all-male sample and Panel B is estimated on an all-black sample. Plotted are the set of ß coefficients and 95% confidence intervals. Standard errors are clustered at the state level. ,** p<0.05.

FIGURE 5: EVENT STUDY ON MORTALITY

Panel A: Male Sample



Panel B: Black Sample



Notes: Event study of the coefficient on the interaction of proximity and a black male indicator for each two-year period. The dependent variable is log chronic disease mortality for demographic groups ages 45 to 74. Data are from the compressed mortality files provided by the CDC. Panel A is estimated on an all-male sample and Panel B is estimated on an all-black sample. Plotted are the set of ß coefficients and 95% confidence intervals. Standard errors are clustered at the SEA level. ,** p<0.05.

9 Main Estimates

Event study coefficients contained in Figures 4 and 5 are consistent with the idea that the Tuskegee disclosure had an impact on both health-seeking behaviors and health outcomes (namely, mortality). To provide a summary measure of the impact of Tuskegee and to subject our results to a battery of placebo and other robustness checks, we move to reporting the results of the DDD specification in equations 1 and 2. Estimates of β are unbiased if places equidistant from Macon County, Alabama would have had the same evolution of black-white health gradients (or black male v. black female gradients), conditional on covariates, in the absence of the Tuskegee disclosure. For now, we continue to use proximity to Macon County as the proximity measure.

Table 1 reports the coefficient on $P_s * post_t * blackmale_g$ for outcomes related to health seeking behavior estimated using first the all-male sample (panel A, columns 1-4) and the all-black sample (panel B, columns 5-8). We report four outcomes of interest: whether an individual reported any outpatient interactions with a physician in the last 12 months ("Any Outpatient Visit"), how many interactions they reported ("Number Outpatient Visits"), whether an individual reported having been admitted to the hospital in the past 12 months ("Any Hospital Admission") and the length of their hospital stay conditional on admission ("Number Nights in Hospital"). We find sharp declines in the probability of receiving both outpatient and inpatient care.

Estimated coefficients in Column 1 indicate that black men experienced sharp decreases in the probability of visiting a doctor in the years following 1972 of 4 percentage points per 1000 kilometers of proximity Tuskegee. Estimates derived from an all-black sample (Column 5) are slightly smaller, but indicate a reduction of 3.8 percentage points per 1000 kilometers of proximity. Approximately 60% of adult black men lived within 1000 kilometers of Tuskegee in 1970, and we estimate the average treatment effect at roughly 9.5 percentage points.[59] These estimates imply that black men in Alabama reduced their probability of physician interactions by $12 - 13\%$ percentage points (The mean value of this variable has not been disclosed to us yet.) A similarly sizable effect is discernible on the intensive margin: the number of physician visits reported over the past 12 months. Estimated values in Columns 2 and 6 indicate that black men reduced the number of physician visits by $1.4 - 1.9$ visits per 1000 kilometers of proximity to Tuskegee. In this case, we can size the reduction in physician visits at roughly $4.5 - 6.5$ visits for Alabama residents, or a reduction

[59] Calculated by multiplying point estimates by the average proximity measure: 2.54.

yy visits represents the estimated DDD coefficient multiplied by the average proximity from Macon County for black men in the estimating sample.

of 145% off of the sample mean. For black men in Illinois, the implied reduction is $3.5 - 4.5$ interactions per year, approximately 100% of the sample mean.

Turning to hospitalization measures, the estimated β coefficient in Column 3 of Table 1 indicates a post-1972 gradient in the probability of hospital admission for black men of -2.5 percentage points per 1000 kilometers of proximity to Tuskegee. When we estimate the same result for an all-black sample (Column 7), the estimate is attenuated somewhat (a point estimate of -1.7) but remains statistically significant. These results indicate a post-Tuskegee disclosure impact on the likelihood that black men received acute medical care, and the estimated impact for men in Illinois is a reduction of between 33 and 45% of the sample mean while black men in Alabama are estimated to have reduced hospital admission rates by $47 - 62\%$. Although hospital admission rates went down, black men appeared to be have more advanced illness on presentation as reflected in longer hospital stays conditional on admission. Columns 4 and 8 imply that a black male admitted to the hospital experienced an increase in duration of stay of about one-half a night per 1000 kilometers of proximity to Tuskegee. These estimates are large and are likely the result of two reinforcing effects. First, the sample of black men admitted to the hospital is likely to be negatively selected given the overall reduction in hospital admissions. Second, the reduction in primary care utilization documented in Columns 1-2 and Columns 5-6 likely led to more serious or advanced conditions at the time of presentation.

In Appendix Table 2, we estimate this baseline equation on another healthcare utilization outcome less likely to have been affected by the Tuskegee disclosure: dental visits over the last 12 months. This outcome is available for a slightly truncated period of time (1969-1975), and shows black men in closer proximity to Tuskegee, Alabama are slightly *more* likely to visit the dentist after 1972. This result provides suggestive evidence that racial gaps in other forms of care stabilized or even converged after 1972 and that relative avoidance was specific to institutions most reminiscent of the Tuskegee study.[60]

The NHIS data are rich, and we can use additional information about men in the sample to explore heterogeneous effects in the utilization effects documented in Table 1. For now, we turn to the mortality data to estimate treatment effects akin to the utilization impacts in Table 1.

[60] We also find some limited evidence that black men were more likely to smoke than their white peers, see Appendix Table 2. Note a harmonized smoking variable is available in only some of the relevant years.

TABLE 1: BASELINE UTILIZATION RESULTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	*Any Outpatient Visit*	*Number Outpatient Visits*	*Any Hospital Admission*	*Number Nights in Hospital*	*Any Outpatient Visit*	*Number Outpatient Visits*	*Any Hospital Admission*	*Number Nights in Hospital*
	PANEL A--Male Sample				**PANEL B---Black Sample**			
P_s*$post_t$*$blackmale_g$	-0.040*** (0.012)	-1.894*** (0.414)	-0.025** (0.010)	0.569* (0.303)	-0.038*** (0.007)	-1.415*** (0.442)	-0.017* (0.008)	0.697** (0.330)
Fixed Effects	State-Year , Race-Year				State-Year , Gender-Year			
Observations	60,837	60,837	60,837	60,837	18,966	18,966	18,966	18,966
No Clusters	49	49	49	49	44	44	44	44
Mean of Dept. Var.	Not Disclosed	4.32	0.13	2.04	Not Disclosed	5.15	0.12	1.98
Adj R-squared	0.029	0.023	0.019	0.016	0.059	0.047	0.028	0.026
California	Not Disclosed	-7%	-3%	4%	Not Disclosed	-4%	-2%	6%
Illinois		-103%	-45%	66%		-65%	-33%	83%
Tennessee		-134%	-59%	85%		-84%	-43%	108%
Alabama		-145%	-62%	95%		-92%	-47%	117%

Notes: OLS Regressions of equation 1. Data are from the harmonized version of the National Health Interview Survey (NHIS) available from IPUMS and merged with restricted identifiers for use in the Restricted Data Center (RDC). The sample includes non-veteran males ages 45-74 and women. Panel A is restricted to men and Panel B is restricted to black individuals. The outcome varies by column and is given by column heading. In columns (1) and (5) the outcome is an indicator variable for any outpatient physician interaction in the past 12 months. In columns (2) and (6) the outcome variable is the number of outpatient physician interactions in the last 12 months. In columns (3) and (7) the outcome variable is an indicator variable for any hospital admission in the last 12 months. In columns (4) and (8) the outcome variable is the number of nights hospitalized in the last 12 months. Controls in every specification include indicator variables for educational status, income category, age, marital status, telephone ownership as well as rural/urban status, the interaction between proximity to Tuskegee and black or male, and black-year or male-year and state-year fixed effects. Regressions are weighted using provided survey weights. Standard errors are clustered at the state level. The triple difference calculated at the black population centroid of various states is shown in the last four rows of each panel. The mean of the dependent variable for the sample in the is also reported. Standard errors are clustered at the state level. ***p<0.01 ,** p<0.05 and *p<0.10, respectively.

Our baseline mortality results, estimates of β from equation 2, are presented in Table 2 . As in Table 1, we report results for an all-male sample and an all-black sample in panels A and B, respectively. Our outcomes include both age-adjusted and cause-specific chronic mortality, as motivated above, and we report results for both the full sample of SEAs and for southern SEAs only. The South-only specifications are motivated by the occasional mortality outliers in the western U.S. apparent in Figure 2; South only specifications assure that a geographic gradient in mortality is apparent within the South and not just as a result of comparing western SEAs to southern SEAs.

Coefficient estimates in the first column of Table 2 represent a comparison of the post-1972 difference in log age-adjusted mortality between black and white older men, relative to the pre-1972 difference and as a function of proximity to Macon County, Alabama.[61] Mortality patterns, like utilization, appear to have moved adversely against blacks after 1972 along a gradient of proximity to Macon County, Alabama, and we estimate a 4.3 log point increase in the mortality rate of black men per 1000 kilometers of proximity in Panel A. Using average values of P_a and of the outcome variable, we estimate an overall increase in older male mortality in Alabama of 5% off of the pre-1972 mean and an increase in Illinois of 3%.[62] These calculations implicitly assume mortality in these locations would have followed patterns apparent in California (where P_a is 0) in the absence of treatment, and our treatment effects are calculated relative to this counterfactual. When we estimate the model using an all-black sample, we find even larger results: a statistically significant 8.5 log points of mortality increase per 1000 kilometers for a calculated effect of 9% in Alabama and 7% in Illinois. These are sizable effects, but not fully out of line with other findings in the literature. Results from **?** show that exposure to a community health center (CHC) in one's county of residence between 1965 and 1974 reduced adult all-cause mortality by as much as 2%. Of course, not all individuals in a county would have utilized a CHC, just as not all individuals in an SEA would have been aware of the Tuskegee study, making comparisons between these two coefficients difficult. Still, the study's results indicate sizable mortality effects related to the utilization of primary care facilities.

The gradient of post-1972 older black male mortality as a function of proximity to Tuskegee is apparent for a restricted sample of southern SEAs as well, an indication that the nationwide results are not a result of mortality differences in the South versus the rest of the country. Because we restricted the nationwide estimates to those SEAs containing black and white male mortality results, and because SEA density is

[61] Our results are fully robust to a levels specification, as reported in Appendix Table 4.

[62] INSERT CALCULATION HERE.

higher in the southern U.S. that elsewhere, the southern restriction leaves more than 40% of the original sample available for estimation For these SEAs, coefficient estimates contained in columns 2 (5.2 log points) and 6 (9.5 log points) of Table 2 show slightly steeper gradients for SEAs within the South, consistent with the observations from Figure 2. As further evidence that the documented effects are not driven by a post-1972 general southern penalty, later in the paper we will show that the gradient we observe is specific to proximity to Macon County and the calculated effect is much weaker when we use proximity to other southern SEAs as placebo proximity measures.

Estimates for chronic mortality, contained in columns 3-4 and again in columns 7-8 are our preferred specification given the hypothesized pathways by which the revelation of Tuskegee would have affected mortality. For the all-male sample, we measure increases in log mortality of 6.3 log points and 10.3 log points for the full sample and the southern sample, respectively. Results for the all-black sample are slightly larger: 10.2 log points for the full sample and 13.4 log points for the South-only sample. Larger log specification results for chronic mortality compared to all age-adjusted mortality is to be expected if the Tuskegee disclosure primarily affects chronic mortality as the former is a subset of the latter. We measure the mean of log mortality at 3.07 (21 deaths per thousand per year) and of log chronic mortality at 2.82 (17 deaths per thousand per year) for the all-male sample, and our estimates of β correspond to an increase of approximately 1 death per year per thousand (and per thousand kilometers of proximity) in both cases For the all-black sample, baseline death rates are 19.7 per thousand from all causes and 14.6 per thousand from chronic conditions, reflecting lower mortality for black women. Our estimates of β in the all-black sample correspond to increases of roughly 1.5 deaths per thousand per year in both chronic and age-adjusted mortality measures.

TABLE 2: BASELINE MORTALITY RESULTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Outcome	*Log Age-Adjusted Mortality*		*Log Chronic Mortality*		*Log Age-Adjusted Mortality*		*Log Chronic Mortality*	
	Full Sample	South Only	Full Sample	South Only	Full Sample	South Only	Full Sample	South Only
	PANEL A--Male Sample				**PANEL B---Black Sample**			
P_a*$post_t$*$blackmale_g$	0.043***	0.052*	0.063***	0.103***	0.085***	0.095***	0.103***	0.134***
	(0.016)	(0.028)	(0.018)	(0.037)	(0.024)	(0.034)	(0.021)	(0.046)
Fixed Effects	SEA-Year, Race-Year				SEA-Year, Gender-Year			
Observations	8568	3708	8400	3694	8126	3694	7916	3668
No. Clusters	455	186	462	186	440	186	437	186
Mean of Dept. Var.	3.07	3.14	2.82	2.86	2.98	3.04	2.68	2.74
Adj R-squared	0.43	0.68	0.27	0.58	0.56	0.78	0.55	0.74
California	0%	0%	0%	1%	0%	1%	1%	1%
Illinois	3%	4%	5%	8%	7%	7%	9%	12%
Tennessee	4%	5%	7%	11%	9%	10%	12%	15%
Alabama	5%	6%	8%	12%	10%	11%	13%	16%

Notes: OLS Regressions of equation 2. Data are from the CDC compressed mortality files on the non-institutionalized US population and are collapsed into two year bins across state economic areas (SEA). Panel A restricts the sample to white and black male demographic groups and Panel B restricts the sample to black male and female demographic groups. In columns (1) (2) (5) and (6) the outcome is the log of the age adjusted mortality rate for those ages 45-74. In columns (3) (4) (7) and (8) the outcome is log of the cause-specific mortality among the 45-75 population from chronic disease, including cardiovascular disease, cancer, smoking-related respiratory disease, gastrointestinal disease and diabetes. Mortality regressions that are population-weighted or in levels are provided in robustness and appendix tables, respectively. Odd numbered columns are estimated on the continental US and even numbered columns are restricted to the South (Census Region III). Controls included in every specification include the interaction between proximity to Tuskegee and black or male and black-year or male-year and sea-year fixed effects. The triple difference calculated at the black population centroid of various states is shown in the last four rows of each panel. The mean of the dependent variable for the relevant sample is also reported. Standard errors are clustered at the SEA level. ***p<0.01 ,** p<0.05 and *p<0.10, respectively.

TABLE 3: HETEROGENEOUS EFFECTS, UTILIZATION

Number Outpatient Visits						
	(1)	(2)	(3)	(4)	(5)	(6)
	By Income Level		By Educational Status		By Prevalence Black Doctors	
	income> median	*income <= median*	*education> median*	*education<= median*	*black MD > median*	*black MD <= median*
PANEL A---Male Sample						
P_s*$post_t$*$blackmale_g$	-1.142**	-2.973***	-0.246	-3.263***	-1.808***	-2.749***
	(0.507)	(0.623)	(0.296)	(0.642)	(0.424)	(0.685)
Observations	41,368	19,469	34,290	26,547	34,017	26,820
No Clusters	49	49	49	49	17	32
Adj R-squared	0.019	0.047	0.022	0.036	0.023	0.024
PANEL B--Black Sample						
P_s*$post_t$*$blackmale_g$	-0.536	-1.892**	0.640**	-3.316***	-1.441***	-2.072**
	(0.557)	(0.752)	(0.297)	(0.976)	(0.463)	(0.791)
Observations	8,551	10,415	8,616	10,350	14,219	4,747
No Clusters	41	43	43	43	17	27
Adj R-squared	0.051	0.065	0.063	0.070	0.043	0.074

Notes: OLS Regressions of equation 1 assessing heterogeneous effects to a) income level b) education level and c) black doctors per capita. Specifically, in the first two columns we divide the sample by median black male income (above (1) or below (2)). In the following two columns we divide the sample by median black male education (above (3) or below (4)). In the last two columns, we calculate black physicians as a percentage of all physicians using occupational data from the Census. The outcome variable across all panels and columns is the number of physician interactions in the last 12 months. Controls in every specification include indicator variables for educational status, income category, age, marital status, telephone ownership as well as rural/urban status, the interaction between proximity to Tuskegee and black or male and black-year or male-year as well as state-year fixed effects (except in column (4)). Regressions are weighted using provided survey weights. Standard errors are clustered at the state level. ***p<0.01 ,** p<0.05 and *p<0.10, respectively.

9.1 Heterogeneous Effects

Our main estimates of post-1972 utilization effects for black men documented in Table 1 mask substantial heterogeneity across locations and individuals. In Table 3, we document this heterogeneity by undertaking a series of sample bifurcations informed by theory and reporting estimates of β for these subsamples in a regression equation First, we explore the roles of income and education in the post-1972 health-seeking behavior of black men and bifurcate the all-male and all-black samples at the median of black male income (columns 1 and 2) and at the median of black male education (columns 3 and 4). In both cases, we hypothesize that black men lower on the socioeconomic ladder would respond more strongly to the news of Tuskegee because they were in closer cultural proximity to the poor sharecroppers who were the Tuskegee study's victims. Indeed, although the estimates of β are negative and significant for both income groups in the all-male sample, the point estimates for lower income black men in column 2 of panel A are substantially larger. Similarly, for the all-black sample, we find no evidence of a post-1972 change in the health-seeking behavior of black men relative to black women for higher income black men (column 1 of panel B), but a larger and statistically significant effect for men with lower incomes (column 2 of panel B) Splitting the sample by educational attainment generates a similar result (columns 3 and 4).

Next, we examine the moderating effect of black physicians on our baseline results in columns 5 and 6 of Table 3. We hypothesize that the availability of a black physician would have reduced the rate at which black men downgraded their expectation of encountering a "good" doctor. For these results, we collected data on the number of black and white physicians in each U.S. state from the 1970 U.S. Census reports. Importantly, these counts are measured before, and therefore not endogenous to, the 1972 disclosure. These numbers are not available at the SEA level, and so we do not perform a similar analysis for mortality results. But for the measures of primary care utilization, when we bifurcate the sample of men and the sample of blacks by their residence in states above and below the median number of black doctors as a percentage of all doctors, we find a substantial moderating effect. The presence of more black doctors in a state is associated with a reduction in the post-1972 reduction in outpatient visits of between 0.5 and 1.0 visits per 1000 kilometers of proximity to Tuskegee off of the full-sample estimated reduction of 1.4 to 1.8 visits per 1000 kilometers.

10 Threats to Identification

We perform several tests to assure a causal interpretation for our results. These results can be divided into a set of placebo tests and a set of robustness tests. First, we use placebo locations to show that the main results are specific to gradients of proximity to Macon County, Alabama. Second, we use placebo populations to demonstrate that the main results for both utilization measures and mortality rates do not obtain when we estimate the baseline equations on younger population samples, namely children aged 1-9. We also show in these placebo population results that neither black women nor white men show the same geographic patterns of mortality divergence after 1972. Third, we test the robustness of our estimates by incorporating migration as an alternative measure and demonstrating similarly signed and statistically significant estimates. We also show that our results are robust to modifications of the geographic proximity measure, to incorporating additional fixed effects, to controlling for distance to other Southern locations and to population weighting (in the mortality results). We describe each of these tests in turn below.

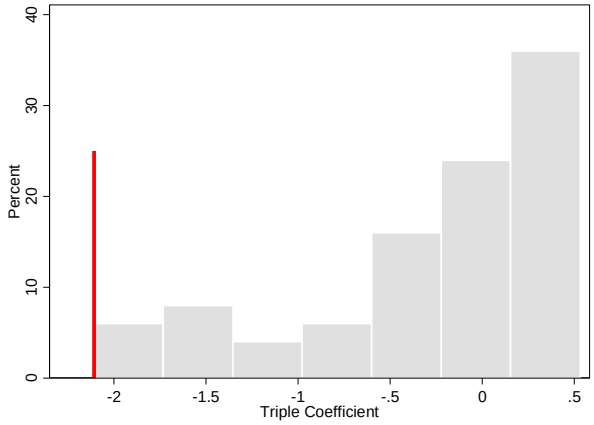
10.1 Placebo Locations

Our baseline estimates, coupled with the lines of best fit in Figure 2, indicate a stark geographic gradient of post-1972 utilization and mortality based on proximity to Macon County, Alabama. South-only results indicate that the mortality results hold for a sample of southern SEAs, and we show similar results for utilization in Table 5 below. We expand on this robustness exercise in Figure 6. For the intensive margin of primary care utilization, we run 47 placebo regressions, substituting for the baseline proximity measure (proximity to Macon County, Alabama) with the proximity to the geographic centroid of every other state in the continental U.S. and re-estimating the model. These tests serve as placebo tests, evaluating whether we find the same (or stronger) utilization effects as a function of the gradient to other locations in the U.S. The top portion of Figure 6 presents the distribution, in histogram form, of the estimated values of β in each of these tests. The vertical red line indicates the estimated coefficient when the geographic proximity measure is based on the SEA housing Macon County, Alabama. The absolute value of the estimate is greater than 96% of placebo estimates for the within-male specification and 98% of estimates for the male-only and black-only specification. We perform the same exercise in the mortality data by substituting the baseline geographic proximity measure (proximity to Macon County, Alabama) with the distance to each of the other SEAs in the mortality dataset and re-estimating equation 2 using log chronic mortality as the outcome.

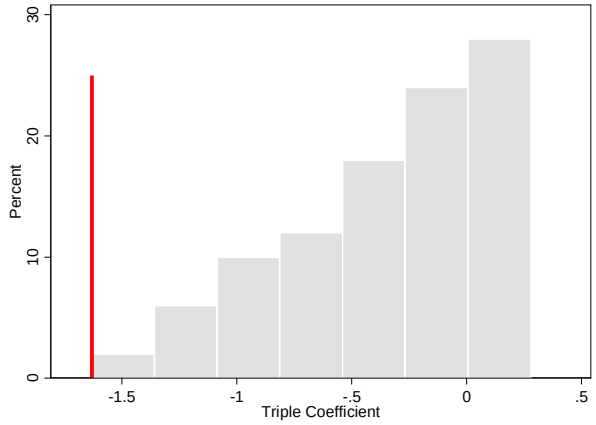
Again, these tests serve as placebos, measuring whether the post-1972 mortality gradients are apparent when we use other U.S. locations as the placebo location of the syphilis study. These results, located in the bottom portion of Figure 6, indicate that the value of β estimated using proximity to Macon County as the center of the geographic proximity metric is greater than 93% (98%) of all other coefficients for the all-male (all-black) samples. A "heat map" of our estimates for Panel C and D is provided in Appendix Figure 3.

FIGURE 6: PLACEBO DISTANCES (EMPIRICAL DISTRIBUTION)

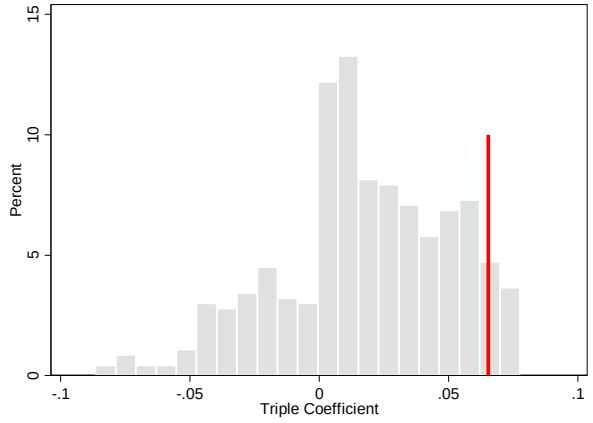
Panel A: Male Sample (Number of Outpatient Visits)



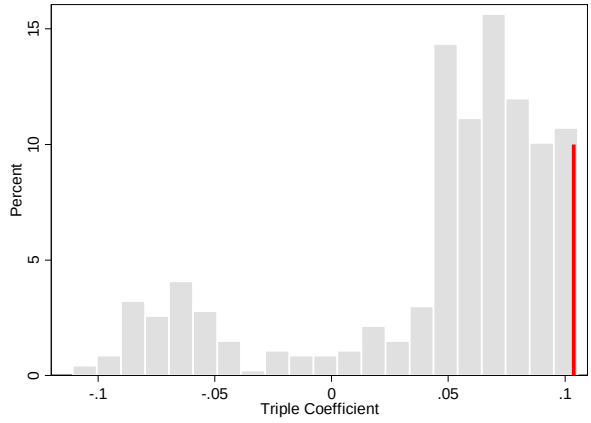
Panel B: Black Sample (Number of Outpatient Visits)



Panel C: Male Sample (Log Chronic Mortality)



Panel D: Black Sample (Log Chronic Mortality)



Notes: Frequency of the β coefficient using distance from every other state or SEA in the sample and estimating equation (1) in Panels A and B or equation (2) in Panels C and D. The vertical (red) line denotes the β_1 coefficient estimate for the true treatment distance.

10.2 Other Demographic Groups

Our theoretical model indicates that individuals in closer cultural proximity to the study's victims should have been more affected by the study's revelation. This motivates our baseline specifications, comparing black men to both white men and black women whom we hypothesize would have been more weakly affected. In this section, we use this prediction to perform additional placebo tests, estimating the baseline utilization regressions on a set of placebo populations who are more distant in their cultural proximity to the Tuskegee study's victims, necessarily restricting our presentation to a subset of utilization outcomes for space reasons. **(The remainder are located in the Appendix. NO THEY ARE NOT. WE HAVE HOSPITALIZATION BUT ONLY FOR ADULTS NOT KIDS, WE'D HAVE TO RENUMBER APPENDIX TABLES, REDO LABELING IN FNs and TEXT MAYBE JUST SAY AVAILABLE ON REQUEST.)** If successful, these placebo tests help rule out confounding factors that may be confounding interpretation of our baseline coefficients as discussed below.

Table 4 contains estimates of β for primary care outcomes: "any outpatient visit" in Panel A and "number outpatient visits" in Panel B. We first limit the sample to all male children aged 1-9 and show that the post-1972 difference in outpatient care between black and white male children exhibits no geographic gradient, as evidenced by coefficients statistically indistinguishable from zero in columns 1 and 5.[63] The same is true for black male children relative to their female peers; the estimates for an all-black sample are contained in columns 2 and 6. The non-result for black male children is consistent with either a lower cultural proximity to the study's victims for younger children and with results to follow showing that black women were not affected by the study's disclosure. Although the perceptions of their fathers could feasibly have affected the demand for children's healthcare utilization, the non-result here may indicate that decision-making for children at this time was mostly driven by maternal preferences.

Our baseline within-black specifications implicitly tests for differential effects for black males relative to their female peers. Here, we test for a post-1972 effect of the Tuskegee disclosure on black women, explicitly by estimating equation 1 above, replacing $P_s * post_t * blackmale_g$ with $P_s * post_t * blackfemale_g$ and using an all-female sample comparing black women to their white peers. The resulting estimates for β are contained in columns 3 and 7. For both an extensive and intensive margin of primary care utilization, we see no statistically significant post-1972 change in the health-seeking behavior of black females as a function

[63] Results for adolescents and teens, aged 11-20, are similar but not disclosed.

of proximity to Tuskegee. We perform a similar analysis for white men, replacing $P_s * post_t * blackmale_g$ with $P_s * post_t * whitemale_g$ in equation 1 and performing estimation over an all-white sample. Estimates for β, contained in columns 4 and 8, show no post-1972 gradient (as a function of proximity to Macon County) in primary care utilization for white men relative to their female peers.

We perform a symmetric set of tests on age-specific and age-adjusted mortality results in the bottom panels of Table 4 and a similar set of non-results obtains. Mortality rates for the population aged 1-9 cannot be age-adjusted, and we report results for age-specific mortality for this group. For the all-female and all-male samples, we report coefficients from age-adjusted mortality regressions; chronic mortality coefficients are not remarkably different from those reported here. In all but one case, we observe post-1972 gradients by proximity to Tuskegee that are insignificantly different from zero.[64] The exception is the coefficient on $P_a * post_t * whitemale_g$which is positive and statistically significant, but economically insignificant at 0.7 log points of increase per 1000 kilometers of proximity to Macon County. This positive result is also not robust; it is not significant in the levels specification contained in the Appendix or in a sample restricted to the South. For black women, relative to their white peers, the point estimates, although insignificant, are negative and consistent with the upward-sloping line in Panel C of Figure 2. Taken together, the scope of identification threats is narrowed substantially to those factors correlated with the primary care utilization and mortality rates of black men in close geographic proximity to Macon County, Alabama post-1972, that did not also affect older black women, older white males and black male children.

[64]Infant mortality results (not shown) indicate a post-1972 decrease in the mortality rates of black male infants relative to white, a result consistent with other literature on mortality convergence in this era. In this case, the gradient is not specific to Macon County, Alabama.

TABLE 4: PLACEBO UTILIZATION AND MORTALITY RESULTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample	*Male Children*	*Black Children*	*Female*	*White*	*Male Children*	*Black Children*	*Female*	*White*
	PANEL A---Any Outpatient Visit				**PANEL B---Number Outpatient Visits**			
P_s*$post_t$*$blackmalechild_g$	0.002	0.010			0.015	-0.119		
	(0.007)	(0.008)			(0.093)	(0.119)		
P_s*$post_t$* $blackfemale_g$			0.004				-0.127	
			(0.006)				(0.140)	
P_s*$post_t$* $whitemale_g$				0.004				-0.008
				(0.005)				(0.093)
Fixed Effects		State-Year, Group-Year				State-Year, Group-Year		
Observations	104,071	28,950	150,864	192,735	104,071	28,950	150,864	192,735
No Clusters	49	46	49	49	49	46	49	49
Adj R-squared	0.109	0.155	0.018	0.022	0.046	0.068	0.016	0.014
	PANEL C--Log Age-Specific and Age-Adjusted Mortality (Full Sample)				**PANEL D--Log Age-Specific and Adjusted Mortality (South)**			
P_s*$post_t$*$blackmalechild_g$	-0.047	-0.023			-0.022	-0.064		
	(0.029)	(0.042)			(0.070)	(0.111)		
P_s*$post_t$* $blackfemale_g$			-0.021				-0.050	
			(0.021)				(0.037)	
P_s*$post_t$* $whitemale_g$				0.007**				-0.005
				(0.003)				(0.014)
Fixed Effects		SEA-Year, Group-Year				SEA-Year, Group-Year		
Observations	6274	4946	8306	9300	3310	2884	3698	3720
No Clusters	415	346	451	465	183	171	186	186
Adj R-squared	0.22	0.53	0.53	0.97	0.31	0.57	0.78	0.97

Notes: OLS Regressions of equation 1 (Panels A and B: Utilization Outcomes) and equation 2 (Panels C and D: Mortality Outcomes) on demographic groups exclusive of older black men for whom we predict the effect of the Tuskegee experiment disclosure should be more muted. The specific sample is given by the column heading. Columns (1) and (2) in the utilization data refer to outpatient physician interactions for children below age 10, whereas the age-specific mortality rates in corresponding columns for Panels C and D exclude infant deaths (thus are estimated on children ages 1-9). Columns (3) and (7) are estimated on a female sample across all panels and columns (4) and (8) are estimated on a sample of white males and females across all panels. Data in Panels A and B are from the harmonized version of the National Health Interview Survey (NHIS) available from IPUMS and merged with restricted identifiers for use in the Restricted Data Center (RDC). Data in Panels C and D are from the CDC compressed mortality files on the non-institutionalized US population and are collapsed into two year bins across state economic areas (SEA). In Panel A the outcome is an indicator variable for any outpatient physician interaction in the past 12 months. In Panel B the outcome variable is the number of outpatient physician interactions in the last 12 months. In Panel C the outcome variable is either the log of the age-specific mortality rate for children ages 1-9 (columns 1 and 2) or the age-adjusted mortality rate for adults ages 45-75 on the female (column 3) or white (column 4) samples. Panel D includes the same outcome variables restricted to the South (Census Region III). Controls included in the utilization regressions include indicator variables for educational status, income category, age, marital status, telephone ownership as well as rural/urban status, the interaction between proximity to Tuskegee and black or male, and black-year or male-year and state-year fixed effects. Regressions are weighted using provided survey weights. Controls included in the mortality regressions include the continuous interaction between distance and black and black-year or male-year and sea-year fixed effects. Mortality regressions that are population-weighted or in levels are provided in robustness and appendix tables, respectively. The triple difference calculated at the black population centroid of various states is shown in the last four rows of each panel. The mean of the dependent variable for the sample in the prior to the Tuskegee disclosure period is also reported. Standard errors are clustered at the State level for Panels and B and the SEA level for Panels C and D. ***p<0.01 ,** p<0.05 and *p<0.10, respectively.

10.3 Robustness Checks

Next, we demonstrate that the baseline results are robust to alternative proximity measures. For this next set of regressions, all-male samples are contained in Panel A and all-black samples in Panel B. Table 5 reports these checks for the utilization results regarding number of outpatient visits and Table 6 provides analogous results for log chronic mortality. Results are similar for age-adjusted mortality and for other utilization measures. The tables are symmetric with the exception of the final column, and we will discuss them in parallel.

In columns 1 and 6 of Table 5, we replace the baseline proximity measure, geographic proximity to Macon County, Alabama, with the percentage of 1935-1940 black migrants to a particular state or SEA who emanated from Alabama. (All Alabama SEAs and the state of Alabama are excluded from this analysis.) We observe statistically significant differences in the post-1972 utilization of primary healthcare and in the post-1972 mortality rates for black men as a function of this proxy. Again, the interpretation of the magnitudes of these coefficients differs from the baseline results using geographic proximity to Macon County, Alabama as the proximity measure. In columns 2 and 7, we replace the continuous geographic proximity measure with proximity "bins", defined as quintiles of proximity to Macon County. The resulting coefficient estimates carry a different interpretation from those in the baseline results as they now measure the post-1972 impact for black males per quintile bin, which is approximately 720 kilometers. As a result, the coefficients should be smaller than in the baseline specification, precisely as we observe.

Columns 3, 4, 8 and 9 of Table 5 provide results for alternative specifications of our baseline results. In column 3, we add group-location (state or SEA) fixed effects in lieu of the $P_s * blackmale_g$ interaction term in equation 1 and in lieu of $P_s * blackmale_g$ in equation 2. In the all-male sample contained in Panel A of both tables, the blackmale-state or blackmale-SEA fixed effect absorbs any time-invariant location-specific conditions that affected black but not white men in a non-parametric manner. (In the baseline specification, these controls are in place as a function of proximity to Macon County, Alabama.). In the all-black sample, the corresponding fixed effects absorb any time-invariant location-specific shocks that affect black men but not black women. These fixed effects nonparametrically control for the presence of government programs which might have differentially enrolled black males, provided those enrollment patterns are static over time. In Column 4 of both tables, we add distance from a placebo location, Dallas, Texas, to our estimating equations, continuing to control for geographic proximity to Macon County, Alabama. If the patterns we

observe are reflective of a more general "southern" phenomenon after 1972, our estimates should attenuate when we also include measures of proximity to Dallas.

In each of these cases, we fail to detect material differences in estimated values of β relative to the baseline. The estimated value of β in the baseline specification contained in Table 1 is -1.9 for the all-male sample and 1.4 for the all-black sample when the outcome is the number of outpatient visits. Robustness check results in Table 5 are in the range of -1.2 to -1.8 for the within-male sample and -0.8 to -1.4 using the all-black sample (columns 7-10).[65] For mortality, the baseline estimate of β for log chronic mortality is 0.063 in the all-male sample and 0.102 in the all-black sample. Results in Table 6 range from 0.048 to 0.077 for the all-male sample and 0.061 to 0.109

Finally, column 5 of Table 5 restricts the utilization analysis to southern residents only. This reduces the scope for identification threats to things correlated with geographic proximity to this particular location in the South and not with the South in general. The estimated value of β is relatively unchanged in this specification compared to our baseline results although the estimates are no longer statistical significant.

In column 5 of Table 6, we weight our baseline estimates by the size of the older black population in each SEA, dropping the SEAs corresponding to the six largest black media markets.[66] Weighting has the additional advantage of implicitly assuring the majority of our results are driven by within-South variation since approximately 50% of the black male population resided in the South at this time. Finally, population weighting helps address the noise in the mortality data (since death rates are more consistently measured in populous places). We find that our mortality results are robust to incorporating population weights in log chronic mortality for both the all-male and all-black analytic samples.

As a final word on threats to identification, we note that the geographic mobility of individuals may serve to complicate the interpretation of β. We cannot identify where individuals were at the time of exposure, but we note that mobility was relatively limited in this period (which followed the Great Migration) and was generally short-distance. Still, it would be preferable to observe outcomes that are correlated with a fixed geographic proximity to Tuskegee, Alabama in 1972 and not in the years following. In Section 12 below, we utilize the only data we have with this information (GSS data on attitudes and perceptions) and regress

[65] We convert the magnitudes of the coefficients in column 3 to values comparable to our baseline results for this exercise, multiplying by 1000/720.

[66] Consistent with Panel C in Table 3 and our hypothesis regarding information transmission via black newspapers, the largest media markets show sharp increases in black male mortality after 1972. These cities (San Francisco, Los Angeles, Chicago, Washington DC, New York, and Atlanta), are generally not in close geographic proximity to Macon County and represent mere outliers in an unweighted regression but dominate the estimation with a weighted sample because of population sizes in these locations.

perceptions of physicians' judgment in 1998 on the location of an individual at the time of TSUS revelation in 1972.

TABLE 5: ROBUSTNESS CHECKS, UTILIZATION

Number Outpatient Visits

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Alternative Instruments		*Alternative Specifications*		*Sample Restrictions*	*Alternative Instruments*		*Alternative Specifications*		*Sample Restrictions*
	Migrants	Bins	Addnl Fixed Effects	Placebo Distance (Dallas)	South Only	Migrants	Bins	Addnl Fixed Effects	Placebo Distance (Dallas)	South Only
	PANEL A --- Male Sample					**PANEL B --- Black Sample**				
P_s*$post_t$*$blackmale_g$	-10.34**	-1.234***	-1.800***	-1.617***	-1.838	-9.506**	-0.829**	-1.389***	-1.334***	-1.459
	(4.575)	(0.276)	(0.352)	(0.350)	(1.437)	(4.386)	(0.360)	(0.410)	(0.361)	(1.591)
$P_{placebo}$*$post_t$*$blackmale_g$				-0.626					-0.184	
				(0.463)					(0.556)	
Fixed Effects	State-Year, Race-Year	State-Year, Race-Year	State-Year, Race-State, Race-Year	State-Year, Race-Year	State-Year, Race-Year	State-Year, Gender-Year	State-Year, Gender-Year	State-Year, Gender-State, Gender-Year	State-Year, Gender-Year	State-Year, Gender-Year
Observations	59,738	60,837	60,837	60,837	18,574	18,076	18,966	18,966	18,966	10,587
No Clusters	48	49	49	49	17	43	44	44	44	17
Adj R-squared	0.022	0.023	0.024	0.023	0.029	0.047	0.047	0.049	0.047	0.041
p-value			Not Disclosed					Not Disclosed		

Notes: OLS Regressions of equation 2 testing robustness to a) alternative treatments b) alternative specifications and c) alternative weighting. Specifically, in columns (1) and (6) we use the fraction of black migrants that are from Alabama, constructed using the 100% 1940 Census from IPUMS (see text for details). In columns (2) and (7) we use a nonparametric measure of distance (quantiles) instead of linear distance as the predictor variable. In columns (3) and (8) we run a fully saturated specification replacing group*distance with group-location fixed effects. In columns (4) and (9) we include proximity to Dallas and its second-order interactions along with proximity to Tuskegee to test whether the effects are driven by being near a city in the South. In columns (5) and (10) we assess robustness to weighting each cell by the size of the relevant older age population. The data source and sample were described in the notes to Table 2. The outcome variable is the log of chronic disease mortality. Controls in every specification include the interaction between proximity to Tuskegee and black or male and black-year or male-year as well as state-year fixed effects (except in columns (3) and (8)). Standard errors are clustered at the SEA level. ***p<0.01 ,** p<0.05 and *p<0.10, respectively.

TABLE 6: ROBUSTNESS CHECKS, MORTALITY

Log Chronic Disease Mortality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Alternative Instruments		*Alternative Specifications*			*Alternative Instruments*		*Alternative Specifications*		
	Migrants	Bins	Addnl Fixed Effects	Placebo Distance (Dallas)	Popln-Weighted	Migrants	Bins	Addnl Fixed Effects	Placebo Distance (Dallas)	Popln-Weighted
	PANEL A --- Male Sample					**PANEL B --- Black Sample**				
P_a*$post_t$*$blackmale_g$	0.176***	0.035***	0.077***	0.054***	0.055***	0.187***	0.044***	0.081***	0.109***	0.087***
	(0.036)	(0.007)	(0.018)	(0.020)	(0.011)	(0.061)	(0.009)	(0.019)	(0.022)	(0.013)
$P_{placebo}$*$post_t$*$blackmale_g$				0.021					-0.014	
				(0.020)					(0.022)	
Fixed Effects	SEA-Year, Race-Year		SEA-Year Race-SEA Race-Year	SEA-Year, Race-Year		SEA-Year, Gender-Year		SEA-Year Gender-SEA Gender-Year	SEA-Year, Gender-Year	
Observations	8426	8400	8376	8400	8400	7952	7916	7882	7916	7916
No. Clusters	461	462	450	462	462	437	437	420	437	437
Adj R-squared	0.27	0.27	0.49	0.27	0.86	0.54	0.55	0.56	0.55	0.92

Notes: OLS Regressions of equation 2 testing robustness to a) alternative treatments b) alternative specifications and c) alternative weighting. Specifically, in columns (1) and (6) we use the fraction of black migrants that are from Alabama, constructed using the 100% 1940 Census from IPUMS (see text for details). In columns (2) and (7) we use a nonparametric measure of distance (quantiles) instead of linear distance as the predictor variable. In columns (3) and (8) we run a fully saturated specification replacing group*distance with group-year fixed effects. In columns (4) and (9) we include proximity to Dallas and its second-order interactions along with proximity to Tuskegee to test whether the effects are driven by being near a city in the South. In columns (5) and (10) we assess robustness to weighting each cell by the size of the relevant older age population. The data source and sample were described in the notes to Table 2. The outcome variable is the log of chronic disease mortality. Controls in every specification include the interaction between proximity to Tuskegee and black or male and black-year or male-year as well as state-year fixed effects (except in columns (3) and (8)). Standard errors are clustered at the SEA level. ***p<0.01 ,** p<0.05 and *p<0.10, respectively.

11 Channels

Our measure of how much trust one has in their doctor comes from the General Social Survey (GSS). The GSS is a repeated cross section extending from 1972 to the present. The earliest year questions were asked about doctors was 1998, when several questions were included. In particular, participants were asked about whether “doctors judgment trusted" and whether "doctors deny me the treatment needed". Although we cannot perform a DDD analysis on such data (there is no “pre” period), we can ask whether individuals living closer to Macon County, Alabama were more affected by it. For this analysis, we ask whether black male survey respondents have systematically different perceptions about physicians relative to their white (or female) counterparts. The (cross-sectional) specification representing survey responses for individual i who lived in state s at age 16 and currently resides state c is given by:

$$Mistrust_MD_{igs} = \alpha + \beta(P_s * blackmale_g) + \theta blackmale_g + \pi_s + \gamma_s X_{is} + \epsilon_{igs} \tag{3}$$

where P_s is the proximity to Macon County and π_s is a current-state fixed effect so that β is the difference-in-differenecs estimate. The sample includes individuals at least 10 years of age at the time of the disclosure. X contains age, marital status and education category. In addition, we condition on an individual’s general level of mistrust in others to isolate the impact of medical mistrust.[67] We estimate the equation on a sample of all men, and then on an all black sample comparing again black men to black women. Standard errors are clustered at the level of treatment (state of residence). The results, contained in Table 7, demonstrate that black men, compared to white men and black women are 12.3 to 15.4 percentage points more likely to mistrust doctors and 10.1 to 32.1 percentage points more likely to be concerned with treatment denial, respectively. We find no corresponding effects for black women relative to white women or for white women relative to white men (see Appendix Table 9).

[67] This variable was constructed from a question on whether people can be trusted. Individuals who replied no or don’t know were coded as 0 and those who replied yes were coded as 1. A similar coding strategy was applied to the outcome variables, so that the coefficients can be easily interpreted as marginal effects. Estimation using an ordered logit produces similar results.

TABLE 7: EFFECT OF TUSKEGEE ON BELIEFS ABOUT MEDICAL CARE

Outcome	(1) *Medical Mistrust*	(2) *Deny Treatment*	(3) *Medical Mistrust*	(4) *Deny Treatment*
	PANEL A --- Male Sample		**PANEL B--- Black Sample**	
P_s*$blackmale_g$	0.1231**	0.1012**	0.1538**	0.3207***
	(0.0565)	(0.0435)	(0.0730)	(0.0675)
Fixed Effects	$State_{1998}$			
Other Controls	Age, Education and Marital Status Indicators, General Mistrust			
Observations	329	329	118	118
Adj R-squared	0.00397	0.0341	0.0208	0.125
No Clusters	36	36	28	28

Notes: OLS estimates of equation 3. The data are from the General Social Survey for the year 1998. The outcome variable for Columns (1) and (3) is whether the respondent disagrees with the statement that doctors can be trusted. The outcome variable in Columns (2) and (4) is whether the respondent believes they will be denied needed treatment by the medical profession. Panel A restricts the sample to black and white male demographic groups and Panel B restricts the sample to black male and female demographic groups. Controls in every specification include indicator variables for educational categories, age categories, marital status, state of current residence fixed effects, black or male and a general measure of mistrust. Standard errors are clustered at the state of residence level. ***p<0.01, **p<0.05, *p<0.10, respectively.

Pursuing further evidence on the channels though which our measured impacts are operating, we note that the model developed above predicted that the impact of Tuskegee would be more poignant for individuals who had limited experience with the healthcare sector prior to the disclosure. To isolate the role of experience as a mitigating factor on belief formation we examine the post-1972 differences between men who have served in the military versus those who have not *within race,* again as a function of proximity to Macon County, Alabama. In other words, we replace $blackmale_g$ in Equation 1 with an indicator variable for whether an individual was *never* drafted into the military, $nonvet_g$. We do not observe veteran status in the mortality data.

We estimate the model for a sample restricted only to black men and again for a sample restricted to white men. A marked geographic gradient in the veteran-non-veteran gap in primary care utilization in the years following the Tuskegee disclosure is apparent in Table 8. The results for black men for the

extensive margin of health seeking behavior (any outpatient visits) is given in Columns 1 and 3, and the intensive margin (number of outpatient visits) is given in Columns 2 and 4. The results are consistent with the predictions of the model and with the prior within-black results in Table 1, which compared non-veteran black men to black women. Black men with the least experience with the medical profession appear to have been more affected by the news of TSUS, in line with the predictions of the theoretical model of belief formation. The same pattern does not hold in Columns 3 and 4, both of which are estimated on an all-white sample.

TABLE 8: BASELINE RESULTS, UTILIZATION VETERANS

Outcome	(1) *Any Outpatient Visit*	(2) *Number Outpatient Visits*	(3) *Any Outpatient Visit*	(4) *Number Outpatient Visits*
	PANEL A - All Black Men		**PANEL B - All White Men**	
P_s*$post_t$*$nonvet_g$	-0.072***	-1.731***	0.194	0.011
	(0.014)	(0.396)	(0.119)	(0.007)
Observations	9,840	9,840	119,650	119,650
No Clusters	42	42	49	49
R-squared	0.065	0.058	0.019	0.023

Notes: OLS estimates of equation 1 testing differences between experienced (veteran) and inexperienced (non-veteran) males in response to the disclosure of the Tuskegee study. Data are from the harmonized version of the National Health Interview Survey (NHIS) available from IPUMS and merged with restricted identifiers for use in the Restricted Data Center (RDC). The sample includes males ages 45-74. The sample varies across columns and is given by the column heading, with columns (1) and (2) representing the black sample and columns (3) and (4) representing the white sample, which functions as a placebo test. In columns (1) and (3) the outcome is an indicator variable for any outpatient physician interaction in the past 12 months. In columns (2) and (4) the outcome variable is the number of outpatient physician interactions in the last 12 months. Controls in every specification include indicator variables for educational status, income category, age, marital status, telephone ownership as well as rural/urban status, the interaction between proximity to Tuskegee and black or male and black-year or male-year as well as state-year fixed effects. Standard errors are clustered at the state level. ***p<0.01 ,** p<0.05 and *p<0.10, respectively.

12 Conclusion

The Tuskegee Study was one of the most egregious examples of medical exploitation in US history. Our estimates indicate that the years following disclosure of the study's tactics brought significantly lower uti-

lization of both outpatient and inpatient medical care by older black men in closer geographic and cultural proximity to the study's subjects. This reduction in healthcare utilization paralleled a significant increase in the probability that a black man in closer geographic and cultural proximity to the study died before the age of 75. The data indicate no corresponding effects for younger black males or for white males or black women. These results are robust to accounting for a wide range of policies, economic forces, and individual characteristics thought to shape health behaviors. Our results imply that Tuskegee can account for half of the outstanding gap black-white male mortality in the immediate aftermath of the disclosure, or approximately 4.7 age-adjusted deaths per 1000 population out of a post-1972 gap of 9.3 Back-of-the-envelope calculations imply the Tuskegee and its revelation reduced black life expectancy at age 50 by up to 1 year.

Our findings underscore the importance of trust for economic relationships involving imperfect information. Typically the literature on trust has focused on trade settings (**?**); however, much of medical care depends on health providers and patients resolving information asymmetries. Trust, therefore, is a key component of this interaction. The theoretical framework in the paper suggests that positive interactions with medical providers in early adulthood can temper the formation of negative beliefs after the revelation of group-specific medical exploitation, a prediction that is also borne out in the paper's main results. Our findings also relate to the often observed low take-up for products with proven health benefits. Understanding the historical rationale for beliefs, even if currently maladaptive, might prove useful to formulating policy aimed at increasing such demand.