

NBER WORKING PAPER SERIES

TUSKEGEE AND THE HEALTH OF BLACK MEN

Marcella Alsan  
Marianne Wanamaker

Working Paper 22323  
<http://www.nber.org/papers/w22323>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 2016

For detailed feedback, we thank Nathan Nunn, Arun Chandrasekhar, Martha Bailey, Pascaline Dupas, William Collins and Matt Harris. We are also grateful to John Parman, Achyuta Adhvaryu, Rebecca Diamond, Claudia Goldin, Melanie Morten, Mark Duggan, Mark Cullen, Melissa Dell, Nancy Qian, Ran Abramitzky, Rema Hanna, Grant Miller and participants at NBER DAE, NBER Cohort Studies, University of Tennessee, Vanderbilt Health Policy, Carnegie Mellon and University of Pittsburgh Joint Seminar, University of Copenhagen, University of Pennsylvania Health Policy, ASSA 2016, PACDEV 2016, Berkeley Population Center, University of Chicago Harris School of Public Policy, Stanford Health Policy, University of California-Davis, University of Maryland Population Center, Stanford Social Science and History Workshop, University of South Carolina and CIREQ Montreal for constructive comments. We thank the CDC for providing access and 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. Address of corresponding author: 117 Encina Commons Room 186 Stanford, CA 94305 email: [malsan@stanford.edu](mailto:malsan@stanford.edu) The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Marcella Alsan and Marianne Wanamaker. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Tuskegee and the Health of Black Men  
Marcella Alsan and Marianne Wanamaker  
NBER Working Paper No. 22323  
June 2016  
JEL No. I14,O15

### ABSTRACT

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 both outpatient and inpatient physician interactions for older black men. Our estimates imply life expectancy at age 45 for black men fell by up to 1.4 years in response to the disclosure, accounting for approximately 35% of the 1980 life expectancy gap between black and white men.

Marcella Alsan  
Stanford Medical School  
Center for Health Policy/PCOR  
117 Encina Commons, Room 218  
Stanford, CA 94304  
and NBER  
malsan@stanford.edu

Marianne Wanamaker  
Department of Economics  
University of Tennessee  
524 Stokely Management Center  
Knoxville, TN 37996  
and NBER  
wanamaker@utk.edu

A data appendix is available at <http://www.nber.org/data-appendix/w22323>

# 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)

African-American men have the worst health outcomes of all major ethnic, racial, and demographic groups in the United States.<sup>1</sup> Although recent trends have shown signs of improvement, particularly at younger ages, the gradients for older men are still sobering: the expectation of life for black men at age 45 is three years less than their white male peers and five years less than for black women (Murphy *et al.*, 2013).<sup>2</sup> Compared to men from other backgrounds, black men have higher death rates from chronic conditions, including HIV/AIDS, heart disease, and cancer (e.g. lung, prostate, and colon) (Kaiser Family Foundation Fact Sheet, 2007). Many factors contribute to such disparities, including lower income and education, lack of health insurance, and higher rates of disengagement from the labor force. But socioeconomic status is not fully determinant of these gaps (Adler *et al.*, 1993; Cutler *et al.*, 2011), and a growing body of qualitative literature suggests that mistrust of healthcare institutions partially contributes to these inequities. Yet empirical evidence on the causal role of medical mistrust for racial health disparities remains thin.

In this study, we aim to fill this gap in the literature using the historical disclosure of an unethical and deadly experiment, the Tuskegee Study of Untreated Syphilis in the Negro Male (TSUS), to identify the relationship between medical mistrust and racial disparities in health-related behaviors and outcomes. For 40 years, between 1932 and 1972, the U.S. Public Health Service (PHS) followed hundreds of poor, black men in Tuskegee Alabama, the majority 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 (Brandt, 1978). Participants were subjected to blood draws, spinal taps, and, eventually, autopsies, by the study's primarily white medical staff. Survivors later reported that study doctors diagnosed them with "bad blood" for which they believed they were being

---

<sup>1</sup>For a comprehensive review of racial inequalities in U.S. medical care, see Institute of Medicine (2003).

<sup>2</sup>Currie and Schwandt (2016a,b) highlight reductions in age-specific mortality rates for black males between 1990 and 2010, particularly for young children and adults 20-49. Gains for children may be related to Medicaid expansions while the Ryan White CARE Act of 1990, which funds HIV/AIDS treatment, may have had an impact on the death rate of young adults. The life expectancy gap at age 45-46 between white and black males was 3.9 in 1980 and 3.3 in 2010 (National Center for Health Statistics, 1985; Arias, 2014).

treated.<sup>3</sup> 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 medical establishment were widespread.<sup>4,5</sup> By that point, the majority of the study's victims were deceased, many from syphilis-related causes.

In the years following 1972, journalists, social scientists, and medical researchers have repeatedly pointed to the Tuskegee experiment as a reason African-Americans remain wary of mainstream medicine. For example, as HIV disproportionately ravaged black communities in the U.S., a number of observers conjectured that the legacy of Tuskegee had hampered public health education efforts in the black community, contributing to the epidemic's spread (Gaston and Alleyne-Green, 2013). A recent resurgence of tuberculosis in Marion County, Alabama is similarly attributed.<sup>6</sup> Qualitative research aimed at understanding the reluctance of black men, in particular, to participate in medical research or engage in preventive care, often finds unprompted reference to Tuskegee (Corbie-Smith et al., 1999).<sup>7,8</sup> In sum, the Tuskegee study is an often-cited contributor to delays and avoidance in care seeking, wariness of public health campaigns, low participation in clinical trials, and overall worse health indicators among black men.<sup>9</sup>

In this paper, we empirically investigate whether the study's disclosure contributed to racial disparities in health and healthcare utilization in the years following 1972. To do so, we rely on a variety of survey and administrative data, including measures of trust in doctors from the General Social Survey (GSS) (Smith et al., 2015), health seeking behavior reported in the National Health Interview Survey (NHIS) (Minnesota Population Center, 2015), and detailed annual mortality data available by race, age group, gender and cause

---

<sup>3</sup>As an example of this outright duplicity, Brandt (1978) quotes a letter subjects received exhorting them to participate in (the often painful) spinal tap procedure: "Some time ago you were given a thorough examination and since that time we hope you have gotten a great deal of treatment for bad blood. You will now be given your last chance to get a second examination."

<sup>4</sup>Contemporary coverage included *The New York Times* and *The Chicago Tribune*, as well as newspapers and magazines with a predominantly black readership.

<sup>5</sup>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," as quoted in Reverby (2000, p.xii).

<sup>6</sup>See coverage of a severe 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 (Blinder, 2016).

<sup>7</sup>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."

<sup>8</sup>Another example comes from an October 2015 episode of "Black-ish", an ABC sitcom which centers on an African-American family. The episode begins with a discussion of the Tuskegee experiment and the effect it has on the health-seeking behavior of the protagonist's aging father.

<sup>9</sup>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 (Paratz and Katz, 2011).

from the Centers for Disease Control and Prevention (CDC) (Centers for Disease Control and Prevention, 2014). We focus our attention on these measures for older (45-74) individuals since the mortality and health-seeking behavior of younger individuals is generally driven by acute conditions such as childbirth or trauma in which the needs for care are urgent and the benefits immediate.<sup>10</sup> 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 conjecture that the disclosure had a stronger impact on the behavior of individuals who more readily identified with the study's mostly illiterate black males subjects, and we test this hypothesis by comparing the behavior of black men to both black females and white males. In addition to race, gender and age, geographic distance from Tuskegee, Alabama is used as a fourth dimension of proximity to the study subjects.<sup>11</sup> As a robustness test, we substitute black migrant inflows from the state of Alabama for geographic distance. Our focus on cultural, geographic and demographic ties to the study's subjects as a measure of treatment intensity is supported both by survey data specific to TSUS highlighting racial and gender differences in related mistrust and by research on the neurological basis of the empathy response.<sup>12</sup> In addition, we propose that individuals possessing less prior experience with the healthcare system (e.g. non-veteran males) would have exhibited a stronger response to the study's disclosure. In the Appendix, we ground our hypotheses in a Bayesian, multi-period model of belief formation regarding physician trustworthiness across heterogeneous agents who differ in their cultural proximity to the study's subjects and in their prior healthcare experience.<sup>13</sup>

These hypotheses are tested empirically using a Difference-in-Difference-in-Differences (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 (Aizer and Stroud, 2010; Alsan and Cutler, 2013; Cutler and Lleras-Muney, 2012; Cutler et al., 2014), income (Deaton, 2002), marital status (Robles and Kiecolt-Glaser, 2003; Holt-Lunstad et al., 2008), and urbanization. The

---

<sup>10</sup>Injuries and violence are the leading cause of death for children and young adults between ages 1 and 44 (CDC, 2010). For the GSS, we use those who were at least 10 years old at the time of the study disclosure.

<sup>11</sup>Our study distinguishes between the non-veteran and veteran population of men. The baseline utilization samples are restricted to non-veteran males, but we do not observe veteran status in the mortality files.

<sup>12</sup>Neurocognitive research has elucidated the physiological basis for empathetic responses (via mirror neurons) (Singer et al., 2006) and documented that neural activity is typically heightened for a member of one's own group (Gutsell and Inzlicht, 2010).

<sup>13</sup>We consider Bayesian updating as the benchmark, although behavioral models could deliver similar results (e.g. Becker and Rubinstein (2011)).

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 demographic 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.

We estimate that a one-standard deviation increase in geographic proximity to Tuskegee, Alabama reduced utilization of routine care (outpatient physician contacts) among older black males by 1.29 interactions per year when compared to their white peers and 0.97 interactions per year relative to black women, a 20-30% reduction in utilization relative to the pre-disclosure mean value for black men. Similar results obtain for alternative measures of geographic and cultural proximity. Our utilization results are driven by the behavior of black men with lower levels of education and income, a population in closer socioeconomic proximity to the victims of TSUS, in line with our hypotheses regarding connectedness to the study population. We also find a reduction in the probability that black men were admitted to the hospital, though their length of stay was longer, consistent with more advanced disease upon presentation. In the mortality data, we find that a one standard deviation increase in geographic proximity to Tuskegee was associated with post-disclosure age-adjusted mortality increases among older black men of 3.5 log points compared to white men and 6.9 log points compared to black women.

Our results indicate that the Tuskegee study and its disclosure were substantial determinants of a life expectancy gap between older black and white men that widened between 1970 and 1980. Our estimates imply a reduced life expectancy attributable to the study's legacy of up to 1.4 years for older black men, approximately 35% of the racial life expectancy gap in 1980.<sup>14</sup>

An alternative explanation for our findings is that access to healthcare was deteriorating for blacks

---

<sup>14</sup>See Appendix Section 5 for details of the life table calculation. 1979-1981 life tables from the CDC show white male life expectancy at age 45 of 29.55 and black male life expectancy of 25.61. Our estimated reduction of 1.4 years in black life expectancy at age 45 is 35% of the difference (National Center for Health Statistics, 1985).

relative to whites in this period, particularly in the South, and 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 for black Americans improved in the years following the Civil Rights Act, 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 (Goodman-Bacon, 2015), the desegregation of hospitals (Almond et al., 2006; Zheng and Zhou, 2008), and political enfranchisement, which led to an increase in the flow of public goods to black communities (Cascio and Washington, 2014).<sup>15</sup> Our estimates imply that the Tuskegee revelation stalled the overall pattern of convergence for older black men, producing a drag on healthcare utilization for this specific demographic group and contributing to an abrupt divergence in their mortality pattern. 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, and our findings on health-seeking behavior and mortality are robust to estimation on a within-South sample. 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.<sup>16</sup> 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 and incarceration rates, does not affect our estimates.<sup>17,18</sup>

---

<sup>15</sup>For an historical account of U.S. black-white health disparities, see Boustan and Margo (2014). Another possible confounder is the Great Migration, which, for this group of men, reached its nadir in the 1940s (when men of age 45-74 in 1972 were of age 13-42). The Great Migration may have affected the geographic gradient in later-life mortality for black men if (1) there was selection into migration based on health or (2) men who spent their early life in Northern cities experienced more positive health outcomes later in life. Neither of these outcomes are likely to be the case. Eriksson and Niemesh (2016) show that there is little evidence of health-based selection into migration and that Northern cities were health depleting, rather than augmenting, for young children. See also Black et al. (2015).

<sup>16</sup>See Appendix Tables 5 and 8. In Appendix Table 8, both hospital beds and Medicaid expenditures show a continuous increase over this time period, and more so in locations closer to Tuskegee. But there is no trend break in 1972 that would have induced utilization and mortality changes in that year.

<sup>17</sup>These two robustness checks (in Appendix Table 5) 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.

<sup>18</sup>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 back to the early 1970s. We therefore use race- and gender-specific unemployment from the 1970 census by county interacted with

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 (Baicker et al., 2004, 2005), 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 additional placebo tests, we show that the baseline coefficients, based on geographic proximity to Tuskegee, Alabama, are larger (in absolute value) than 92% to 98% of placebo tests substituting proximity to all other SEAs or states in the dataset. Thus the effects we measure are specific to geographic proximity to Tuskegee and not to a post-1972 condition affecting the South overall.

Finally, to evaluate 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. When we interact race and gender indicators with a measure of the distance of an individual from Macon County, Alabama, we find the same geographic gradient apparent in our baseline utilization and mortality results. This gradient in mistrust by geographic proximity to Tuskegee is present even when conditioning on the overall level of mistrust revealed in the respondent's answers to other survey questions.<sup>19</sup>

This paper builds on and contributes to several literatures in economics. First, the effects we describe shed light on factors influencing healthcare demand and medical outcomes among minorities in the United States and deepen our understanding of how this watershed event affected the relationship between black Americans and the U.S. healthcare system. Second, and more broadly, our study is motivated by the theoretical contributions of Guiso et al. (2008) regarding the transmission and updating of beliefs and empirical work by Nunn and Wantchekon (2011) 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 (Knack and Keefer, 1997; Fafchamps, 2006). Finally, the research sheds light on questions in development economics regarding low demand for products that have been shown to improve health (Dupas, 2011; Chapter 3 of Banerjee and Duflo, 2012). The findings presented herein suggest that historical exploitation and its enduring impact on beliefs may explain some of the uptake paradox.<sup>20</sup>

---

a time-trend to control for macroeconomic conditions for each group. See Appendix Table 5.

<sup>19</sup>Although attrition is an issue in such a sample, given the findings on mortality described above, these estimates likely represent a lower bound if more distrustful individuals were more likely to suffer early mortality.

<sup>20</sup>This is also provided as a potential explanation for aforementioned low demand in Banerjee and Duflo (2012, p.58): "Faith, or to use the more secular equivalents, a combination of beliefs and theories, is clearly a very important part of how we all navigate the health system."



## 2 Context

The following section places the study in historical and medical context by providing background on the pathology of syphilis, the logistics of the TSUS experiment, and health patterns across demographic groups in the years leading up to the disclosure. In addition, we describe some of the qualitative research performed by clinical psychologists on medical mistrust and preventive care delays among African-American men.

### 2.1 Background on the Tuskegee Study

The Tuskegee Study was designed to trace the course of untreated syphilis. The organism that causes the disease is related to that causing Lyme disease, and both bacteria 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 teeming 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 often leading to hemorrhage. Neurosyphilis (an attack on the nervous systems) presents in late-stage syphilis with paresis, gait disturbance, blindness, and dementia (Mandell et al., 2009).<sup>21</sup>

According to Jones (1992), 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."<sup>22</sup> 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 (~2/3 of whom had 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. The study was halted that same year, but few of the original test subjects were alive and dozens of their wives and children had been infected (Heintzelman, 2003). Medical historian Allan Brandt (1978)

---

<sup>21</sup> Syphilis can also be transmitted mother to child and cause severe congenital problems including stillbirth.

<sup>22</sup> According to Thomas and Quinn (1991, 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."

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 Tuskegee Syphilis Study Legacy Committee (1996) 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-Americans in clinical trials, organ donation efforts, and routine preventive care."<sup>23</sup>

## 2.2 African-American Men and Medical Mistrust

As discussed in the introduction, black men consistently experience a deficit in many health indicators, including life expectancy, survival rates within disease category, and health-promoting behavior. Black men are far less likely to obtain preventive care than white men and are therefore less likely to know their cholesterol levels, engage in blood pressure monitoring, or otherwise benefit from early detection of chronic conditions (Sandman et al., 2000; Viera et al., 2006).<sup>24</sup>

Although reasons for these racial disparities in preventive care utilization are multifactorial, the legacy of medical exploitation (represented in part by the Tuskegee study) has led to research on whether medical mistrust is particularly heightened among black men (Carlson and Chamberlain, 2004). In a large number of recent studies, black men do, indeed, exhibit higher levels of medical mistrust that are correlated with reduced probabilities of preventive and early-stage disease care.<sup>25</sup> For example, Hammond et al. (2010) recruited 600 African-American men from barber shops across the country for the purpose of assessing predictors of self-reported delays in medical care. The most important measured factor was medical mistrust, which was associated with doubled or tripled odds ratios of delays in routine checkups and in blood pressure and cholesterol screenings.<sup>26</sup> This finding echoes others in the literature related to primary care underutilization (Thom et al., 2006; Hertz et al., 2005; Nelson et al., 2002; Keating et al., 2004; Halbert et al., 2006).

---

<sup>23</sup>This lack of minority representation in clinical trials may have important spillover effects including on the speed and direction of innovation for ailments that heavily afflict their communities. See Hamilton et al. (2016) for a structural estimation of these losses.

<sup>24</sup>Indeed, the life expectancy gap between black and white men can be substantially reduced by controlling for stage at diagnosis. See American Cancer Society (2008, 2009) and Silber et al. (2014).

<sup>25</sup>See Boulware et al. (2003); Brandon et al. (2005); Hood et al. (2012) and Wiltshire et al. (2011).

<sup>26</sup>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)."

Medical mistrust among black men is also correlated with poor compliance with treatment plans, conditional on diagnosis. In a 2016 study of low adherence to antiretroviral treatment among HIV-positive black men, researchers documented that 63% of the study's subjects held a "conspiracy belief", for example that the federal government was responsible for HIV's introduction into the black population. These beliefs, in turn, were associated with a reduced likelihood of adhering to a physician-prescribed treatment regimen (Bogart et al., 2016). In addition to HIV, medical mistrust has also been associated with poor treatment adherence among cancer, cardiology, and hypertension patients (Halbert et al., 2009; Kayaniyil et al., 2009; LaVeist et al., 2000; Do et al., 2010; Thomas, 2013).

Although these studies highlight the heightened mistrust among black men and its correlation with healthcare utilization, to our knowledge, we are the first to relate racial disparities in healthcare utilization and health outcomes to medical mistrust in a quasi-experimental framework.<sup>27</sup>

### 2.3 Patterns of Healthcare Utilization and Mortality

The 1960s marked an era of rapid convergence in healthcare access and utilization for black Americans relative to white Americans. 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 President 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 (Reynolds, 1997; Zheng and Zhou, 2008). These decades also witnessed expansions in 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.<sup>28</sup> Using data from the NHIS, among

---

<sup>27</sup>The most recent (and perhaps only) attempt to calculate the observational correlation between Tuskegee knowledge and medical mistrust (Brandon et al., 2005) has been roundly criticized for methodological failures (White, 2005).

<sup>28</sup>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 (Byrd and Clayton, 2002, p. 391).

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 60% to 69%. The gains for black men were similarly impressive; the same metric rose from 54% to 64%. These increases were twice as large as the increases experienced by white males and females over the same time period.<sup>29</sup>

The convergence of black and white healthcare utilization rates in the years prior to the Tuskegee disclosure 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 convergence at the beginning of the period followed by 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 differences between black and white women. The mortality pattern for most other ages (contained in Appendix Figure 2) mimics that for infants and children.<sup>30</sup> 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 and physician involvement, including timely diagnosis and management of chronic illness.<sup>31</sup> To facilitate comparisons across age groups, the mortality rates represented by Figure 1 are those from all causes of death, although specific cause-of-death is available in the underlying data and utilized later in the paper. These stark mortality patterns for black men relative to both black women and white men motivate the paper's focus on health behaviors and outcomes of black men in particular. We next turn to outlining a conceptual framework for interpreting the differential behavior of black men.

---

<sup>29</sup>The level of utilization for black (and white) men is lower than for women in every period. Note that we do not have a consistent, continuous series of utilization data in the pre-1969 period.

<sup>30</sup>The Appendix also contains figures for Southerners alone (Appendix Figure 1) and all other age groups (Appendix Figure 2). For younger black men, an increase in relative mortality is apparent beginning in the 1980s. This is likely due to the evolving HIV/AIDS crisis in this community. Appendix Section 6.1 contains more detail.

<sup>31</sup>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.

### 3 Conceptual Framework

To better conceptualize how the Tuskegee disclosure may have differentially affected population subgroups, the Appendix of this paper contains a fully-specified model of how prior beliefs, experience with the medical profession, and proximity to the Tuskegee study subjects interact to determine health-seeking behavior. The model's assumptions are that agents are homogenous in their prior beliefs over the trustworthiness of doctors, but heterogeneous along two dimensions: their experience with the medical profession in early life and their cultural, geographic, and demographic proximity to those exploited by the study. In particular, we assume that individuals who engage with the medical system as young adults, either because of military service for men or obstetric care needs for women, receive information that the medical system, on average, is "good", i.e. contains predominantly trustworthy doctors. We also assume that, upon TSUS disclosure, individuals who are "near" in proximity to the study subjects receive and incorporate information indicating that the medical system is, on average "bad" for treating people of their type, i.e. contains predominantly untrustworthy doctors. Thus, experience and proximity determine an agent's posterior belief regarding the trustworthiness of the medical profession.

This framework generates testable predictions for the health-seeking behavior of older adults that inform our empirical specifications. First, the model predicts that, conditional on proximity, agents who are experienced with the medical profession will be less sensitive to the news of Tuskegee, thereby having higher posterior beliefs over the trustworthiness of doctors and higher post-disclosure rates of care-seeking. The resulting prediction is that non-veteran black men would have been more affected by the news of TSUS than veteran black men and more affected than black women as the latter groups would have routinely encountered the medical system as young adults.<sup>32</sup>

Second, our model indicates that, conditional on medical experience in young adulthood, older individuals in closer proximity to the study subjects will harbor greater medical mistrust following the disclosure. Other models of trust formation and social identification (Dixit, 2003; Shayo and Zussman, 2011; Tabellini, 2008) and a vast psychology literature (Gutsell and Inzlicht, 2010; Singer et al., 2006) indicate that indi-

---

<sup>32</sup>By 1960, 97% of all births were in hospitals (Feldhusen, 2000) and less than 10% of U.S. women in 1970 were childless by their early forties (Livingston and Cohn, 2010). We are unable to observe veteran status in the mortality data or in the GSS survey data.

viduals tend to be more affected by news if they can identify with the subject.<sup>33,34,35</sup> Thus, if we assume that race, age, gender and geography together define connectedness to the study subjects, we derive the prediction that black men should have been more affected by the disclosure than either white males or black females and this difference varies over distance.<sup>36</sup> The empirical approach uses a DDD framework to test these predictions, comparing black men before and after the Tuskegee disclosure (first difference), relative to a peer group of white men or black women (second difference) across an additional measure of cultural proximity to the study subjects (third difference). For this third difference, we primarily use geographic proximity to Tuskegee, Alabama, but also incorporate the local density of black migrants from Alabama.<sup>37</sup> These measures are described in more detail below.

## 4 Data

### 4.1 Health-Seeking Behavior and Health Outcomes

The paper's empirical approach focuses on changes in health-seeking behavior and health outcomes as a result of the Tuskegee disclosure. We measure health-seeking behavior using data on healthcare utilization from the NHIS. The survey is a repeated cross-sectional sample of Americans which began in 1963. The survey is taken by household visit, and we use the individual-level respondent data in our analysis.<sup>38</sup> We

---

<sup>33</sup>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.

<sup>34</sup>As evidence of this, although the Rodney King beating, which took place in March 1991, was widely publicized in 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 18% among black men, 7% among black women, and declined among whites (ABC News/The Washington Post, 1992; NBC, 1989).

<sup>35</sup>Shayo and Zussman (2011) in their study of how terrorism affects judicial bias in Israel, endogenize group identification and define each group as a vector of attributes and describe how perceived distance, the distance between oneself and a given group, may differ depending on the salience (e.g. attention weight) associated with an attribute.

<sup>36</sup>Other results from empirical social psychology also support the notion that black men would have been more greatly affected by the news of Tuskegee. In recent work on the formation of trust, psychologists have noted that men tend to form collective trust while women form relational trust. Women's relational trust, based on particular relationships formed over past experiences, would have prevented news about a collective, but unknown, group of people from affecting their trust of the medical system. See Maddux and Brewer (2005).

<sup>37</sup>Data from the Survey of Black Americans support 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 (Jackson et al., 1979). We thank Trevon Logan for pointing out this source.

<sup>38</sup>The U.S. Public Health Service of the United States — the same entity that carried out the Tuskegee study — sponsored the survey, but the interviews were conducted by the U.S. Census Bureau. It is reasonable to suppose we are estimating a lower bound since mistrustful individuals might have refused to answer the survey.

rely on the harmonization provided by the Integrated Health Interview Series (IHIS), which is based on the NHIS public-use data. Harmonized responses regarding the utilization of medical services are available from 1969.<sup>39</sup> The public-use samples are stripped of geographic identifiers necessitating the use of restricted access for these data from the National Center for Health Statistics (NCHS) (Minnesota Population Center, 2015).<sup>40</sup>

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.<sup>41</sup> The data also contain detailed demographic information, including age, family structure, income, education, and veteran status.<sup>42</sup> As discussed previously, we focus on men aged 45 to 74 in order to capture the effects on men (and women) in the age range most likely to benefit from preventive and early-stage disease care.

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.<sup>43</sup> In doing so, we follow a large literature that uses mortality to assess the efficacy of health interventions.<sup>44</sup> 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 (Centers for Disease Control and Prevention, 2014). 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 (Curtin and Klein, 1995, p.3), and the NCHS suggests aggregating over space or time if small numbers are encountered. For this reason, we aggregate our annual

---

<sup>39</sup>The main question we use on outpatient interactions is derived from a question that was unchanged between 1969 to 1981. In 1978 the race categorization changed, with multiple categories included, making 1977 a natural stopping point. But we have also extended our analysis to 1981 (the year the question on doctor visits changes and harmonization is no longer possible) and obtain similar (though not disclosed) results.

<sup>40</sup>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.

<sup>41</sup>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. Individuals that do not know whether they own a telephone were excluded from the analysis.

<sup>42</sup>The NHIS data do not contain consistent measures of individual health insurance coverage over our time period; however, in our mortality regressions we control for Medicare and Medicaid trends and do not obtain statistically different results. See Appendix Table 5.

<sup>43</sup>Self-reported health status and other plausible metrics of morbidity in the NHIS are unavailable prior to 1973.

<sup>44</sup>Recent examples include Almond et al. (2006); Alsan and Goldin (2015); Bhalotra and Venkataramani (2012); Bailey and Goodman-Bacon (2015) and Goodman-Bacon (2015).

county-level values to the state economic area (SEA) level and measure biennial mortality rates.<sup>45</sup>

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, for individuals aged 45 to 74. 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 plausibly prevent or delay mortality.<sup>46</sup> We also use all-cause age-adjusted mortality rates (not restricted to chronic diseases) to account for differences in age structures across groups.<sup>47</sup> 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.<sup>48</sup> 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 their place of birth. To the extent our current proximity measures capture information spread via social networks, 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. We will, however, make use of group-level migration patterns, as discussed in the next section.

## 4.2 Proximity Measures

In addition to race, gender and age, our empirical results rely on two different measures for proximity of individuals to study subjects and, thus, their subsequent belief updating process following the disclosure. Our primary measure is linear proximity to Macon County, Alabama (Figure 2 Panel A), supported by the linear relationship between post-1972 mortality changes and distance, as discussed in Section 5 below.

---

<sup>45</sup>Results using county-level and annual mortality rates are in Appendix Table 6. Using data from the NCHS' Multiple Cause of Death file for earlier 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.'"

<sup>46</sup>Some examples include 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 4 for a discussion of medical and public health innovations over this time period.

<sup>47</sup>We follow the demographic literature and use the standard 1940 population for reference. See the Data Appendix for further details.

<sup>48</sup>After 1987, the paper's conclusions are likely to be compromised by both selective migration and by the evolving HIV/AIDS epidemic. But in the 1981-1987 period, HIV/AIDS was a nascent health threat, particularly among older black men. In this period, there were roughly 9,000 AIDS-infected men older than 45, of which perhaps 25% were black. See CDC (2001). Young people were at higher risk of acquiring HIV, and by 1994 AIDS had become the leading cause of death among black males aged 25-44. This increased mortality among younger black males in the late 1980s is apparent in Appendix Figure 2.



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 we note that geography may capture information spread through formal and informal networks or cultural similarity as discussed above.<sup>49</sup> This proximity may also reflect the information content of the signal if, for example, individuals at greater geographic distances do not believe that the Tuskegee experiment disclosure provided 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 (Figure 2, Panel B) as a robustness check.<sup>50</sup> (We drop Alabama observations from the sample in this analysis.) 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 be more culturally proximate to the event. 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.<sup>51</sup> 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. The 1940 Alabama migration rates are additionally relevant as migration patterns are roughly stable over time, making the 1940 patterns a good proxy for patterns in the years subsequent and, therefore, for the sources of new information flows in each SEA.<sup>52</sup>

---

<sup>49</sup>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 Alabama migrants measure.

<sup>50</sup>We normalize by black migrants for two reasons. First, doing so gives a measure of 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.

<sup>51</sup>Using later census years would preclude us from using detailed geographic information.

<sup>52</sup>We also considered whether formal information networks, such as media exposure, could serve as an alternative treatment, but prefer geographic 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. (Newspaper circulation figures that are available are rarely audited, in contrast to the more mainstream local daily newspapers from Gentzkow et al. (2014).) Our results are robust to including controls for the prevalence of black newspapers. (See Appendix Table 5.) Appendix Section 3 contains more discussion on media coverage of the Tuskegee study. 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, suggests 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

## 5 Empirical Strategy

### 5.1 Motivating Difference-in-Differences Estimates

We motivate the paper's main empirical specifications with simple difference-in-differences (DD) regressions comparing the post-1972 mortality rates of older black men (ages 45-74) relative to their white or female peers. Recall from Figure 1 that black male mortality was converging relative to white men from 1968 through the early 1970s before diverging thereafter. We therefore anticipate that the DD results below will highlight these trends.

The coefficient  $\theta_1$  in the equation below represents the post-1972 "treatment" effect of being a black male:

$$\ln(Y_{gat}) = \alpha + \theta_1(\text{post}_t * \text{blackmale}_g) + \theta_2(\text{blackmale}_g) + \pi_t + \varepsilon_{gat}. \quad (1)$$

Mortality rates for older adults (ages 45-74) are measured at the level of each SEA (denoted  $a$ ) for each of four demographic groups (denoted  $g$ ): black males, black females, white males, white females. The variable  $\text{blackmale}_g$  is equal to one for each SEA-year observation for black males, and zero otherwise. We use chronic mortality data for this exercise although age-adjusted mortality results are similar.

We first estimate equation 1 on an all-male sample, comparing the experiences of black men to those of white men before and after 1972. If mortality rates for black men had continued to converge in the post-1972 period, we would measure a value of  $\theta_1$  less than zero. Alternatively, stalled convergence would generate a value of  $\theta_1$  greater than or equal to zero. As expected given the results in Figure 1,  $\theta_1$  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  $\theta_1$  and the associated 95% confidence interval are given in the heading for Figure 3, Panel A.

This *average* treatment effect of being a black male after 1972 masks significant heterogeneity. We find a strong geographic gradient in the value of  $\theta_1$  along a measure of distance from Macon County, Alabama, the county that houses Tuskegee. When we estimate a  $\theta_1$  coefficient for every SEA in the sample, utilizing multiple pre- and post- observations within each SEA for identification, the results indicate a strong negative correlation between distance from Macon County and the value of these coefficients. In Figure 3, Panel A, we plot the value of these coefficients averaged across SEAs in ventile distance bins of approximately 180 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 Ayer (1971), as well as the distribution of televisions, as control rather than treatment variables.

kilometers, as well as a linear line of best fit for the underlying disaggregated data (in blue).<sup>53</sup> The size of the data bubbles are indicative of the number of older black men living in each group of SEAs. Although the average difference in black male chronic mortality relative to white male mortality increases by 6.2 log points following 1972, 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 a geographic gradient in the post-1972 divergence of black male mortality, we estimate  $\theta_1$  for an empirical specification restricted to black male and black female mortality observations. As indicated in the figure header for Panel B of Figure 3, 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 again upwards of 20 log points in locations closest to Tuskegee and declines linearly in distance from Tuskegee.<sup>54</sup>

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<sub>g</sub>* with *blackfemale<sub>g</sub>* and restricting the sample to be all female. 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  $\theta_1$  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. A similar exercise in Panel D examines the experience of white men with respect to white women. 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. 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

---

<sup>53</sup>For graphical exposition, we averaged coefficients across SEAs in 20 buckets representing equal distance intervals. Because fewer SEAs in the North and West housed black residents, coefficients in buckets further than 1500 kilometers from Tuskegee (10 buckets total) are quite noisy and were averaged across distance buckets twice as wide. This results in 10 buckets for distances 1500 kilometers and less and 5 buckets for distances greater than 1500 kilometers.

<sup>54</sup>Using the same estimating equation as above, utilization data from the NHIS show similar patterns for an extensive measure (probability of seeing a doctor in the past 12 months) and an intensive measure (the number of physician interactions in the last 12 months) of healthcare utilization. For an all-black analytical sample, the overall DD estimate for  $\theta_1$  is negative and insignificant. The estimated coefficients within the South region are more negative and statistically significant. For the within-male sample, the pattern is similar. The estimated value of  $\theta_1$  is positive for both utilization outcomes and significant for the intensive margin, but for survey respondents within the South region, the  $\theta_1$  coefficient is negative, indicating lower utilization in locations closer to Tuskegee. These results can be gleaned from publicly available data, where region of residence is observable. A geographic gradient figure (relying on state-level data) similar to Figure 2 cannot be produced for utilization outcomes given disclosure concerns. But the DDD estimates presented later in the paper capture the geographic gradient described here.

kilometers from Alabama; the estimated values of  $\theta_1$  within the South region 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, instituting a false definition of *post* within the set of observations prior to 1972. We plot estimates of  $\theta_1$  for the same set of distance buckets in panels E and F of Figure 3, defining *post* to be one for all observations after the mid-point of the pre period (1970). These falsification tests indicate no geographic gradient in the estimates of  $\theta_1$ , 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 highly statistically significant slope coefficient in panels A and B. The geographic gradients documented in Figure 3, both across all regions and within the South, inform the empirical specifications to follow.

## 5.2 Baseline Estimating Equations

We employ a Difference-in-Difference-in-Differences (DDD) estimator with differences taken over time, race or gender, and proximity to Macon County, Alabama. We denote this proximity measure  $P_k$  where  $k \in s, a$  represents measurement at the state or SEA level; respectively. The relevant estimating equation for healthcare utilization for individual  $i$  of demographic group  $g$  measured in state  $s$  at time  $t$  ( $Y_{igst}$ ) is:

$$Y_{igst} = \alpha + \beta(P_s * post_t * blackmale_g) + \mu(P_s * blackmale_g) + \theta_{gt} + \gamma_{st} + X_{igst}\delta + \epsilon_{igst}, \quad (2)$$

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

$$Y_{gat} = \alpha + \beta(P_a * post_t * blackmale_g) + \mu(P_a * blackmale_g) + \theta_{gt} + \gamma_{at} + \epsilon_{gat}. \quad (3)$$

In both cases,  $\theta_{gt}$  represents group-year fixed effects and  $\gamma_{st}$  and  $\gamma_{at}$  represent state-year and SEA-year fixed effects, respectively. In utilization 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,  $\beta$  measures the post-1972 geographic gradient in the cost of being a member of group  $g$ , controlling for any geographic gradient in the cost of being a member of group  $g$  in the years prior to 1972, as well as for any geography-specific time effects and for any group-specific time

effects.<sup>55</sup>

As for the DD estimates above, we estimate Equations 2 and 3 separately for an all-male sample (comparing black men to white men) and for an all-black sample (comparing black men to black women). We limit our analysis on health to a window around the disclosure of 1969-1977 for utilization measures and 1968-1987 for mortality.<sup>56</sup> We do not attempt to model the carry-on effects of TSUS to future cohorts of black men given that the identifying assumption behind the DDD approach becomes more tenuous as one moves out in time.<sup>57</sup>

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 in the years prior for each of the study outcomes.<sup>58</sup> 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 black population-weighted state centroids which are used to construct proximity as above.<sup>59</sup>

Our baseline mortality results contain one final sample adjustment: 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 SEA is unavailable. Practically, this implies eliminating all-white SEAs from the analytic sample as the primary cause of missing observations is a 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).

---

<sup>55</sup>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.

<sup>56</sup>1968 is the first year mortality data are recorded and NHIS data do not include consistent, harmonized utilization measures prior to 1969. The longer window for mortality accounts for the lag between health behaviors and death.

<sup>57</sup>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.

<sup>58</sup>Specifically, we use  $\max(\text{distance}) - (\text{distance}_i)$  so that locations furthest from Tuskegee receive a proximity value of 0.

<sup>59</sup>These centroids represent the average latitude and longitude of black individuals in each state based on the black population of counties (Haines, 2010).

### 5.3 Identification

In order for the resulting estimates to reflect the effect of the Tuskegee disclosure, there must be no other systematic shocks to black men that affected healthcare utilization and mortality correlated with proximity to Macon County, but not due to the timing of the study's disclosure. As discussed in Section 2.3 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 men differed in this period, any time-invariant geographic differences in these factors are absorbed by the  $P * blackmale$  interaction term in the all-male sample estimates. Additionally, any geography-invariant time effects for black men are absorbed by race-year or gender-year fixed effects in the all-male and all-black estimates, respectively. Location-year effects net out any time-varying health advantages or disadvantages associated with geography that affected all black Americans or all males, such as the rollout of Medicare and Medicaid and their various expansions. These location-year effects reduce the scope for unobserved factors to those affecting only black men, but not black women or white men, in closer proximity to Macon County, Alabama. Still, we note that there is no major expansion or contraction of public insurance in or around 1972 that would serve to confound interpretation of our results. Further, Goodman-Bacon (2015) estimates minority adult eligibility for Medicaid at 4.5%, a fraction unlikely to drive the results below. 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.

## 6 Results

### 6.1 Event Study Estimates

We first present the results of event study specifications for Equations 2 and 3 in which we generate a coefficient on  $P * blackmale$  as well as a coefficient on  $P * blackmale$  interacted with each year (except

1972). Specifically, we estimate:

$$Y_{igst} = \alpha + \sum_{n \neq 1972} \beta^n (P_s * I_t^n * blackmale_g) + \mu (P_s * blackmale_g) + \theta_{gt} + \gamma_{st} + X_{igst} \delta + \epsilon_{igst}, \quad (4)$$

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 log chronic disease mortality. (Recall that the mortality data have been grouped to measure biennial averages.) Just as in Figure 4, the plotted  $\beta$  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  $\beta$  coefficients for years after 1973. The post-disclosure biennial estimates are frequently statistically significant.

## 6.2 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 2 and 3.

Table 1 reports the coefficient on  $P_s * post_t * blackmale_g$  in equation 2 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 outpatient physician interactions they reported in the last 12 months ("Number Outpatient Visits"), whether an individual reported

having been admitted to the hospital in the past 12 months ("Any Hospital Admission") and the number of nights spent in the hospital in the past 12 months ("Number Nights in Hospital").

Estimated coefficients in Column 1 indicate that black men experienced sharp declines in the probability of visiting a doctor in the years following 1972 as a function of their proximity to the Tuskegee study's location. For the all-male sample, each thousand kilometer increase in proximity brought a reduction in outpatient visit probabilities of 4.0 percentage points after 1972. Accordingly, a one standard deviation increase in proximity to Macon County, Alabama was associated with a 2.7 percentage point decline in the probability of having an outpatient doctor's visit.<sup>60</sup> Estimates derived from an all-black sample (Column 5) are similar. A sizable post-1972 effect is also discernible on the intensive margin, the number of physician visits reported over the past 12 months. We estimate a reduction of 1.4 (Column 6) to 1.9 visits (Column 2) per thousand kilometers of proximity to Macon County, depending on the estimating sample. These estimates indicate that a standard deviation increase in proximity to the study's home county was associated with reduced outpatient interactions of 0.97 to 1.29 visits per year, approximately 20 – 30% of the pre-disclosure black male sample mean.

Turning to hospitalization measures, the estimated  $\beta$  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 thousand kilometers. In Column 7, using an all-black sample, the reduction is  $-1.7$  percentage points per thousand kilometers. These estimates translate to  $-1.7$  to  $-1.2$  percentage points per standard deviation increase in proximity, depending on the comparison group. These results indicate a post-Tuskegee disclosure reduction in acute medical care probabilities of 12 – 17%. Although hospital admission rates went down, black men appeared to have more advanced illness on presentation as reflected in longer hospital stays. Columns 4 and 8 imply that, despite the reduced probability of being admitted, black men experienced an increase in duration of stay of between 0.57 and 0.70 nights per thousand kilometers, or 0.4 – 0.5 nights per standard deviation. These estimates are large, representing 21 – 26% of the pre-1972 sample mean.

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 Macon County, Alabama are slightly *more* likely to visit the dentist after 1972. This result provides

---

<sup>60</sup>The standard deviation of closeness to tuskegee for the utilization data is 0.68 (000 kilometers) and for the mortality data is 0.81 (000 kilometers). For reference, these values approximately reflect the distance from Tuskegee, AL to Louisville, KY and Indianapolis, IN, respectively. Additional summary statistics are available in Appendix Table 1.



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.

Our baseline mortality results, estimates of  $\beta$  from equation 3, 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 address both the occasional mortality outliers in the western U.S. apparent in Figure 3 and, more importantly, the concern that effects we measure are a result of comparing western SEAs to southern SEAs, the latter experiencing a generalized post-1972 reduction in black male health.

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.<sup>61</sup> 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 an increase in post-1972 mortality for black men of 4.2 log points per thousand kilometers (a 3.4 log point increase per one standard deviation increase in proximity). When we estimate the model using an all-black sample, we find even larger results: 8.4 log points per thousand kilometers, or a 6.8 log point increase per standard deviation. These are sizable effects, but not fully out of line with other findings in the literature. Results from Bailey and Goodman-Bacon (2015) show that exposure to a community health center in one's county of residence between 1965 and 1974 reduced adult all-cause mortality by as much as 2%. The authors' results indicate sizable mortality effects related to the utilization of primary care facilities consistent with our findings that both outpatient utilization and mortality suffered in the study's aftermath.

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. Larger results for log chronic mortality compared to all age-adjusted mortality are to be expected if the Tuskegee disclosure primarily affects chronic diseases and the former is a subset of the latter. Our estimates in columns 3 and 7 imply an increase in post-1972 mortality for older black men of 6.5 to 10.5 log points per thousand kilometers of proximity to Macon County. As a result, a standard deviation increase in proximity is associated with an increased mortality rate of 5.3 to 8.5 log points. Notably, we do not observe an effect of the Tuskegee disclosure on mortality from violent causes of death. (See Appendix Table 7.)

---

<sup>61</sup>Our results are robust to a levels specification, as reported in Appendix Table 4.

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 (Table 2, columns 2, 4, 6 and 8), an indication that the nationwide results are not merely reflecting 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 higher in the southern U.S. than elsewhere, the southern restriction leaves more than 40% of the original sample available for estimation. For these SEAs, coefficient estimates contained in columns 2, 4, 6 and 8 are slightly larger than the corresponding values for the entire U.S., consistent with the observations from Figure 3. 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.

### 6.3 Heterogeneous Effects

In Table 3, we explore heterogeneous effects by undertaking a series of sample bifurcations and reporting estimates of  $\beta$  for these subsamples. First, we explore the roles of income and education by dividing the all-male and all-black samples at the median of black male household income (columns 1 and 2) and at the median of black male education (columns 3 and 4). (The entire sample is bifurcated at the black male median, so above- and below- median samples are not equally sized.) 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 study's victims. Indeed, although the estimates of  $\beta$  are negative and significant for both income groups in the all-male and all-black sample: the point estimates for lower income black men in column 2 of Panels A and B are substantially larger. Splitting the sample by educational attainment generates a similar result (columns 3 and 4); the estimates by education are significantly distinguishable from each other at the 1% level for both samples.

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 (US Census Bureau, 1970). 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 samples by places above and below the median number of black doctors (as a percentage of all doctors), we find suggestive evidence for a moderating effect of black doctors though the differences are not statistically significant.<sup>62</sup>

## 7 Threats to Identification and Robustness Checks

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. 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 networks as an alternative measure of proximity 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.

### 7.1 Placebo Locations

To reinforce that the geographic gradients documented in our baseline DDD results are specific to distance from Macon County, Alabama, we run 49 placebo regressions, substituting for the baseline proximity measure (proximity to Macon County, Alabama) with the proximity to the geographic centroid of every other state and re-estimating the model. The outcome variable of this exercise is the intensive margin of primary care utilization (columns 1 and 5 of Table 1). These regressions 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  $\beta$  in each of these tests. The vertical red line indicates the estimated coefficient from Table 1 when the geographic proximity measure is based on the SEA housing Macon County, Alabama. The absolute value of the estimate is greater than 92% of placebo estimates for both the within-male and within-black specification.

---

<sup>62</sup>We do not find the same pattern if we use black physicians per black capita, suggesting that black men responded to the "whiteness" of the local medical profession, rather than to the availability of a black physician.

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 (excluding the SEA containing Macon County) and re-estimating equation 3 using log chronic mortality as the outcome. These results, located in the bottom portion of Figure 6, indicate that the value of  $\beta$  estimated using proximity to Macon County as the center of the geographic proximity metric is greater than 92% (98%) of all other coefficients for the all-male (all-black) samples. "Heat maps" of our mortality estimates for Panel C and D are provided in Appendix Figure 3. (The same maps cannot be disclosed for the utilization data.)

## 7.2 Other Demographic Groups

In this section, we perform additional placebo tests related to the paper's main assumptions that older black men would have been most affected by the Tuskegee disclosure, 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. We restrict our presentation to a subset of utilization outcomes due to space constraints. If successful, these placebo tests help rule out factors that may be confounding the interpretation of our baseline coefficients.

Table 4 contains estimates of  $\beta$  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. 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 both 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. Results for adolescents, teens, and for young adults (less than 44) are also statistically indistinguishable from zero.<sup>63</sup>

<sup>63</sup>Available upon request. We created an age-adjusted mortality rate for individuals 15 to 44 and did not find a significant effect in either a within-male or within-black estimating sample. Similarly, the adolescent and teen group, ages 10 to 19, mirror the insignificance of the 1-9 age group. Estimates for  $\beta$  when the outcome is infant mortality are negative and significant for the all-male sample (but not for the all-black sample), a result consistent with other literature on southern infant mortality convergence in this era. However, in this case, the gradient is not specific to Macon County, Alabama.

Our baseline within-black specifications implicitly test 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 2 above, replacing *blackmale<sub>g</sub>* with *blackfemale<sub>g</sub>* and using an all-female sample comparing black women to their white peers. The resulting estimates for  $\beta$  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 of proximity to Tuskegee. We perform a similar analysis for white men, replacing *blackmale<sub>g</sub>* with *whitemale<sub>g</sub>* in equation 2 and performing estimation over an all-white sample. Estimates for  $\beta$ , 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 insignificant 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. Because chronic mortality is not largely relevant for young children, these are not reported here. In all but one case, we observe post-1972 gradients by proximity to Tuskegee that are insignificantly different from zero. The exception is the coefficient on  $P_a * post_t * whitemale_g$  which is positive and statistically significant, but economically insignificant and an order of magnitude smaller than our estimate for black men compared to black women.<sup>64</sup> This positive result is also not robust; it is not significant in the levels specification or in a sample restricted to the South. For black women compared to their white peers, the point estimates, although insignificant, are negative and consistent with Panel C of Figure 3; the mortality of black women generally improved in the South relative to the rest of the country after 1972, but the gradient from Macon County, Alabama, in particular, is weak.

Taken together, the scope of identification threats is narrowed substantially to those factors correlated with the primary care utilization and mortality rates of older 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.

---

<sup>64</sup>The coefficient on the child sample is also marginally significant when restricted to the South but has the "wrong" sign.

### 7.3 Robustness Checks

Next, we demonstrate that the baseline results are robust to alternative empirical specifications and alternative measures of proximity. Table 5 reports these checks for the utilization results regarding number of outpatient visits and Table 6 provides analogous results for log chronic mortality. Again, all-male samples are contained in Panel A and all-black samples in Panel B. 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 both tables, 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. The interpretation of the magnitudes of these coefficients differs from the baseline results using geographic proximity to Macon County, Alabama as the proximity measure, but the estimated effect per standard deviation is remarkably similar, 3.6 to 3.8 log points of chronic mortality in Table 6.<sup>65</sup> In columns 2 and 7, we replace the continuous geographic proximity measure with proximity "bins", defined as cardinal measures of geographic adjacency, by state, to Alabama.<sup>66</sup> 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 bin.<sup>67</sup>

Columns 3, 4, 8 and 9 of both Tables 5 and 6 provide results for alternative specifications of our baseline results. In column 3 of both tables, we add group-location (state or SEA) fixed effects in lieu of the  $P_s * blackmale_g$  interaction term in equation 2 and in lieu of  $P_s * blackmale_g$  in equation 3. 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 parametric, linear 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

<sup>65</sup>The standard deviation of the migrant variable in the mortality sample is 0.2. See Appendix Table 1.

<sup>66</sup>Bordering states of Alabama are scored as "9", states that border bordering states that do not themselves border Alabama are scored as "8", etc.

<sup>67</sup>The bin proximity values range from 1 to 9. As a result, the coefficients in this robustness check should be smaller than in the baseline specification (based on geographic proximity, ranging from 0 to 3.5), precisely as we observe.

control for the presence of government programs which might have differentially enrolled black males, provided those enrollment patterns are static over time. In all cases, we fail to reject the null hypothesis that the estimated coefficients are the same under the baseline model and the models with additional fixed effects.<sup>68</sup>

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  $\beta$  relative to the baseline.

Finally, columns 5 and 10 of Table 5 restrict 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  $\beta$  is relatively unchanged in this specification compared to our baseline results although the estimates are no longer statistically significant.

In column 5 of Table 6, we weight our baseline estimates by the size of the older black population in each SEA. 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.

## 8 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 "doctor's 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

---

<sup>68</sup>The p-value of the test comparing the triple coefficient using the main specification versus a fully saturated specification is  $>0.10$  in both columns 3 and 8 of Table 5. This is the finest level of detail we have disclosed. In the mortality results, the p-value is 0.43 for the within-male sample and 0.15 for the within-black sample.

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 residing in state  $s$  is given by:

$$\text{Medical Mistrust}_{igs} = \alpha + \beta(P_s * \text{blackmale}_g) + \mu(\text{blackmale}_g) + \delta \text{mistrust}_i + \pi_s + \gamma X_i + \epsilon_{igs} \quad (3)$$

where  $P_s$  is the proximity to Macon County and  $\pi_s$  is a current-state fixed effect. These state fixed effects ensure that  $\beta$  captures the geographic gradient in mistrust for black males net of any gradient apparent in the comparison population (white males of black females). The sample includes individuals at least 10 years of age at the time of the disclosure.<sup>69</sup>  $X$  contains age, marital status and education categories. In addition, we condition on an individual's general level of mistrust in others,  $\text{mistrust}_i$ , to isolate the impact of medical mistrust.<sup>70</sup> 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 exhibit a strong, statistically significant geographic gradient in both mistrust of doctors and concern regarding treatment denial. A one standard deviation increase in the proximity to Tuskegee, Alabama results in an increased probability of medical mistrust on the order of 9 percentage points compared to white men and 12 percentage points compared to black women. For concerns regarding treatment denial, the magnitudes are 8 and 24 percentage points, respectively. We find no statistically significant corresponding effects for black women relative to white women or for white men relative to white women (see Appendix Table 9).

Pursuing further evidence on the channels through 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

<sup>69</sup>This implies that age in the sample ranges from 36 to 89 with an average in the mid-fifties. Note that, if we restrict to age < 10 in 1972, so those ages 18 to 35, our coefficients on medical mistrust and deny treatment generally have the "wrong sign" and are not statistically significant. We also created a histogram of placebo locations as in Figure 6 and find that our estimates are greater than 92% and 76% of all possible coefficient estimates.

<sup>70</sup>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 for medical mistrust, so that the coefficients can be easily interpreted as marginal effects. See Appendix section 2.3.3. Estimation using an ordered logit produces similar results.



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 2 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.

## 9 Conclusion

The Tuskegee Study was one of the most egregious examples of medical exploitation in U.S. history. Our estimates indicate that the years following disclosure of the study's tactics brought significantly lower utilization 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 such men 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 and its revelation reduced black life expectancy at age 45 by 1.4 years, accounting for 35% of the black to white-male life expectancy gap in 1980.

Our findings underscore the importance of trust for economic relationships involving imperfect information. Typically the literature on trust has focused on trade settings (Greif, 1989); 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.

Marcella Alsan, *Stanford University and NBER*

Marianne Wanamaker, *University of Tennessee Knoxville and NBER*

## References

- ABC News/The Washington Post (1992). ABC News/The Washington Post: Los Angeles Race Riots.
- Adler, N., W. Boyce, M. Chesney, et al. (1993). Socioeconomic inequalities in health: no easy solution. *JAMA* 269(24), 124–163.
- Aizer, A. and L. Stroud (2010, March). Education, Knowledge and the Evolution of Disparities in Health. Working Paper 15840, National Bureau of Economic Research.
- Almond, D., K. Y. Chay, and M. Greenstone (2006, December). Civil Rights, the War on Poverty, and Black-White Convergence in Infant Mortality in the Rural South and Mississippi. Working Paper 07-04, MIT Department of Economics.
- Alsan, M. and C. Goldin (2015, June). Watersheds in Infant Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880 to 1915. Working Paper 21263, National Bureau of Economic Research.
- Alsan, M. M. and D. M. Cutler (2013). Girls' education and HIV risk: Evidence from Uganda. *Journal of Health Economics* 32(5), 863 – 872.
- American Cancer Society (2008). Colorectal Cancer Facts and Figures. Technical report, American Cancer Society, Atlanta.
- American Cancer Society (2009). Cancer Facts and Figures 2009-2010. Technical report, American Cancer Society, Atlanta.
- Arias, E. (2014). U.S. Life Tables 2010. *National Vital Statistics Reports* 63(7).
- Ayer (1971). *Ayer Directory of Newspapers, Magazines and Publications*. Ayer press.
- Baicker, K., A. Chandra, and J. Skinner (2005). Geographic Variation In Health Care And The Problem Of Measuring Racial Disparities. *Perspectives in biology and medicine* 48(1), 42–S53.
- Baicker, K., A. Chandra, J. S. Skinner, and J. E. Wennberg (2004). Who You Are And Where You Live: How Race And Geography Affect The Treatment of Medicare Beneficiaries. *Health* 23(VAR-33).
- Bailey, M. J. and A. Goodman-Bacon (2015). The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans. *American Economic Review* 105(3), 1067–1104.

- Banerjee, A. and E. Duflo (2012, March). *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty* (Reprint ed.). PublicAffairs.
- Becker, G. and Y. Rubinstein (2011). Fear and the Response to Terrorism: An Economic Analysis. Technical report, Centre for Economic Performance Working Paper No. 1079.
- Bhalotra, S. R. and A. Venkataramani (2012, December 19). Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions. *Human Capital Investments, and Institutions*, 1–65.
- Black, D. A., S. G. Sanders, E. J. Taylor, and L. J. Taylor (2015). The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South. *American Economic Review* 105, 477–503.
- Blinder, A. (2016). In Rural Alabama A Longtime Mistrust of Medicine Fuels A Tuberculosis Outbreak. *The New York Times*.
- Bogart, L. M., G. J. Wagner, and H. D. G. J. et al. (2016). Medical Mistrust among social network members may contribute to antiretroviral nonadherence in African Americans living with HIV. *Social Science & Medicine Available Online 19 March 2016*.
- Boulware, L. E., L. A. Cooper, L. E. Ratner, T. A. LaVeist, and N. R. Powe (2003, July-August). Race and Trust in the Health Care System. *Public Health Reports* 118, 358–365.
- Boustan, L. and R. A. Margo (2014, December). Racial Differences in Health in Long-Run Perspective: A Brief Introduction. Working Paper 20765, National Bureau of Economic Research.
- Brandon, D. T., L. A. Isaac, and T. A. LaVeist (2005). The Legacy of Tuskegee and Trust in Medical Care: is Tuskegee Responsible for Race Differences in Mistrust of Medical Care? *Journal of the National Medical Association* 97(7), 951–956.
- Brandt, A. M. (1978). Racism and Research: The Case of the Tuskegee Syphilis Study. *The Hastings Center Report* 8(6), 21–29.
- Byrd, W. M. and L. A. Clayton (2002, August). *An American Health Dilemma: Race, Medicine, and Health Care in the United States 1900-2000*, Volume II. Routledge.
- Carlson, E. D. and R. M. Chamberlain (2004). The Black White Perception Gap and Health Disparities. *Public Health Nursing* 21(4), 372–379.
- Cascio, E. and E. Washington (2014). Valuing the Vote: The Redistribution of Voting Rights and State Funds following the Voting Rights Act of 1965. *The Quarterly Journal of Economics* 129(1), 379–433.
- CDC (2001). HIV/AIDS in the United States.
- CDC (2010). Ten leading causes of death by age groups. Technical report.
- Centers for Disease Control and Prevention (2014, December). Compressed Mortality File (CMF 1968-1988, Series 20, No. 2A, 2000) on CDC WONDER online.
- Corbie-Smith, G., S. B. Thomas, M. V. Williams, and S. Moody-Ayers (1999). Attitudes and Beliefs of African Americans Toward Participation in Medical Research. *Journal of General Internal Medicine* 14(9), 537–546.

- Currie, J. and H. Schwandt (2016a). Inequality in mortality decreased among the young while increasing for older adults, 1990-2010. *Science* 352, 708–712.
- Currie, J. and H. Schwandt (2016b). Mortality Inequality: The Good News from a County-Level Approach. *Journal of Economic Perspectives* 30(2), 29–52.
- Curtin, L. R. and R. J. Klein (1995, March). *Direct Standardization (Age-Adjusted Death Rates)* (6 ed.). US Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics.
- Cutler, D., W. Huang, and A. Lleras-Muney (2014, May). When Does Education Matter? The Protective Effect of Education for Cohorts Graduating in Bad Times. Working Paper 20156, National Bureau of Economic Research.
- Cutler, D., A. Lleras-Muney, and T. Vogl (2011). Socioeconomic status and health: Dimensions and mechanisms. *Oxford Handbook of Health Economics*, 124–163.
- Cutler, D. M. and A. Lleras-Muney (2012, January). Education and Health: Insights from International Comparisons. Working Paper 17738, National Bureau of Economic Research.
- Deaton, A. (2002). Policy Implications of the Gradient of Health And Wealth. *Health Affairs* 21(2), 13–30.
- Dixit, A. (2003). Trade Expansion and Contract Enforcement. *Journal of Political Economy* 111(6), 1293–1317.
- Do, Y., W. Carpenter, P. Spain, J. Clark, et al. (2010). Race, health access and physician trust among prostate cancer patients. *Cancer Causes Control* 21(1), 31–40.
- Dupas, P. (2011). Health Behavior in Developing Countries. *Annual Review of Economics* 3(1), 425–449.
- Eriksson, K. and G. Niemesh (2016). Impact of Migration on Infant Health: Evidence from the Great Migration. *Working Paper*.
- Fafchamps, M. (2006). Development and Social Capital. *The Journal of Development Studies* 42(7), 1180–1198.
- Feldhusen, A. E. (2000). The History of Midwifery and Childbirth in America: A Time Line.
- Gaston, G. B. and B. Alleyne-Green (2013). The Impact of African Americans' Beliefs About HIV Medical Care on Treatment Adherence: A Systematic Review and Recommendations for Interventions. *AIDS Behavior* 17, 31–40.
- Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2014). United States Newspaper Panel, 1869-2004. Icpsr30261-v6, Inter-university Consortium for Political and Social Research [distributor], Ann Arbor, MI.
- Goodman-Bacon, A. (2015). Public Insurance and Mortality: Evidence from Medicaid Implementation. *University of California, Berkeley*.
- Greif, A. (1989). Reputation and coalitions in Medieval Trade: Evidence on the Maghribi Traders. *The journal of economic history* 49(04), 857–882.
- Guiso, L., P. Sapienza, and L. Zingales (2008). Alfred Marshall Lecture: Social Capital as Good Culture. *Journal of the European Economic Association* 6(2-3), 295–320.

- Gutsell, J. N. and M. Inzlicht (2010). Empathy constrained: Prejudice predicts reduced mental simulation of actions during observation of outgroups. *Journal of Experimental Social Psychology* 46(5), 841 – 845.
- Haines, M. (2010). Historical, Demographic, Economic, and Social Data: The United States, 1790-2002 Computer file. ICPSR 02896-v3. Technical report, Inter-university Consortium for Political and Social Research, Ann Arbor, MI.
- Halbert, C. H., K. Armstrong, O. Gandy, et al. (2006). Racial Differences in Trust in Health Care Providers. *Archives of Internal Medicine* 166(8), 896–901.
- Halbert, C. H., B. Weathers, E. Delmoor, et al. (2009). Racial Differences in Medical Mistrust among Men Diagnosed with Prostate Cancer. *Cancer* 115(11), 2553–2561.
- Hamilton, B., A. Hincapie, R. A. Miller, and N. W. Papageorge (2016). Innovation and Diffusion of Medical Treatment. Technical report, Working Paper.
- Hammond, W. P., D. Matthews, D. Mohottige, A. Agyemang, and G. Corbie-Smith (2010). Masculinity, Medical Mistrust, and Preventive Health Services Delays Among Community-Dwelling African-American Men. *Journal of General Internal Medicine* 25(12), 1300–1308.
- Heintzelman, C. (2003). The Tuskegee Syphilis Study and Its Implications for the 21st Century. *The New Social Worker* 10(4).
- Hertz, R. P., A. N. Unger, J. A. Cornell, and E. Saunders (2005). Racial Disparities in Hypertension Prevalence, Awareness and Management. *Archives of Internal Medicine* 165(18), 2098–2104.
- Holt-Lunstad, J., W. Birmingham, and B. Q. Jones (2008). Is There Something Unique about Marriage? The Relative Impact of Marital Status, Relationship Quality, and Network Social Support on Ambulatory Blood Pressure and Mental Health. *Annals of Behavioral Medicine* 35(2), 239–244.
- Hood, K. B., A. Hart, F. Belgrave, R. H. Tademy, and R. A. Jones (2012). The Role of Trust in Health Decision Making Among African American Men Recruited From Urban Barbershops. *Journal of the National Medical Association* 104, 351.
- Institute of Medicine (2003). *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care (full printed version)*. Washington, DC: The National Academies Press.
- Jackson, J. S., H. W. Neighbors, and G. Gurin (1979). National Survey of Black Americans.
- Jones, J. H. (1992, December). *Bad Blood, The Tuskegee Syphilis Experiment, New and Expanded Edition*. Free Press.
- Kaiser Family Foundation Fact Sheet (2007). The Health Status of African American Men in the United States. Technical report, Kaiser Family Foundation.
- Kayaniyil, S., S. Gravely-Witte, D. E. Stewart, et al. (2009). Degree and correlates of patient trust in their cardiologist. *Journal of evaluation in clinical practice* 15(4), 634–640.
- Keating, N., T. Gandhi, E. Orav, et al. (2004). Patient characteristics and experiences associated with trust in specialist physicians. *Archives of Internal Medicine* 164(9), 1015–1020.
- Knack, S. and P. Keefer (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly journal of economics*, 1251–1288.

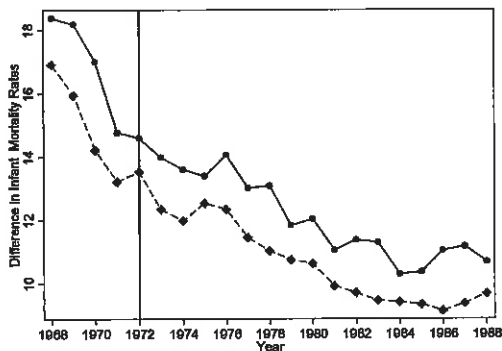
- LaVeist, T. A., K. Nickerson, and J. Bowie (2000). Attitudes about racism, medical mistrust and satisfaction with care among African American and white cardiac patients. *Medical Care Research Review* 57(1), 146–61.
- Livingston, G. and D. Cohn (2010, June). Childlessness Up Among All Women; Down Among Women With Advanced Degrees.
- Maddux, W. W. and M. B. Brewer (2005). Gender Differences in the Relational and Collective Bases for Trust. *Group Processes and Intergroup Relations* 8(2), 159–171.
- Mandell, G. L., J. E. Bennett, and R. Dolin (2009, October). *Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases* (7 ed.), Volume 2. Churchill Livingstone.
- Minnesota Population Center (2015). Integrated Health Interview Series: Version 6.11. Technical report, Minneapolis: University of Minnesota.
- Murphy, S. L., J. Xu, and K. D. Kochanek (2013). Deaths: Final Data for 2010. *National Vital Statistics Reports* 61(4).
- National Center for Health Statistics (1985). U.S Decennial Life Tables for 1979-81 . *DHHS Pub. No. (PHS) 85-1150-1 I(1)*.
- NBC (1989). NBC News Poll: R.A.C.E.–Racial Attitudes and Consciousness Exam [USNBC1989-RACE].
- Nelson, K., K. Norris, and C. Mangione (2002). Disparities in the Diagnosis and Pharmacologic Treatment of High Serum Cholesterol by Race and Ethnicity: Data from the Third National Health and Nutrition Examination Survey. *Archives of Internal Medicine* 162(8), 929–935.
- Nunn, N. and L. Wantchekon (2011). The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review* 101(7), 3221–52.
- Paratz, E. and B. Katz (2011). Ageing Holocaust Survivors in Australia. *Medical Journal of Australia* 194(1), 194–197.
- Reverby, S. M. (2000, July). *Tuskegee's Truths: Rethinking the Tuskegee Syphilis Study* (1 ed.). Studies in Social Medicine. The University of North Carolina Press.
- Reynolds, P. P. (1997). The Federal Government's Use of Title VI and Medicare to Racially Integrate Hospitals in the United States, 1963 Through 1967. *American Journal of Public Health* 87(11), 1850–1858.
- Robles, T. F. and J. K. Kiecolt-Glaser (2003). The Physiology of Marriage: Pathways to Health. *Physiology & Behavior* 79(3), 409 – 416. A Tribute to Paul MacLean: The Neurobiological Relevance of Social Behavior.
- Sandman, D., E. Simantov, and C. An (2000). Out of Touch: American Men and the Health Care System. Technical report, Commonwealth Fund Men's and Women's Health Survey Findings.
- Shayo, M. and A. Zussman (2011). Judicial Ingroup Bias in the Shadow of Terrorism. *The Quarterly Journal of Economics* 126, 1447–1484.
- Silber, J. H., P. R. Rosenbaum, R. N. Ross, B. A. Niknam, J. M. Ludwig, W. Wang, A. S. Clark, K. R. Fox, M. Wang, O. Even-Shoshan, et al. (2014). Racial Disparities in Colon Cancer Survival: A Matched Cohort Study. *Annals of Internal Medicine* 161(12), 845–854.

- Singer, T., B. Seymour, J. P. O'Doherty, K. E. Stephan, R. J. Dolan, and C. D. Frith (2006, January). Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439(7075), 466–469.
- Smith, T., P. Marsden, M. Hout, and J. Kim (2015). General Social Surveys, 1972-2014. Cumulative file, The Roper Center for Public Opinion Research, University of Connecticut.
- Tabellini, G. (2008). The Scope of Cooperation: Values and Incentives. *The Quarterly Journal of Economics* 123(3), 905–950.
- Thom, T., N. Haase, W. Rosamond, V. Howard, J. Rumsfeld, T. Manolio, Z.-J. Zheng, K. Flegal, C. O'Donnell, S. Kittner, D. Lloyd-Jones, D. C. Goff, and Y. Hong (2006). Heart Disease and Stroke Statistics. *Circulation* 113, e85–e151.
- Thomas, S. B. and S. C. Quinn (1991, November). The Tuskegee Syphilis Study, 1932 to 1972: Implications for HIV Education and AIDS Risk Education Programs in the Black Community. *American Journal of Public Health* 81(11), 1500–1501.
- Thomas, S. L. (2013). Understanding the Predictors of Patient Trust and Its Role in Utilization and Outcomes. Technical report, University of Alabama-Birmingham.
- Tuskegee Syphilis Study Legacy Committee (Ed.) (1996). *Final Report of the Tuskegee Syphilis Study Legacy Committee*, <http://exhibits.hsl.virginia.edu/badblood/report/>.
- US Census Bureau (1970). Census of Population and Housing Vol 1 Characteristics of the Population. Technical report, US Census Bureau.
- Viera, A. J., J. M. Thorpe, and J. M. Garrett (2006). Effects of sex, age, and visits on receipt of preventive healthcare services: a secondary analysis of national data. *BMC Health Services Research* 6, 1–8.
- White, R. M. (2005). Misinformation and Misbeliefs in the Tuskegee Study of Untreated Syphilis Fuel Mistrust in the Healthcare System. *Journal of the National Medical Association* 97(11), 1566–73.
- Wiltshire, J., S. D. Person, and J. Allison (2011, September/October). Exploring Differences in Trust in Doctors Among African American Men and Women. *Journal of the National Medical Association* 103(9), 845–851.
- Zheng, H. and C. Zhou (2008). The Impact of Hospital Integration on Black-White Differences in Mortality: A Case Study of Motor Vehicle Accident Death Rates. *Unpublished working paper*.

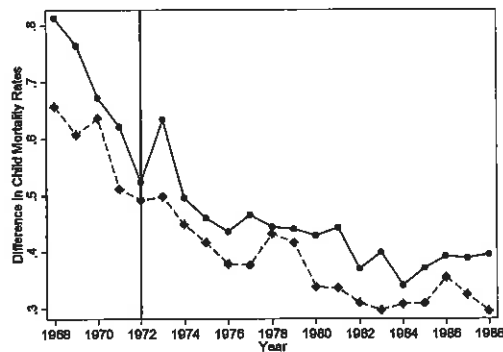
## 10 Figures

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

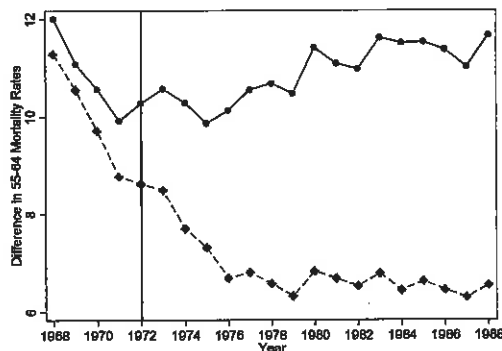
Panel A. Infant Mortality Rate



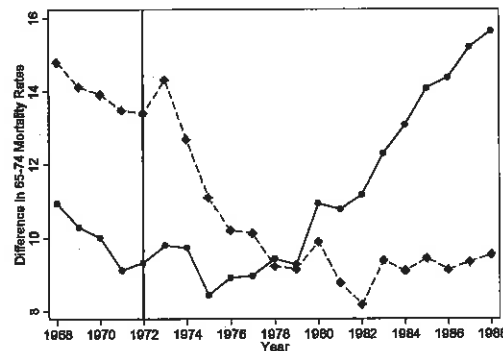
Panel B. Child Mortality Rate



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.



**FIGURE 2: PROXIMITY MEASURES**

**Panel A: Distance to Tuskegee**



**Panel B: Fraction Black Migrants from Alabama**

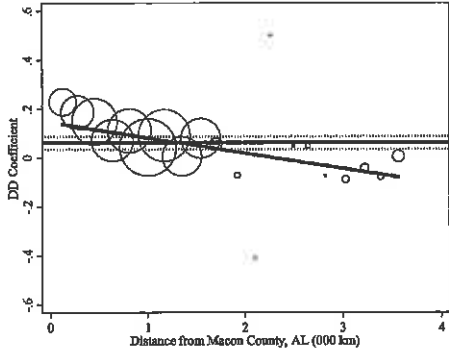


*Notes:* Figure 2 plots proximity measures. Panel A represents distance from Macon County for each SEA calculated using ArcGIS proximity tools. Panel B represents the fraction of black migrants from Alabama by SEA calculated using the 1940 census. Darker tones reflect closer to Tuskegee or a higher fraction of Alabama migrants using 20 natural breaks. The white circle in Panel A demarcates approximately 1000 kilometer radius from Macon County.

**FIGURE 3: DIFFERENCE-IN-DIFFERENCE COEFFICIENTS AND GEOGRAPHIC GRADIENTS**

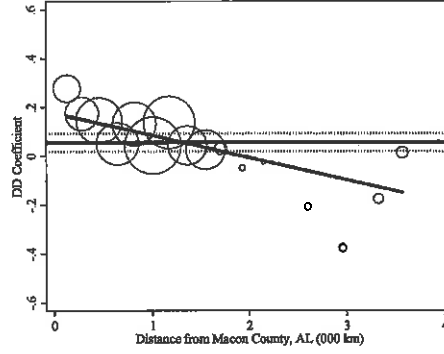
**Panel A.  $\theta_1$  - All Male Sample**

DD Coefficient: 0.062  
95% CI: [0.036, 0.088]



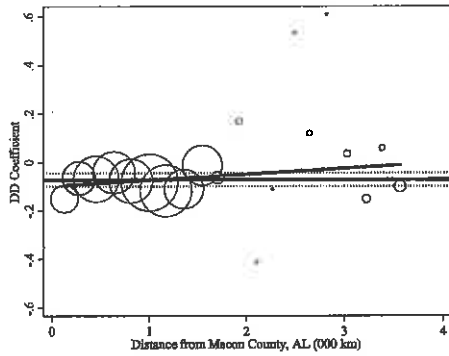
**Panel B.  $\theta_1$  - All Black Sample**

DD Coefficient: 0.056  
95% CI: [0.019, 0.093]



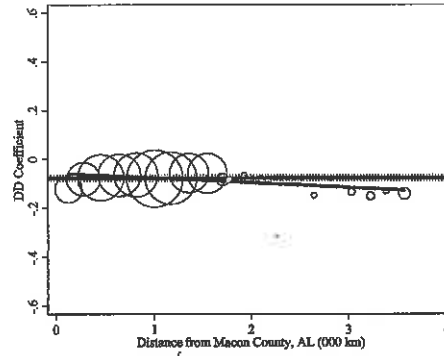
**Panel C.  $\theta_1$  - All Female Sample**

DD Coefficient: -0.073  
95% CI: [-0.101, -0.046]



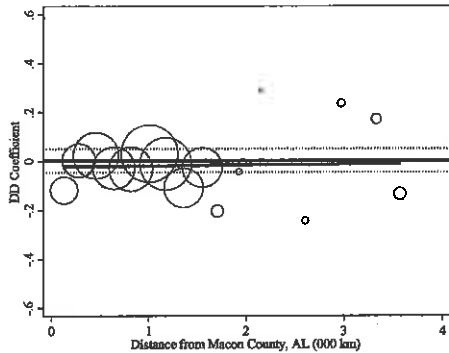
**Panel D.  $\theta_1$  - All White Sample**

DD Coefficient: -0.078  
95% CI: [-0.089, -0.066]



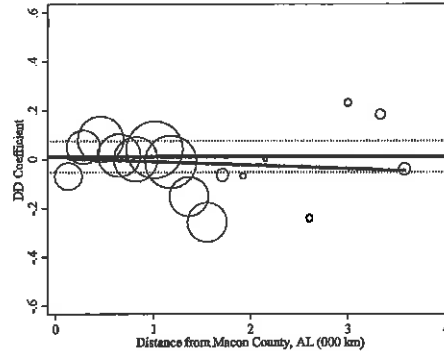
**Panel E.  $\theta_1$  - All Male Sample with False Post**

DD Coefficient: 0.002  
95% CI: [-0.049, 0.049]



**Panel F.  $\theta_1$  - All Black Sample with False Post**

DD Coefficient: 0.010  
95% CI: [-0.054, 0.074]

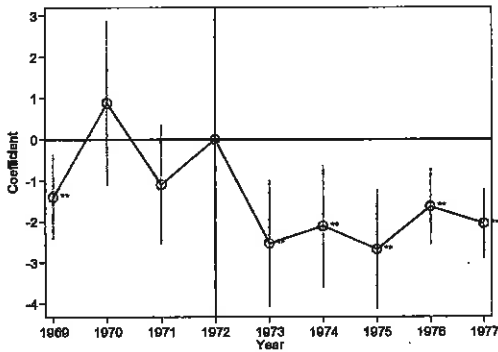


*Notes:* These figures represent the coefficient in each SEA-specific regression for log chronic mortality averaged across all SEAs in intervals of approximately 180 kilometers of distance from Macon County, Alabama. (The maximum observed distance from Tuskegee is divided into 20 intervals of equal distance 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 of 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, respectively.

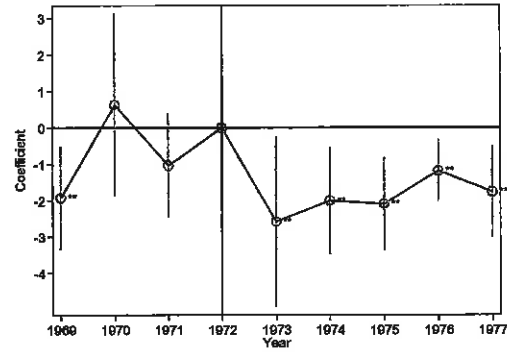
f

FIGURE 4: EVENT STUDY ON UTILIZATION

Panel A: Black vs. White Males



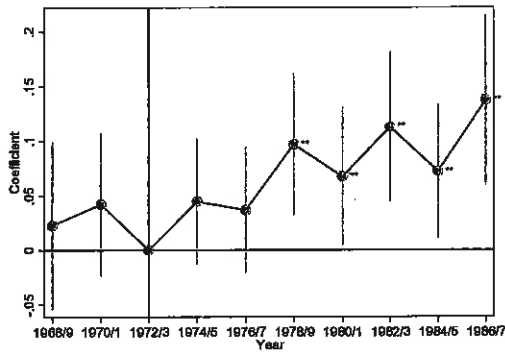
Panel B: Black Men vs. Black Women



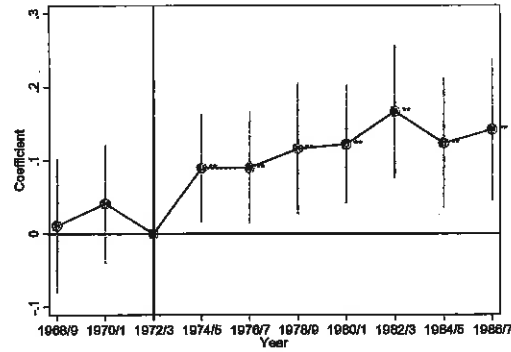
Notes: Event study of the coefficient on the interaction of proximity and a black male indicator for each year. The dependent variable is the number of outpatient visits within the last 12 months. Data are from the harmonized version of the National Health Interview Survey (NHIS) available from IPUMS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

FIGURE 5: EVENT STUDY ON MORTALITY

Panel A: Black vs. White Males



Panel B: Black Men vs. Black Women



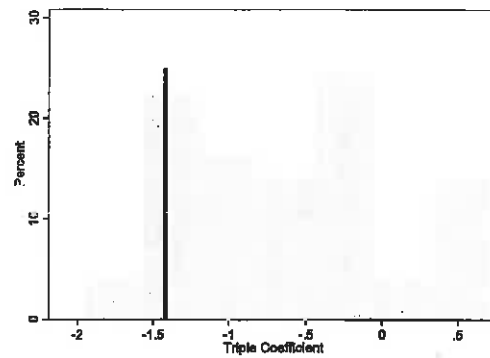
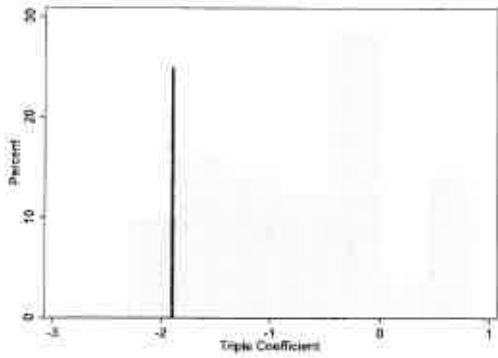
Notes: Event study of the coefficient on the interaction of proximity and a black male indicator for each two-year period. The dependent variable is log chronic disease mortality for demographic groups ages 45 to 74. Data are from the compressed mortality files provided by the CDC. Panel A is estimated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

FIGURE 6: PLACEBO DISTANCES (EMPIRICAL DISTRIBUTION)

Outcome: Number Outpatient Visits

Panel A: Male Sample

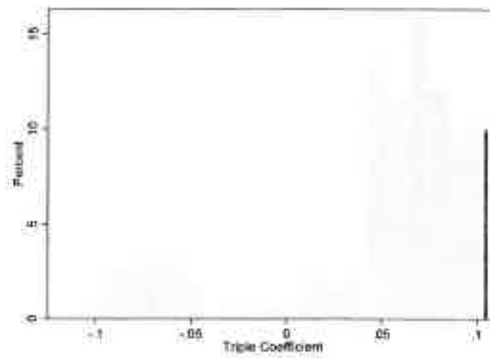
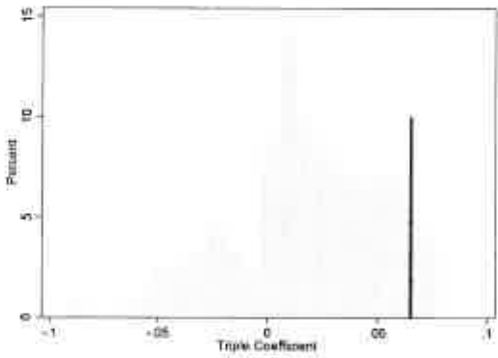
Panel B: Black Sample



Outcome: Log Chronic Mortality

Panel C: Male Sample

Panel D: Black Sample



*Notes:* Frequency of the  $\beta$  coefficient using distance from every other state (exclusive of Alabama) or SEA (exclusive of the one containing Macon County) 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  $\beta$  coefficient estimate for the true treatment distance (reported in Tables 1 and 2).

TABLE 1: BASELINE UTILIZATION RESULTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable:							
	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
<i>Mean of Dep. Var for Black Men before 1973</i>	0.76	4.23	0.10	1.82	0.76	4.23	0.10	1.82
	PANEL A--Male Sample				PANEL B--Black Sample			
$P_s^* \text{post}_t^* \text{blackmale}_g$	-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
Adj R-squared	0.029	0.023	0.019	0.016	0.059	0.047	0.028	0.026

Notes: OLS Regressions of equation 2. 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 ages 45-74. 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 mean of the dependent variable for black men up to and including 1972 is also reported. Standard errors are clustered at the state level. \*\*\*p<0.01, \*\* p<0.05 and \*p<0.10, respectively.

**TABLE 2: BASELINE MORTALITY RESULTS**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Dependent variable:							
	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
<i>Mean of Dep. Var for Black Men before 1973</i>	3.37	3.41	2.89	2.88	3.37	3.41	2.89	2.88
	<b>PANEL A--Male Sample</b>							
$P_a$ post <sub>t</sub> *blackmale <sub>g</sub>	0.042*** (0.016)	0.052* (0.028)	0.065*** (0.018)	0.103*** (0.037)	0.084*** (0.024)	0.095*** (0.034)	0.105*** (0.021)	0.134*** (0.046)
Fixed Effects		SEA-Year, Race-Year				SEA-Year, Gender-Year		
Observations	8548	3708	8382	3694	8106	3694	7898	3668
No. Clusters	454	186	461	186	439	186	436	186
Adj R-squared	0.43	0.68	0.26	0.58	0.56	0.78	0.55	0.74
	<b>PANEL B---Black Sample</b>							

*Notes:* OLS Regressions of equation 3. Data are from the CDC compressed mortality files on the non-institutionalized U.S. 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 45-74 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 U.S. 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 mean of the dependent variable for black men in the years up to and including 1972 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**

(1)	(2)	(3)	(4)	(5)	(6)	
<i>Dependent Variable: Number Outpatient Visits</i>						
By Income Level		By Educational Status		By Prevalence Black Doctors (as pct of all doctors)		
<i>income &gt; black male median</i>	<i>income &lt;= black male median</i>	<i>educ &gt; black male median</i>	<i>educ &lt;= black male median</i>	<i>black MD &gt; median</i>	<i>black MD &lt;= median</i>	
<b>PANEL A—Male Sample</b>						
$P_s * post_t * black_{male}_g$	-1.142** (0.507)	-2.973*** (0.623)	-0.0625 (0.345)	-3.578*** (0.766)	-1.808*** (0.424)	-2.749*** (0.685)
Fixed Effects			State-Year, Race-Year			
Observations	41,368	19,469	46,179	26,547	34,017	26,820
No. Clusters	49	49	49	49	17	32
Adj R-squared	0.019	0.047	0.021	0.036	0.023	0.024
<b>PANEL B—Black Sample</b>						
$P_s * post_t * black_{male}_g$	-0.536 (0.557)	-1.892** (0.752)	0.0749 (0.411)	-3.831*** (1.354)	-1.441*** (0.463)	-2.072*** (0.791)
Fixed Effects			State-Year, Gender-Year			
Observations	8,551	10,415	11,040	7,926	14,219	4,747
No. Clusters	41	43	44	43	17	27
Adj R-squared	0.051	0.065	0.058	0.075	0.043	0.074

*Notes:* OLS Regressions of equation 2 assessing heterogeneous effects to a) income level b) education level and c) black doctors. Specifically, in the first two columns we divide the sample by median black male income. In the following two columns we divide the sample by median black male education. 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. 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.



**TABLE 4: PLACEBO UTILIZATION AND MORTALITY RESULTS**

Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male Children	Black Children	Female	White	Male Children	Black Children	Female	White
	<b>PANEL A--Any Outpatient Visit</b>							
$P_s \text{post}_t^* \text{blackmalechild}_g$	0.002 (0.007)	0.010 (0.008)			0.015 (0.093)	-0.119 (0.119)		
$P_s \text{post}_t^* \text{blackfemale}_g$			0.004 (0.006)				-0.127 (0.140)	
$P_s \text{post}_t^* \text{whitemale}_g$				0.004 (0.005)				-0.008 (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
	<b>PANEL B--Number Outpatient Visits</b>							
	<b>PANEL C--Log Age-Specific and Age-Adjusted Mortality (Full Sample)</b>							
$P_a \text{post}_t^* \text{blackmalechild}_g$	0.002 (0.038)	0.014 (0.033)			-0.083 (0.078)	-0.132* (0.076)		
$P_a \text{post}_t^* \text{blackmalechild}_g$			-0.021 (0.021)				-0.050 (0.037)	
$P_a \text{post}_t^* \text{blackmalechild}_g$				0.007** (0.003)				-0.005 (0.014)
Fixed Effects		SEA-Year, Group-Year				SEA-Year, Group-Year		
Observations	5808	4862	8286	9280	3104	2822	3698	3720
No. Clusters	408	341	450	464	182	171	186	186
Adj R-squared	0.38	0.44	0.53	0.97	0.42	0.40	0.78	0.97

*Notes:* OLS Regressions of equation 2 (Panels A and B: Utilization Outcomes) and equation 3 (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) (2) (5) and (6) refer to outpatient utilization 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 U.S. 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 45-74 in 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. Standard errors are clustered at the State level for Panels A and B and the SEA level for Panels C and D. \*\*\* $p < 0.01$ , \*\* $p < 0.05$  and \* $p < 0.10$ , respectively.

**TABLE 5: ROBUSTNESS CHECKS, UTILIZATION**

*Dependent Variable: Number Outpatient Visits*

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	<i>Alternative Instruments</i>		<i>Alternative Specifications</i>		<i>Sample Restrictions</i>	<i>Alternative Instruments</i>		<i>Alternative Specifications</i>		<i>Sample Restrictions</i>
		Bins	Addnl Fixed Effects	Placebo Distance (Dallas)	South Only	Migrants	Bins	Addnl Fixed Effects	Placebo Distance (Dallas)	South Only
$P_{s} \cdot \text{post}_t \cdot \text{blackmale}_g$	-10.34** (4.575)	-1.234*** (0.276)	-1.800*** (0.352)	-1.617*** (0.350)	-1.838 (1.437)	-9.506** (4.386)	-0.829** (0.360)	-1.389*** (0.410)	-1.334*** (0.361)	-1.459 (1.591)
$P_{\text{placebo}} \cdot \text{post}_t \cdot \text{blackmale}_g$				-0.626 (0.463)					-0.184 (0.556)	
<b>Fixed Effects</b>										
		State-Year, Race-Year	State-Year, Race-State, Race-Year	State-Year, State-Year, Race-Year	State-Year, Race-Year	State-Year, Gender-Year	State-Year, Gender-Year	State-Year, 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

*Notes:* OLS Regressions of equation 2 testing robustness to a) alternative treatments b) alternative specifications and c) sample restrictions. 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 cardinal measure of geographic adjacency instead of linear distance as the predictor variable. (Neighboring states are highest value, neighbors of neighbors are second highest, etc.) In columns (3) and (8), we run a fully saturated specification replacing group\*proximity with location-group 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 restrict the sample to the South (Census Region III). The data source, sample and outcome variables were described in the notes accompanying Table 1. 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 columns (3) and (8)). 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.



**TABLE 7: EFFECT OF TUSKEGEE ON BELIEFS ABOUT MEDICAL CARE**

<i>Outcome</i>	(1) <i>Medical Mistrust</i>	(2) <i>Deny Treatment</i>	(3) <i>Medical Mistrust</i>	(4) <i>Deny Treatment</i>
	<b>PANEL A — Male Sample      PANEL B— Black Sample</b>			
$P_g$ blackmale <sub>g</sub>	0.1231** (0.0565)	0.1012** (0.0435)	0.1538** (0.0730)	0.3207*** (0.0675)
Fixed Effects	State <sub>1998</sub>			
Observations	329	329	118	118
R-squared	0.15	0.18	0.36	0.42
No. Clusters	36	36	28	28

*Notes:* OLS estimates of equation 4. 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 males and Panel B restricts the sample to black males and black females. 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.

**TABLE 8: BASELINE RESULTS, UTILIZATION VETERANS**

Outcome	(1)	(2)	(3)	(4)
	Any Outpatient Visit	Number Outpatient Visits	Any Outpatient Visit	Number Outpatient Visits
	PANEL A - All Black Men		PANEL B - All White Men	
$P_s \text{ post}_t^* \text{ nonvet}_g$	-0.072*** (0.014)	-1.731*** (0.396)	0.194 (0.119)	0.011 (0.007)
Fixed Effects	State-Year, Group-Year		State-Year, Group-Year	
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 2 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 nonveteran and nonveteran-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.