Tuskegee and the Health of Black Men

Marcella Alsan and Marianne Wanamaker*

First version: November 2015
PRELIMINARY. PLEASE DO NOT CIRCULATE.

Abstract

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

*We are grateful to William Collins, Joe Ferrie, Nathan Nunn, Achyuta Adhvaryu, Arun Chandrasekhar, Martha Bailey and participants at NBER DAE, University of Tennessee-Knoxville and Vanderbilt for constructive comments. We thank Andrew Goodman-Bacon for facilitating the use of geographic identifiers in the NHES data. We thank the CDC for providing access and to the administrators at the Atlanta and Stanford Census Research Data Centers for their help in navigating the restricted data.
1 Introduction

The Tuskegee study became a symbol of their mistreatment by the medical establishment, a metaphor for deceit, conspiracy, malpractice, and neglect, if not outright genocide.

Corbie-Smith et al (1999)

The Tuskegee Study of Untreated Syphilis in the Negro Male (TSUS) is frequently cited as an important factor contributing to the alienation of black Americans from mainstream medicine. For 40 years, between 1932 and 1972, the U.S. Public Health Service followed hundreds of black men in Tuskegee Alabama, many of whom had syphilis, for the stated purpose of understanding the natural history of the disease. The men were denied highly effective treatment for their condition (most egregiously, penicillin, which became standard of care in 1947) 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. Survivors later reported that study doctors diagnosed them with "bad blood" for which they believed they were being treated. Compensation for participation included hot meals, the guise of treatment and burial payments.

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

In this paper, we test the hypothesis that Tuskegee contributed to racial health disparities in the years following 1972. To do so, we rely on a variety of survey and administrative data, including measures of trust in doctors in the General Social Survey (GSS) (Tom Smith and Peter Marsden and Michael Hout and Jibum Kim, 2015), health seeking behavior reported in the National Health Interview Survey (NHIS) (Minnesota Population Center and State Health Access Data Assistance Center, 2015), and detailed annual mortality

---

1 Contemporary coverage of TSUS was widespread, including major U.S. newspapers such as The New York Times and The Chicago Tribune.

2 An October 2015 episode of ABC sitcom “Black-ish”, a sitcom which centers on an African-American family, begins with a discussion of the Tuskegee experiment and the effect it has on the health seeking behavior of the protagonist’s aging father.

3 For a comprehensive review of racial inequalities in US medical care see the Institute of Medicine (2003) Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care. Black men have the lowest life expectancy of any demographic group in the United States. Although gaps are closing, as of 2010 black men can expect to live 5 years less than their white male peers and 6.3 fewer years than black females (US Department of Health and Human Services, 2013 p. 8).
data available by race, age group, and gender 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 behaviors of younger individuals are more closely related to acute conditions such as childbirth or trauma for which medical treatment (in contrast to primary or secondary prevention) is paramount.\(^4\)

A variety of sources document that the Tuskegee study disclosure disproportionately affected the psyche of individuals who identified most closely with the black male subjects (Boulware et al., 2003; Brandon et al., 2005; Corbie-Smith et al., 1999, 2002; Eaton et al., 2015; Hood and Hart Jr., 2012; Wiltshire et al., 2011).\(^5\) We utilize geographic distance from Tuskegee as well as black migration from Alabama to capture connectedness to the study’s subjects, interacted with indicator variables for black and male. Our Difference-in-Difference-in-Difference (DDD) framework compares the health utilization and outcomes of demographic groups (e.g. black men versus white men or black women, the first difference) before and after 1972 (second difference) with an interaction for cultural or geographic proximity to Tuskegee, Alabama (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, 2010, 2012; Cutler et al., 2014) income, (Deaton, 2001, 2002), marital status (Robles and Kiecolt-Glaser, 2003; Holt-Lunstad et al., 2008), and urbanization. Our main specifications include state-year fixed effects to capture the state-specific implementation and diffusion of healthcare programs. For mortality outcomes, which are measured at the level of state economic areas (SEAs), we can also include SEA-year fixed effects.

The impact of the TSUS disclosure are interpreted within a multi-period model of health belief formation and healthcare utilization with heterogeneous agents. The agents vary along two dimensions: their medical experience as young adults and their connectedness to Tuskegee. Agents move through life’s stages updating their beliefs about whether doctors can be trusted, incorporating both positive and negative signals via Bayes’ Rule. Subgroups who are exposed to routine medical care in young adulthood, such as women during pregnancy or military personnel during routine physicals for battle preparedness, are predicted to be have a more muted behavioral response to the Tuskegee disclosure relative to black men who lack such prior

---

\(^4\)Injuries and violence are the leading cause of death for US children and young adults between ages 1 and 44 (CDC, 2010).

\(^5\)The physiological basis of these findings has recently been elucidated by neurocognitive research on how empathetic responses are modulated (Singer et al., 2006); and how such responses are stronger for a member of one’s own group (Gutsell and Inzlicht, 2010).
exposure.\(^6\) We empirically test the short-run effects of the disclosure as predicted by the model. Tuskegee may have long-run effects that are difficult to quantify given our empirical strategy and data limitations. Instead, the empirical analysis focuses on the historical impact of the disclosure, which has not been rigorously identified, yet is often described as a watershed moment in the relationship between black Americans and the healthcare system.\(^7\)

We estimate that the Tuskegee disclosure reduced utilization of routine care (physician contacts) among older black males by 1.9 interactions per year per 1000 kilometers of proximity to Tuskegee when compared to their white peers and 1.4 interactions per year relative to black women. These estimates imply that older black men residing near Macon County, Alabama reduced their physician interactions by approximately one-third after 1972. Our mortality results imply an annual increase in black male mortality of between 6 and 9 percent per thousand miles of proximity to Tuskegee, the equivalent of between 2.1 and 3.6 deaths per thousand.

An alternative explanation for our findings is that access to healthcare was lower for blacks versus whites and this might have precluded black men from obtaining routine medical care, particularly in the South, regardless of their underlying demand. However, a body of other economic research exploring black-white differences over a similar time period has provided evidence on increased access and a reduction in racial health disparities. 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\(\ 2014)\).\(^8\) Our results indicate that the Tuskegee revelation stalled the overall pattern of convergence for black men. Our within-black results comparing men to women directly addresses the concern that our strategy is picking up differential access or rollout to programs after 1972 for blacks living in the South. Our findings on health seeking and mortality are robust to estimation on a within-South sample, as are various transformations of the distance variable. Furthermore, adding high frequency county-level data on Medicare and Social Security expenditures does

---

\(^6\)The study focuses on the non-veteran population of men unless the adjective veteran is specified.

\(^7\)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 (Reverby 2000, p.xii)."

\(^8\)For an historical account of US black-white health disparities see Boustan and Margo (2014).
not meaningfully affect the results (see Appendix Table 7). Alternatively, there may have been other factors, such as high unemployment of unskilled labor, that could be correlated with proximity to Tuskegee as well as measures of health care utilization and outcomes and might disproportionally affect black men. We therefore condition on a rich set of individual and location-specific socioeconomic characteristics, including unemployment, in our estimates.\footnote{We thank Jacob Vigdor for this comment. Note that CPS only has longitudinal data for 11 states and approximately 19 MSAs dating back to the early 1970s. We therefore use unemployment from the 1970 census by county interacted with a time-trend. See Appendix Table 7.}

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 \cite{Baicker2004,Baicker2005}, we demonstrate that our findings are absent in comparisons between white men versus women and muted in comparisons between black women versus white women. In addition, we show that the coefficient magnitudes from the SEA including Tuskegee are larger than \(>90\%\) of placebo tests in the male subsample and \(>95\%\) of placebo tests on the black subsample using distance from any SEA in the dataset other than Tuskegee’s.

To test whether the behavioral responses we observe are driven by medical mistrust, we utilize survey data from the 1998 wave of the GSS on whether individuals trust a doctor’s judgment and whether they suspect that the medical establishment will deny them necessary treatment or services. We show that, unsurprisingly, the level of mistrust reflected in responses to each of those questions is greater for black men than for black women and for black men than white men. When we interact race- or gender-specific effect with the distance an individual was from Macon County, Alabama in 1972, we find the same geographic gradient apparent in our baseline utilization and mortality results. This increased mistrust by geographic proximity to Tuskegee is present even when conditioning on the current state of residence (in 1998) and the overall level of mistrust revealed in the respondent’s answers to other survey questions.\footnote{Although attrition is an issue in such a sample, given the findings on mortality described above, these estimates likely represent a lower bound.}

This paper builds on and contributes to several literatures in economics. First, our study is motivated by the theoretical contributions of \cite{Guiso2008} regarding the intergenerational transmission of beliefs and empirical work by \cite{Nunn2011} regarding the role of history (namely, the export slave trade) in shaping interpersonal mistrust in Africa. Our findings connect to a rich empirical evidence on the importance of trust for economic development \cite{Knack1997,Fafchamps2006} and firm management \cite{Bloom2008}. Finally, the research sheds light on questions in development economics
regarding low demand or uptake for products that have been proven to improve health (see review by Dupas (2011) and 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 puzzle.\textsuperscript{11}

The paper is organized as follows. In section 2 we provide historical background on the Tuskegee experiment and patterns of divergence and convergence in health across racial groups. In section 3 we describe the data and empirical framework, section 4 presents our results section 5 interprets the results using our model and concludes.\textsuperscript{12}

2 Literature Review and Historical Context

2.1 The Tuskegee Study and Medical Mistrust

The Tuskegee Study was designed to trace the course of untreated syphilis. The organism that causes the disease is closely related to that causing Lyme disease, and both bacterial diseases manifest themselves in stages. The first stage of sexually acquired syphilis is often an ulcer, followed by a full body rash that includes the palms and soles. However, it is the tertiary (or late-stage) syphilis that inflicts the most damage. The third stage is characterized by gummas (syphilitic tumors teaming with the bacteria) which coalesce and eat away at bone (frequently the nasal bridge) as well as other organs, and show a predilection for the arch of the aorta. Neurosyphilis (an attack on the nervous systems) presents in late-stage syphilis with paresis, gait disturbance, blindness and dementia (Mandell et al., 2009). 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."\textsuperscript{13} These incentives included physical exams, hot meals, and burial stipends that would be paid to their survivors. Most of the men also believed they were receiving some form of treatment. Approximately 600 black men (399 with syphilis)

\textsuperscript{11}This is also provided as a potential explanation for aforementioned low demand in Banerjee and Duflo (2012, p.58): "Faith or its secular equivalents, beliefs and theories, is clearly a very important part of how we all navigate the health system."

\textsuperscript{12}WE WANT TO CITE Chandra and Skinner (2004), NHDR (2003) http://www.ahrq.gov/qual/nhdr03/fullreport/, More recent papers that attempt to identify the casual mechanisms for these disparities include Chay/Guryan/Mazumder unpublished), Aizer, Lleras-Muney and Stabile (2005).

\textsuperscript{13}According to Thomas and Crouse p 409 "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."
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. "In addition to what they read in newspapers and magazines or heard on the radio and television, many blacks learned about the study by word-of-mouth" [Jones (1992, p.220)]. By the end of the study that year, only 74 of the test subjects were alive. Of the original 399 men with syphilis, 28 had died of syphilis, 100 were dead of related complications, 40 of their wives had been infected, and 19 of their children were born with congenital syphilis [Heintzelman (2003)]. Medical historian Allan Brandt (1978) 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 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." Further evidence of the role of mistrust in the health of black Americans comes from clinical psychology. Hammond et al. (2010) recruited 600 African-American men from around the country from local barber shops. Their purpose was to assess predictors of self-reported delays in medical care. The most important factor was medical mistrust—which doubled or tripled the odds ratio of delays in blood pressure and cholesterol screening. The authors were motivated to study medical mistrust as a potential predictor of preventive health seeking behavior because it "is higher among African-Americans" and "is linked to visible incidents of race-based malice towards this group (e.g. The Tuskegee Study of Untreated Syphilis in the Negro Male)."

Despite this, the empirical evidence linking Tuskegee to medical mistrust or health behaviors/outcomes is much more limited and does not attempt to establish causality. Corbie-Smith et al. (1999, p. 541) used

---

14 This is the full quote: , replete with the embellishments and distortions that usually attend oral traditions. Many blacks (and whites) were told that the government deliberately inoculated black sharecroppers with syphilis, while others were given to understand that the experiment was performed on black prisoners. Despite such errors, most black Americans got the gist of the story right. For forty years, their government had withheld treatment from men with syphilis so science could learn more about the disease. Many of the men died from syphilis while others had gone blind or insane. Confronted with the experiment’s moral bankruptcy, many blacks lost faith in the government and no longer believed health officials who spoke on matters of public concern." [p. 220 Jones]

15 Cite: http://exhibits.hsl.virginia.edu/badblood/report/

16 The most recent (and perhaps only) attempt to calculate the observational correlation between Tuskegee knowledge and
a discussion group setting to understand black men’s reluctance to participate in medical research. In the group, “participants made [unprompted] reference to the Tuskegee Syphilis Study.” Similarly, participants in the same study used the Tuskegee event to justify “their belief that ‘doctors value your life less than their own’” and that "African Americans still need to be suspicious when dealing with the medical establishment." To our knowledge, we are the first to relate racial disparities in healthcare utilization and health outcomes to this historical event.

2.2 Patterns of Healthcare Utilization and Mortality

The 1960’s marked an era of rapid convergence in healthcare access for black Americans relative to their white peers. Hospitals, initially racially separated or segregated according to designated white and black beds, gradually integrated over the 1950s and early 1960s, a process that culminated not with the Civil Rights Act of 1964, but with Lyndon Johnson’s insistence that any hospital receiving Medicare funding fully desegregate by July 1966. The process was quicker in the North than in the South; separate hospital wings were present in 75% of hospitals in the South as late as April of 1966 before full compliance was achieved three months subsequent (Reynolds 1997; Zheng and Zhou 2008).

Using data from The National Health Interview Survey (NHIS), among individuals ages 45 and 74, the percent of black women with less than a year interval to visit the doctor rose between 1963-64 and 1967-68 from 64% to 73%. The gains for black men were similarly impressive and rose from 57% to 65%. These increases of around 9 percentage points were over twice as large the increases experienced by white males and females over the same time period. The level of utilization for black (and white) men is lower than for women in every period, but the gap is one-third larger for those who have never served in the military.

The convergence of black and white healthcare utilization rates in the years prior to Tuskegee are closely mirrored by convergence in mortality rates. Annual county-level mortality statistics by age, race and gender are available from the CDC beginning in 1968. We plot the difference in age-specific mortality rates (ASMR) for blacks and whites by gender in Figure 1. Panels (A) and (B) demonstrates 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. Age-specific mortality rates generally exhibit a U-shaped pattern, highest in infancy and reaching a nadir in the data for young adults. The average ASMR for infants is roughly equivalent to medical mistrust (Brandon et al. 2005) has been roundly criticized for methodological failures (White 2005).

In the empirical strategy, both black women and white men are used as a comparator for black men.
that for adults 55 to 64 (around 19); and yet the pattern observed in Panels C and D is distinct from that presented in A and B. Unlike younger ages, where parental behavior (for early childhood) and trauma (for young adults) play key roles in determining mortality, for older adults, earlier mortality is driven partly by individual lifestyle choices and health-seeking behavior. At these ages, we observe a striking divergence in the mortality rates for black men versus white men following news of the Tuskegee disclosure (1972), which is not reflected in the difference between black and white women. An alternative way to have plotted these data would be to plot the differences in ASMR against median age at the time of the Tuskegee disclosure. From that perspective, year 1982 corresponds to those roughly 60 years old (in panel D) 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.
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: Data are from the NCHS/CDC and represent the black-white difference in the age-specific mortality rates. Each mortality rate is calculated dividing the number of the deaths in the relevant population and dividing by the at risk population. The blue line represents the difference for males, and the red line represents the difference for females. The vertical line (1972) represents the year of the Tuskegee disclosure. For additional graphs—including for South only; see Appendix Figures 1 and 2.
We repeat the exercise for the South (Appendix Figure 1) and for all other age groups (Appendix Figure 2). These additional mortality figures reinforce the notion that widening racial mortality disparities were particular to older black men and were present even within the South. Our estimation strategy will measure the health utilization and outcomes for black men relative to a counterfactual trend defined by either white males or black females in the years following the revelation of TSUS.

3 Theoretical Framework

As a framework for our empirical analysis, we provide a model of the interaction between prior beliefs, information about Tuskegee, and experience with the medical profession. In this regard, our model relates to work by [Bronnenberg et al. 2012, 2015] on how prior experience can influence beliefs and behavior and to belief formation models from XX AND YY. The model’s equilibrium provides testable predictions which will inform our empirical analysis.

We consider a multi-period model of belief formation where individuals engage in Bayesian updating regarding the trustworthiness of medical providers.\(^\text{18}\) Agents live for three periods (childhood, young adulthood, and old adulthood). They update their beliefs over two possible states of the world: the good state \((\theta = 1)\) where fraction \(q > \frac{1}{2}\) of the doctor’s are trustworthy, and the bad state \((\theta = 0)\) where the fraction is only \(1 - q'\). To simplify the exposition, we set \(q = q'\).

At \(t = 0\), agents are born with a prior belief over the distribution of the bad state: \(\pi^{\text{prior}}\). At \(t = 1\), some young adult agents are induced to visit the doctor. This inducement might come either from gynecological or obstetric care associated with being a young woman in her reproductive prime or from military enlistment for a young man. We assume such individuals, via this experience, receive a signal, \(S\), indicating the state of the world is \(\theta = 1\); the signal takes the same value as the state of the world it implies, so \(S = 1\). A negative signal (indicating that the state of the world is \(\theta = 0\)) would come in the form \(S = 0\). These agents then rationally update their beliefs regarding the state of the world.\(^\text{19}\)

Because our focus is on the behavior of older men, we derive predictions over the probability of seeking medical care in the last period, conditional on an individual’s experience at \(t = 1\). At \(t = 2\), both experienced

\(\text{\textsuperscript{18}}\)Fafchamps (2006, p. 1183) defines trust as “an optimistic expectation or belief regarding other agents’ behavior”. Fafchamps further discusses the origins of trust and uses this to distinguish two types: personalized trust which forms from “repeated interpersonal interactions” and generalized trust which has its origins in “general knowledge about the population of agents.” It is the latter we seek to model.

\(\text{\textsuperscript{19}}\)Such individuals cannot convincingly share their information with inexperienced agents. For the experienced group, their doctors visits are induced and therefore uninformative.
and unexperienced agents enter middle age, at which time those within a suitable range receive news of Tuskegee, \( S = 0 \), a "bad" signal indicating the state is \( \theta = 0 \). Those distant from Tuskegee, due to geographic or cultural isolation, do not receive this signal. Therefore, we can depict the posterior information structure as a 2x2 matrix based on proximity to Tuskegee:

<table>
<thead>
<tr>
<th></th>
<th>&quot;Near&quot; to Event</th>
<th>&quot;Far&quot; from Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experienced</td>
<td>( \Pr(\theta = 1</td>
<td>S = 1, S = 0, \pi_i^{prior}) )</td>
</tr>
<tr>
<td>Not-experienced</td>
<td>( \Pr(\theta = 1</td>
<td>S = 1, \pi_i^{prior}) )</td>
</tr>
</tbody>
</table>

Older agents must now decide whether to seek out medical care. Their payoffs are a function of their choice of action and the state of the world, and the cost is dependent only on their actions:

\[
u_i = u(a_i, \theta) = \begin{cases} 1 & \text{if } a_i = 1 & \theta = 1 \\ 0 & \text{if } a_i = 1 & \theta = 0 \end{cases} \text{ and } \begin{cases} 0 & \text{if } a_i = 0 \end{cases}
\]

\[c = c(a_i) = \begin{cases} c & \text{if } a_i = 1 \text{ where } c < q \\ 0 & \text{if } a_i = 0 \end{cases} \]

A rational agent seeks out medical care in old age if and only if \( \text{EU}(a_i = 1) > \text{EU}(a_i = 0) \equiv 0 \) where \( \text{EU} \) denotes expected utility. This inequality hinges on a threshold rule involving the posterior belief, and the agent seeks care if and only if:

\[
\Pr(\theta_1|\text{signals}, \pi_i^{prior}) > \Omega.
\]

where \( \Omega \equiv \frac{q+c-1}{2q-1} \). If agents share a common prior, \( \pi_i^{prior} \), the ordering of posterior beliefs in the matrix above is \( \Pr(\theta = 1|S = 1, \pi^{prior}) > \Pr(\theta = 1|S = 1, S = 0, \pi^{prior}) > \Pr(\theta = 1|S = 0, \pi^{prior}) \). Because these posterior beliefs map directly into actions, we can make empirical predictions about the behavior of individuals in older age.\(^{20}\)

In the aftermath of the Tuskegee disclosure, older adults will exhibit health-seeking behavior such that:

1. Experienced agents near to Tuskegee will exhibit lower health-seeking behavior rates than experienced agents in the same locations.

\(^{20}\)See appendix for proof. One can also show that the threshold posterior above which agents visit the doctor, \( \pi^* \) is declining in the benefits doctors provide and increasing in the cost of the visit. Generalizing the utility of visiting a good doctor from 1 to \( B \), : 

\[
\frac{\partial \pi^*}{\partial B} = \frac{-c}{2q-1} < 0 \text{ and } \frac{\partial \pi^*}{\partial c} = \frac{1}{B(2q-1)} > 0 \text{ for } q > \frac{1}{2}.
\]
2. Experienced agents near to Tuskegee will exhibit lower health-seeking behavior rates than experienced agents far from Tuskegee.

Further, if "nearness" to Tuskegee represents cultural, rather than geographic, proximity, the implication is that inexperienced black males will exhibit lower health-seeking behavior rates than their inexperienced white male peers.

4 Data

4.1 Health Seeking Behavior and Health Outcomes

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

Each individual in the survey self-reported the interval since their last doctor’s visit or physical, the number of doctor’s visits and other physician interactions (including phone calls) in the last 12 months, as well as the frequency and duration of hospital admission. The data also contain detailed demographic information, including family structure, income, education, and veteran status. From the surveys we have 165,864 respondents between 1969 and 1977. We lack morbidity data and rely on mortality data to assess whether the changes in beliefs and behaviors which we document below, translate into an effect on longevity.

We follow a large literature that uses mortality to assess the efficacy of health interventions. 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 that the number of deaths in a particular year is very low (and sometimes 0). Age-adjusted mortality rates based on a sparse number of events exhibits a large amount of random variation (Curtin and Klein, 1995, p.3) and the NCHS suggests aggregating over space or time if this small numbers are encountered. For this reason, we aggregate our annual county-level values to the state enumeration area (SEA) level. We report the SEA results in the main paper. To account for differences in age structure across regions and racial groups, we generate age-adjusted mortality rates. We follow the demographic literature and use the standard 1940 population as our reference. Since we anticipate the impact of Tuskegee to more slowly affect mortality than utilization, we utilize mortality rates through 1988 in our analysis.

### 4.2 Treatment Intensity

As described in the model above, as well as in other models of trust formation (Dixit, 2003; Tabellini, 2008) and in a vast psychological literature, individuals tend to be more affected by news if they can identify with the subject (Gutsell and Inzlicht, 2010; Singer et al., 2006). Tabellini models agents around a circle, then randomly couples them to play prisoner dilemma’s. 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." In this paper, we use geographic distance to Tuskegee to approximate who was likely to be affected by news of the study.

Although news of the study was national, geographic proximity correlates with the about the TSUS event through formal and informal news networks. It is also possible that the distance measure reflects the proximity of TSUS information to an individual’s own probable experience, for example if the pool of doctors serving black individuals in Alabama is fundamentally different from the pool elsewhere in the

---

25 See the data appendix for further details.

26 As an alternative measure of distance, Spolade and Wacziarg (2009) find that genetic differences predict income differences. The authors emphasize their measure of distance does not measure fitness, rather neutral genetic change, and therefore interpret their results as evidence that cultural relatedness facilitates diffusion of ideas and technology. "More closely related societies are more likely to learn from each other and adopt each other’s innovations.”

27 Note that linear distance from Tuskegee creates concentric circles which radiate out from the centroid; all points on a given circle have similar levels of exposure.
country. In this case, although information about Tuskegee may have been unrelated to geographic distance, *ex post* beliefs about the quality of the medical system may have been downgraded more severely in areas geographically closer to Macon County, Alabama.\textsuperscript{29}

\textsuperscript{29}We also have tried collecting data on local black newspapers and other publications. However, reliable circulation statistics are generally not available for many of these smaller publication outlets. Only a fraction of the black publications were members of the Audit Bureau of Circulations (ABC) and those which were tended to be larger national magazines (e.g. Jet and Ebony) which have national circulation. We searched for the mention of the Tuskegee study in www.newspaperarchive.com, a repository of archived newspaper articles, and assigned them to the state in which the newspaper was based for a state-level metric of media-based Tuskegee knowledge. This is not our preferred measure of media-based knowledge as, unlike black periodicals, it is not clear that the readership of these periodicals were predominantly black (indeed this database often excludes the black publications listed in Ayer and the information is often syndicated). Moreover, actual reporting could be more vigorous in places that were on a path of increasing mistrust of the medical profession, and thus is likely endogenous.
Figure 2: Geographic & Cultural Proximity

Panel A: Distance to Tuskegee (Macon County, Alabama)

Panel B: Fraction of Black Migrants from Alabama (1940)

A second candidate for measuring relatedness to the Tuskegee study disclosure is the fraction of black migrants from Alabama. We take advantage of the complete count version of the 1940 U.S. Census, which contains a question about 5-year migration patterns, to calculate this statistic.\textsuperscript{30} The migration patterns from 1935-1940 will be reflective of cultural proximity 30 years hence if individuals remain connected to their

\textsuperscript{30} Using later census years would require using state of birth as a proxy for geographic proximity to Macon County, as well as calculating the statistic from a much smaller sample of black Americans, a tradeoff we find too great.
locations of origin or if migration patterns are roughly stable over time and new migrants determine the cultural proximity of the broader population in their destination location \[\text{\#\#CITE MARIANNE}\].

5 Empirical Strategy

The goal of our empirical work is to identify the effect of Tuskegee on health-related beliefs, behavior and outcomes. Our hypothesis is that revelation of the Tuskegee syphilis study increased mistrust of the medical system for black Americans who identified most closely with the subjects and consequently reduced their propensity to engage with medical professionals. As predicted by our model, this impact is expected to be greater for groups that had limited exposure to the medical field prior to the disclosure (e.g. black non-veteran males).\(^\text{31}\) Furthermore, because the Tuskegee experiment enrolled only men, we predict that the effect will be greater for black males than for females.\(^\text{32}\)

To test our predictions, we employ a Differences-in-Differences-in-Differences (DDD) estimator with differences taken over time, race or gender, and proximity to Tuskegee, Alabama. We begin with an all-male sample and estimate the impact of TSUS on black males relative to white males, before and after 1972, and across a continuous measure of exposure to Tuskegee. Our identifying assumption requires there be no other systematic shocks to black men that affect the health outcomes of interest in geographic entities closer to Tuskegee that are correlated but not due to the timing of the study’s disclosure. As discussed in Section 2.2 on patterns of healthcare utilization, most policy changes coinciding with the timing of the disclosure were increasing access for disadvantaged populations and would likely bias our estimates towards the null. Although access to care and insurance coverage for black and white individuals differed in this period, any time-invariant differences in these factors are absorbed in the DDD design. 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. Nevertheless, our inclusion of location-year fixed effects nonparametrically controls for local-year specific shocks, such as the rollout of Medicare and Medicaid

\(^{31}\)CITE STATS on how many black women use reproductive services in 1970 and the draft lottery being in place up until the Vietnam war. Note that we are unable to observe veteran status in the mortality date but control for the number of veterans by county from CITE### in Appendix Table 7 TBD.

\(^{32}\)Our gender-specific predictions are further supported by public opinion survey data following other events which victimized black males. For example, the response of Americans to the events in Ferguson, MO in August, 2014 further justifies that black women were less likely to change their opinion of the medical profession in the wake of TSUS than black males. [CHECK OUT PEW DATA HERE.] On the other hand, if our gender-specific prediction is incorrect and the behavior of black women was equally affected by the news of TSUS, this will bias our results towards zero.
and their various expansions. These measures are not race-specific, but will also be included in regressions, described below, which are restricted to black and white women and to black veteran and non-veteran males. Extensions of our main estimating equation which more flexibly control for race-specific year and geography fixed effects are statistically indistinguishable from the baseline results (see Tables 5 and 6 column 4).

Although the assumptions of DDD estimator are fairly weak, an additional advantage of our framework is that it generates two natural falsification tests: white males versus white females and white females versus black females. Differences between white men and white women should have experienced no significant shift in the post-1972 period differentially by geographic distance to Macon County, Alabama if our proposed mechanism is driving the black male health disparities we observe. Similar, the disparity between black women and white women should be more muted than the same comparison for black and white men.

5.1 Main Estimating Equations

We denote distance to Macon County, AL by $D_k$ where $k \in s, a$ measured at the state or SEA level, respectively. For mortality data, we measure the distance from each SEA centroid to Macon County. To facilitate interpretation of regression coefficients, we multiply distance measures by $\frac{1}{1000}$ so that coefficients represent the effect of “proximity” (per 1000 kilometers) rather than distance. For NHIS outcomes, which are available at the state level, we utilize the county-level population of black adults and geographic county centroids to create population-weighted state centroids which are used to construct proximity as above.\textsuperscript{33}

Our data contain health behaviors and outcomes for both individuals (in the utilization data) and demographic groups (in the mortality data). Given the intensity proxy defined above, the relevant estimating equation for health behaviors or outcomes for individual $i$ of demographic group $g$ (e.g. black-male, black-female, white-male white-female) measured in state $s$ at time $t$ ($Y_{igst}$) is:

\begin{equation}
Y_{igst} = \alpha + \beta_1(D_s * post_t * group_g) + \beta_2(group_g * post_t) + \beta_3(D_s * group_g) + \gamma_g + \gamma_{st} + X_{igst} + \epsilon_{igst}.
\end{equation}

\textsuperscript{33}These centroids represent the average latitude and longitude of black individuals in each state based on the black population of counties (Haines, university Consortium for Political, and Research, Haines et al.).
while the estimating equation for demographic group \( g \) measured in SEA \( a \) at time \( t \) is:

\[
Y_{gat} = \alpha + \beta_1 (D_{a \cdot post} \cdot \text{group}_g) + \beta_2 (\text{group}_g \cdot \text{post}_t) + \beta_3 (D_{a \cdot \text{group}_g} \cdot \gamma_g + \gamma_{at} + \epsilon_{gat}).
\] (2)

In regressions at the individual level, \( X \) include 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, \( \beta_1 \) measures the differential impact of the Tuskegee disclosure on black men relative to white men in the years following 1972, per thousand miles of proximity to Macon County, relative to the difference between the two groups in the years prior for each outcome in question. When we test the intra-racial prediction that black men were more strongly affected by the news of TSUS than black women, we estimate the above equations on a sample of black individuals only and each \( \text{group}_g \) indicator refers to a binary variable for male. The regressions are otherwise unchanged.

Our analytical sample is limited to individuals 45 to 74 in order to facilitate comparisons between black men and women, the latter being more likely to seek healthcare in their reproductive years. This stipulation also avoids health-seeking behavior and mortality associated with violence in youth. We limit our analysis on health to a window around the disclosure (slightly longer for mortality effects to account for the lag between health behaviors and death). We do not attempt to model the carry-on effects of TSUS to future cohorts of black men given that our identifying assumption becomes more tenuous as one moves out in time. Still, this evidence suggests that Tuskegee ushered in an increase in racial health disparities specifically for older

---

34 In both cases, \( \gamma_g \) is an indicator for the race or sex of the individual or demographic group and \( \text{post}_t \) is a post-1972 year indicator. The model provides non-parametric controls for location-specific time effects that are common across demographic groups (\( \gamma_{at} \) and \( \gamma_{at} \)) such as the local rollout of federal and state-specific policies as well as controlling for time-varying demographic group shifts (\( \beta_2 \)) and differential effects of proximity to Tuskegee by demographic group (\( \beta_3 \)). The resulting coefficient of interest, \( \beta_1 \), represents the differential in post-1972 behavior by group and proximity to Tuskegee, Alabama, net of these other controls.

The model above yields estimates of \( \beta_1 \) that are statistically indistinguishable from alternative specifications that allow for full non-parametric control by replacing \( \text{group}_g \cdot \text{post}_t \) with group-year fixed effects, \( \gamma_{gt} \), and by replacing \( D_{a \cdot \text{group}_g} \) or \( D_{a \cdot \text{group}_g} \) with group-location fixed effects, \( \theta_{gs} \) or \( \theta_{ga} \) [Angrist and Pischke 2009, p. 242]. Specifically:

\[ Y_{igt} = \alpha + \beta_1 (D_{a \cdot \text{post}_t} \cdot \text{group}_g) + \gamma_{at} + \lambda_{gt} + \theta_{gs} + X_{igt} + \epsilon_{igt}. \] (3)

and

\[ Y_{gat} = \alpha + \beta_2 (D_{a \cdot \text{post}_t} \cdot \text{group}_g) + \gamma_{at} + \lambda_{gt} + \theta_{ga} + X_{gat} + \epsilon_{gat}. \] (4)

Estimates from the fully parametric versions of these equations can be found in Tables 5 and 6 column 4. We thank Martha Bailey for this suggestion.

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

36 Specifically mention AIDS here.
black men and might still be relevant today.37

6 Results

6.1 Event Study Estimates

We first present the results of event analysis specifications for 1 and 2 in which we generate a coefficient on $D \times group$ interacted with each year separately in lieu of estimating $\beta_1$. We continue to include $D \times group$ as a separate regressor, and we omit the interaction with the year of disclosure, 1972. All coefficients should be interpreted relative to that year.38 Figure 3 panel A thus presents 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. 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 affect all blacks. The pre-Tuskegee disclosure estimates are statistically indistinguishable from zero, but there is a statistically significant and sustained change beginning in 1972.39

Figure 3: Event Study on Number of Physician Visits

Panel A: Within Male

Panel B: Within Black

---

37 Appendix Table 2 presents difference-in-difference results on a sample of restricted to black men.
38 For the utilization data, we estimate: $Y_{ist} = \alpha + \sum_{n \neq 1972} (D_n \times I^n_{post}) + \beta_1 (group \times post) + \beta_2 (D \times group) + \gamma_g + \epsilon_{ist}$; and a similar equation for mortality data.
39 Event studies on the extensive margin of physician interactions can be found in Appendix Figure 3.
Notes: Data are from the NCHS/CDC NHIS Survey. The coefficient on the triple interaction of demographic group \(*\) post 1972 \(*\) proximity to Tuskegee is plotted as with 95% confidence interval. ** represents \(p<.05\).

In the next set of figures, we ascertain whether the differences in health-seeking behavior prompted by Tuskegee potentially translated into a widening racial gradient in mortality. Estimates from the comparable event study specification for the outcome of age-adjusted mortality is presented in Figure 4. We group the mortality data into two year bins on either side of 1972 to reduce noise in the estimates, though the yearly version can be found in Appendix Figure 4 and demonstrate a similar pattern. As anticipated, the change in health-seeking behavior translates into a stark increase in mortality rates for black men relative to both black women and white men in the years following 1972. This pattern is also apparent within the South, where the pre-1972 coefficients are again indistinguishably different from zero (Appendix Figure 5).

**Figure 4: Event Study on Mortality**

Panel A: Within Male

Panel B: Within Black

Notes: Data are from the NCHS/CDC compressed mortality files. The coefficient on the triple interaction of demographic group \(*\) post 1972 \(*\) proximity to Tuskegee is plotted as with 95% confidence interval. ** represents \(p<.05\).

### 6.2 Main Estimates

The event study coefficients 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 triple difference specification in equations 1 and 2. Estimates of \(\beta_1\) are identified
if places equidistant from Macon County Alabama would have had the same evolution of black-white health gradients, conditional on covariates, in the absence of the Tuskegee disclosure.
**TABLE 1: BASELINE RESULTS, UTILIZATION ADULTS**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Visit 12</th>
<th>No Visit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Male</td>
<td>Male</td>
<td>Black</td>
</tr>
<tr>
<td>D_c<em>post</em>black_{g}</td>
<td>-1.867***</td>
<td>-0.126</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.14)</td>
</tr>
<tr>
<td>D_c<em>post</em>male_{g}</td>
<td>-1.415***</td>
<td>-0.00736</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.09)</td>
</tr>
</tbody>
</table>

Controls: All Specifications Include Indicator Variables for Educational Status, Income Category, Age, Marital Status and Telephone Ownership as well as Urban Rural Status, the interaction between Distance Tuskegee and Black/Male and Post and Black/Male State-Year.

Fixed Effects

<table>
<thead>
<tr>
<th>Observations</th>
<th>60,837</th>
<th>18,966</th>
<th>150,864</th>
<th>192,735</th>
<th>60,837</th>
<th>18,966</th>
<th>150,864</th>
<th>192,735</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Clusters</td>
<td>49</td>
<td>44</td>
<td>49</td>
<td>49</td>
<td>49</td>
<td>44</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.022</td>
<td>0.046</td>
<td>0.016</td>
<td>0.014</td>
<td>0.029</td>
<td>0.058</td>
<td>0.018</td>
<td>0.022</td>
</tr>
</tbody>
</table>
Table 1 reports the outcomes for health seeking behavior on the intensive (columns 1-4) and extensive (columns 5-8) margins, respectively. Column 1 suggests that black men within 1000 kilometers of Tuskegee had at least 1.9 fewer physician visits relative to their white male peers in the period following the 1972 disclosure. This estimate is approximately 40% of the sample mean (4.4 interactions per 12 months). Similar, though slightly smaller reductions are observed when comparing black men to black women in column 2. Since approximately XX% of adult black men and women lived within 1000 km of Tuskegee in 1970, these point estimates include a sizable fraction of the affected population. Columns 5 and 6 repeat the exercise, this time for the outcome of whether the respondent reported no interactions with physicians in the last 12 months and demonstrate sharp increases, of about 4 percentage points, in the probability black men report no physician interaction in the last 12 months relative to either control group. These estimates represent approximately 13% of the sample mean. Columns 3, 4, 7 and 8, provide evidence that these results are specific to adult black men; the same patterns are not present for black adult women (relative to white women) or for white men relative to white women. For the latter two groups, the coefficient on the triple difference is statistically indistinguishable from zero for both sets of outcomes. Accordingly, any confounding factor that affects all black Americans or all men cannot be driving our results. Instead, threats to identification, which we discuss in more detail below, will be factors affecting only black men closer to Tuskegee that are correlated with changes in our outcome measures relative to white men or black females.

Our baseline mortality results are presented in Table 2 and represent estimates of $\beta_1$ from equation 2. As before, the first column represents a within-male comparison of the post-1972 difference between black and white men, relative to the pre-1972 difference and as a function of proximity to Macon County, Alabama. Mortality patterns, like utilization, appear to have moved adversely against blacks after 1972, and we estimate a 6.5% increase in the mortality rate of black men per 1000 miles of proximity in Panel B of the table. When we estimate the model using levels rather than log mortality (Panel A), this corresponds to an increase of 2.1 deaths per thousand miles, or a gradient of XX deaths per thousand from Macon County to Los Angeles, CA. Both results are statistically significant at the 0.025 level. In Column 2, we again compare black men to black women to ensure our mortality estimates are not simply picking up adverse results for blacks in general. The results here are even stronger; the mortality gap between black men and black women grew by 8.8% per 1000 miles of proximity in the years following 1972. In Panel A, we can translate this into an increase of 3.6 deaths per thousand per thousand miles, or a predicted YY additional deaths per
thousand in Macon County relative to Los Angeles. In the final two columns, we again find no significant difference in black females relative to white females, but a positive result for white men relative to white women. In addition to being small in magnitude (an increase of 1.0%), this latter result is not robust to a level specification.

<table>
<thead>
<tr>
<th>TABLE 2: BASELINE RESULTS, MORTALITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>PANEL A --- MORTALITY PER THOUSAND</td>
</tr>
<tr>
<td>(D_{c}*post_{t}*black_{g})</td>
</tr>
<tr>
<td>(D_{c}*post_{t}*male_{g})</td>
</tr>
<tr>
<td>Fixed Effects</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
<tr>
<td>PANEL B --- LOG MORTALITY PER THOUSAND</td>
</tr>
<tr>
<td>(D_{c}*post_{t}*black_{g})</td>
</tr>
<tr>
<td>(D_{c}*post_{t}*male_{g})</td>
</tr>
<tr>
<td>Fixed Effects</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>R-squared</td>
</tr>
</tbody>
</table>

In Appendix Tables 3-5 we assess other outcome variables including whether or not the individual was hospitalized (extensive margin), how many nights the respondent was hospitalized (intensive margin) and dental visits. Appendix Table 3 demonstrates that black men were less likely to be hospitalized, however column 1 in Appendix Table 4 shows that, conditional on hospitalization, older black men have a longer hospital stay, providing suggestive evidence that their reason for admission was more advanced or serious at the time of presentation to the hospital. Appendix Table 5 (dental visits over the last 12 months) is available for a slightly truncated period of time (starting in 1970 ##CONFIRM); and shows black men are slightly more likely to visit the dentist. Though this result is marginally significant—it provides suggestive evidence that black men did demand some forms of healthcare more than their white peers, and their relative avoidance was specific to institutions most reminiscent of the Tuskegee study.\(^{40}\)

\(^{40}\)Appendix Table 6 examines heterogenous effects of the effect of Tuskegee on health care utilization by urbanicity and self-reported health status.
6.3 Threats to Identification

We perform several tests to assure a causal interpretation for our results. These tests fall into one of several categories: first, we use placebo outcomes (in addition to within female and within white samples as shown above) demonstrating that our results do not obtain on different samples of black males; second, we demonstrate our results are robust to slight modification of our treatment variable as well as alternative measures of "proximity" that exploit black migration patterns; third we demonstrate that our results obtain when restricting to within the South and when absorbing more variation with fixed effects for black(or male)*year and black(or male)*state; and fourth, we show that similar results are not found when using the universe of alternative placebo geographic distances. We describe each of these tests in turn.

6.3.1 Placebo Outcomes

Here we show that our results are only found among older black male adults. Differences between black and white males or black males and females are not found among children (Tables 3). This is likely because healthcare utilization decisions for children at this time were mostly driven by their mothers’ preferences.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Visit 12</th>
<th>No Visit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>Black</td>
</tr>
<tr>
<td>D_c,post,black_g</td>
<td>0.0179</td>
<td>0.127</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>D_c,post,male_g</td>
<td>-0.117</td>
<td>0.0338</td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.04)</td>
</tr>
</tbody>
</table>

As with the utilization data, we estimate $\beta_1$ from equation $\mathcal{E}$ for younger age groups, specifically infants, children aged 1-4 and children aged 5-14. The resulting estimates of $\beta_1$, located in Table 4 and indicate no

---

41 These categories reflect the mortality and population intervals available in the underlying data.
impact on the mortality levels of black male children relative to either white male children or black female children. (We estimate these regressions in levels rather than logs as there are many zero observations in youth mortality statistics.) Results in Column 1 for infant mortality, on the other hand, show no change in the relative mortality of black and white male infants but a declining rate of black male relative to black female infants relative to 1972. The economic significance of this coefficient is relatively small (representing an infant mortality reduction of 1.9%), and its sign is opposite that for adult men.

TABLE 4: BASELINE RESULTS, CHILDREN'S MORTALITY

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Placebo Infant Mortality</td>
<td>Placebo Young Children (1-4)</td>
<td>Placebo Young Children (5-14)</td>
</tr>
<tr>
<td>Sample</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
</tr>
<tr>
<td>Dc<em>postt</em>blackg</td>
<td>-0.259</td>
<td>3.19E-03</td>
<td>4.22E-04</td>
</tr>
<tr>
<td>(0.198)</td>
<td>(2.38E-03)</td>
<td>(4.53E-03)</td>
<td></td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>SEA-Year</td>
<td>SEA-Year</td>
<td>SEA-Year</td>
</tr>
<tr>
<td>Observations</td>
<td>10274</td>
<td>10260</td>
<td>10254</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.76</td>
<td>0.01</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Black</td>
<td>Black</td>
<td>Black</td>
</tr>
<tr>
<td>Dc<em>postt</em>maleg</td>
<td>-0.0618***</td>
<td>7.52E-04</td>
<td>1.34E-03</td>
</tr>
<tr>
<td>(0.0171)</td>
<td>(3.54E-03)</td>
<td>(4.59E-03)</td>
<td></td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>SEA-Year</td>
<td>SEA-Year</td>
<td>SEA-Year</td>
</tr>
<tr>
<td>Observations</td>
<td>10274</td>
<td>10254</td>
<td>10222</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.97</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Appendix Table 2 includes double-difference results on a sample of older black men. These results demonstrate that even within the sample of older black men, it is those closer to Tuskegee that are affected in both utilization and mortality. These results do not obtain for black women, white men or white women.

6.3.2 Treatment Variable

Next, we demonstrate that our findings are robust to different constructs of the treatment variable that capture cultural and geographic proximity. For this next set of regressions, we place the within male comparisons
in Panel A and the within Black comparisons in Panel B. Table 5 reports the utilization results robustness checks and Table 6 reports similar results for mortality.

<table>
<thead>
<tr>
<th>TABLE 5: ROBUSTNESS CHECKS, UTILIZATION</th>
<th>Outcome Visit 12</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Bins</td>
<td>Migrants</td>
</tr>
</tbody>
</table>

**PANEL A --- Male Samples**

<p>| | Dc<em>postt</em>blackg | -1.226*** | -1.798 | -1.800*** |</p>
<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Placebo Distance (Dallas)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State-Year</td>
<td>60,837</td>
</tr>
<tr>
<td>State-Year</td>
<td>30,072</td>
</tr>
<tr>
<td>State-Year</td>
<td>18,574</td>
</tr>
<tr>
<td>State-Year</td>
<td>60,837</td>
</tr>
<tr>
<td>State-Year</td>
<td>60,837</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.023</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.024</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.028</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.024</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.022</td>
</tr>
<tr>
<td>p-value</td>
<td>0.023</td>
</tr>
<tr>
<td>p-value</td>
<td>0.024</td>
</tr>
<tr>
<td>p-value</td>
<td>0.028</td>
</tr>
<tr>
<td>p-value</td>
<td>0.024</td>
</tr>
<tr>
<td>p-value</td>
<td>0.022</td>
</tr>
</tbody>
</table>

**PANEL B --- Black Sample**

<p>| | Dc<em>postt</em>blackg | -0.833** | -1.455 | -1.389*** |</p>
<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Placebo Distance (Dallas)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State-Year</td>
<td>18,966</td>
</tr>
<tr>
<td>State-Year</td>
<td>13,918</td>
</tr>
<tr>
<td>State-Year</td>
<td>10,587</td>
</tr>
<tr>
<td>State-Year</td>
<td>18,966</td>
</tr>
<tr>
<td>State-Year</td>
<td>18,966</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.046</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.039</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.040</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.049</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.046</td>
</tr>
<tr>
<td>p-value</td>
<td>0.046</td>
</tr>
<tr>
<td>p-value</td>
<td>0.049</td>
</tr>
<tr>
<td>p-value</td>
<td>0.046</td>
</tr>
<tr>
<td>p-value</td>
<td>0.046</td>
</tr>
</tbody>
</table>
In Column 1 of Tables 5 and 6, we replace our fine distance measures with distance bins, defined as quintiles of distance from Tuskegee. In Column 2, we utilize a different proxy for cultural proximity to the disclosure—the number of black migrants from Alabama (Table 5 and 6 columns 2). Column 3 is restricted to the South. If our results are driven by something other than the revelation of TSUS, that something must also be correlated with geographic proximity to this particular location in the South and not with the South in general. In Column 4, we add in fixed effects for black-location (state or SEA) and black-year in the within-male specifications (Panel A) and male-location (state or SEA) and male-year in the within-black specification. In Panel A specification, the black-location fixed effects absorb any local geographic shocks that affected black but not white men. In the within-black specification, the corresponding fixed effects absorb any location-specific shocks that affect this group. These fixed effect specifications are useful as they non-parametrically control for the rollout of government programs which

<table>
<thead>
<tr>
<th>Sample</th>
<th>Male</th>
<th>Male</th>
<th>Male</th>
<th>Male</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_{post,black} &amp; 0.931** (0.431) &amp; 6.291** (2.85) &amp; 1.288* (0.700) &amp; 2.013** (0.948) &amp; 1.900* (1.016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_{placebo,post,black} &amp; 0.727 (0.892)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fixed Effects
- SEA-Year
- SEA-Year
- SEA-Year
- SEA-Year Race-SEA
- SEA-Year Race-Year

Observations
- 10100
- 9794
- 4,092
- 10076
- 10078

R-squared
- 0.17
- 0.17
- 0.56
- 0.13
- 0.17

p-value

<table>
<thead>
<tr>
<th>Sample</th>
<th>Black</th>
<th>Black</th>
<th>Black</th>
<th>Black</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_{post,male} &amp; 1.511*** (0.318) &amp; 11.183*** (3.759) &amp; 2.290** (0.954) &amp; 3.349** (1.183) &amp; 3.489*** (1.306)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D_{placebo,post,male} &amp; 0.348 (1.034)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fixed Effects
- SEA-Year
- SEA-Year
- SEA-Year
- SEA-Year Male-SEA
- SEA-Year Male-Year

Observations
- 9990
- 9696
- 4,092
- 10296
- 9968

R-squared
- 0.23
- 0.23
- 0.66
- 0.16
- 0.231
might have differentially enrolled black individuals. In the last column of both tables, we utilize a distance from a placebo location: Dallas, Texas which is in the same latitudinal band as Tuskegee, Alabama and fail to reproduce statistically significant results. We expand on this exercise in Figure 5, where we calculate the distance from the centroid of every SEA in our dataset and use this as the placebo location of the syphilis study, plotting the resulting $\beta_1$ coefficient. Our results indicate that the Macon County triple coefficient is greater than 91% (97%) of all potential distances for the sample of men (blacks).

Figure 5: Placebo Distances (Mortality)

Panel A: Within Male

Panel B: Within Black

Finally, the geographic mobility of individuals may serve to complicate the interpretation of $\theta_1$ and $\beta_1$. We cannot identify where individuals were at the time of exposure, but we note that mobility was relatively limited in this period (which followed the Great Migration) and was generally short-distance. Still, it would be preferable to observe outcomes that are correlated with a fixed geographic proximity to Tuskegee, Alabama in 1972 and not in the years following. In Section 5.3 below, we utilize the only data we have with this information (GSS data on attitudes and perceptions) and regress perceptions of physicians’ judgement in 1996 on the location of an individual at the time of TSUS revelation in 1972. We show that the

---

30
proximity to Macon County, Alabama in 1972 is correlated with physician mistrust even after controlling for location in 1996. In other words, for a sample where we can control for migration, we observe heightened mistrust of physicians that correlates with the proximity measure $D_k$ above.

### 6.4 Channels

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

$$Mistrust_{MD} = \alpha + \theta_1(D_{16}^{16} \times black_i) + \theta_2 black_i + \theta_3 D_{16}^{16} + \pi_s + \gamma_s X_{i gs} + \epsilon_{i gs}$$ (5)

where $D_{16}^{16}$ is the distance from the state of the respondent at age 16 to Macon County. The sample includes individuals at least 10 years of age at the time of the disclosure.\(^{43}\) $X_{i gs}$ contains indicator variables for individual’s age, marital status, urbanization and level of education as well as an indicator for their current state of residence so that $\theta_1$ is identified conditional on current location.\(^{44}\) In addition, we condition on an individual’s general level of mistrust in others to isolate the impact of medical mistrust. 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 at age 16).\(^{45}\) The results are presented in Table 7 and demonstrate that black men compared to white men and black women, are more likely to

\(^{43}\)The sample consists of never movers, those whose state when they were 16 is the same as their state today- and those who have moved, but we know their location when they were 16. For the former individuals, we include all those at least 10 years of age at the time of the disclosure. For the latter, we include those in a narrow age band around 16 years old (e.g 10-22).

\(^{44}\)Adding respondent income to this regression reduces the sample size considerably (for example, from 260 to 186 in column 1) though the magnitude of the results for men is slightly larger (as are the standard errors)—0.707 (s.e. 0.368, p=.063).

\(^{45}\)We can identify location of residence in 1972 for two groups of individuals in the GSS data. (1) The survey asks for location of residence at age 16, thereby allowing us to identify location of residence for individuals who are xx-yy in 1998. (2) The survey also inquires about state of birth, and we assume individuals who have the same state of birth and state of current residence are never-movers so that their location in 1972 is also known.
disagree with the statement that doctors' judgement can be trusted and agree with the statement that doctors deny needed treatment. Though both of the estimates are only statistically significant for the within male comparison, the within black sample is limited by a much smaller size.

**TABLE 7: EFFECT OF TUSKEGEE ON BELIEFS**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Medical Mistrust</th>
<th>Deny Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td><strong>Male</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Black</strong></td>
<td>0.469***</td>
<td>-0.173</td>
</tr>
<tr>
<td><strong>Female</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>White</strong></td>
<td>(0.134)</td>
<td>(0.198)</td>
</tr>
<tr>
<td><strong>Fixed Effects</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Other Controls</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>260</td>
<td>353</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.42</td>
<td>0.29</td>
</tr>
<tr>
<td><strong>No Clusters</strong></td>
<td>37</td>
<td>26</td>
</tr>
</tbody>
</table>
7 Conclusion

Contemporary and historic disparities are certainly multifactorial, reflecting differences in socioeconomic status, geography, access, insurance coverage, culture, health-related behaviors, and other persistent differences. This paper examined how a historical event, the Tuskegee Study disclosure, contributed to medical mistrust and health-seeking behavior and outcomes. ##NEED STATISTIC HERE ON HOW MUCH OF RACIAL MORTALITY DISPARITY WE EXPLAIN.

Our findings have implications for development and behavioral economics which are often concerned with questions of low take-up for products with proven health benefit despite subsidized prices and intense social marketing (cite Poor Economics, Dupas, others). Our results suggest that efforts to understand the cultural-specific evolution of health perceptions and their influence on behaviors may account for some of these puzzles.
References


38