

NBER WORKING PAPER SERIES

THE GIFT OF A LIFETIME:
THE HOSPITAL, MODERN MEDICINE, AND MORTALITY

Alex Hollingsworth
Krzysztof Karbownik
Melissa A. Thomasson
Anthony Wray

Working Paper 30663
<http://www.nber.org/papers/w30663>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 2022, Revised January 2024

We thank the editor and four anonymous referees whose comments have greatly improved the paper. We also thank Brian Beach, Brandyn Churchill, Jane Greve, Maggie Jones, Carl Kitchens, Peter Nencka, Greg Niemesh, and seminar/conference participants at the Atlanta-Athens Health Economists Research Conference, the Danish Society for Economic and Society History and Scandinavian Society of Economic and Social History annual meeting, the Essen Health Conference, the National Bureau of Economic Research (NBER) Health Economics Fall Meeting, the NBER Summer Institute Development of American Economy and Children's sessions, Duke University, Emory University, Indiana University, The Insper Institute of Education and Research, Lund University, Marquette University, Miami University, Ohio State University, and the University of Copenhagen for helpful comments and feedback. We are grateful to Patrick Carlin and Mallory Dreyer for research assistance and Hardish Bindra at Paradigm Data Services. We would like to especially thank John Parman for sharing the North Carolina death certificate data. Thomasson acknowledges support from the Julian Lange Professorship. Wray appreciates financial support from the Japan Society for the Promotion of Science (KAKENHI Young Scientists B Grant Number J160100115), research funds from the Hitotsubashi Institute for Advanced Study (HIAS) at Hitotsubashi University, and an Arthur Cole Grant from the Economic History Association. 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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2022 by Alex Hollingsworth, Krzysztof Karbownik, Melissa A. Thomasson, and Anthony Wray. 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.

The Gift of a Lifetime: The Hospital, Modern Medicine, and Mortality
Alex Hollingsworth, Krzysztof Karbownik, Melissa A. Thomasson, and Anthony Wray
NBER Working Paper No. 30663
November 2022, Revised January 2024
JEL No. I14,J13,N32

ABSTRACT

We explore how access to modern hospitals and medicine affects mortality by leveraging efforts of The Duke Endowment to modernize hospitals in the early-twentieth century. The Endowment helped communities build and expand hospitals, obtain state-of-the-art medical technology, attract qualified medical personnel, and refine management practices. We find that Duke support increased the size and quality of the medical sector, fostering growth in not-for-profit hospitals and high-quality physicians. Duke funding reduced both infant mortality - with larger effects for Black infants than White infants - and long-run mortality. Finally, we find that communities aided by Duke benefited more from medical innovations.

Alex Hollingsworth
Department of Agricultural,
Environmental, and Development
Economics, Department of
Economics, and the John Glenn
College of Public Affairs
The Ohio State University 2120
Fyffe Road
Columbus, OH 43210
and NBER
hollingsworth.126@osu.edu

Krzysztof Karbownik
Department of Economics
Emory University
Rich Memorial Building
1602 Fishburne Drive
Atlanta, GA 30322
and NBER
krzysztof.karbownik@emory.edu

Melissa A. Thomasson
Miami University/FSB
Department of Economics MSC 1035
800 E. High Street, Rm. #2054
Oxford, OH 45056
and NBER
thomasma@miamioh.edu

Anthony Wray
Department of Economics
University of Southern Denmark
Campusvej 55 Odense M DK-5230
Denmark
wray@sam.sdu.dk

As medicine has advanced over the last century, spending on healthcare has risen to consume an ever-increasing share of resources. In 1900, healthcare represented only 2% of US gross domestic product (Craig 2006). By 2020, this share soared to 20%, with spending on hospital care alone accounting for 6% of US economic activity (CMS 2022). This past century has also marked the rise of modern medicine and hospitals, transforming the institution “from places of dreaded impurity and exiled human wreckage into awesome citadels of science and bureaucratic order” (Starr 2017). Building on gains over the mid- to late-nineteenth century – attributable to rising incomes and better nutrition – health has improved dramatically over the twentieth century, with infant mortality rates falling by 95% and life expectancy at birth rising by 45% (Costa 2015). A large body of work has explored the drivers and the timing of this health transition. One claim from this literature is that modern medicine – with the exception of sulfa drugs – did not begin to materially affect health and mortality until after the 1950s (Cutler 2005). Despite this viewpoint, little is known about the origins of the modern hospital or the contribution of medical care in improving health in the first half of the twentieth century.

We seek to understand the role of funding for formal health care in improving health outcomes. Did financial resources help to expand hospitals and attract higher-quality physicians with modern medical training? Did these improvements to medical care contribute to declining mortality during infancy and in the long run? Given the many racial inequities in society and in health, were these effects similar or different by race? And finally, did the modernization of hospitals allow for more productive use of new medical technologies?

We leverage a combination of novel data and a unique quasi-experiment: a large-scale hospital modernization program introduced by The Duke Endowment in North Carolina during the first half of the twentieth century. Beginning in 1927, The Duke Endowment funded capital projects to upgrade hospitals, thereby expanding access to modern medicine and hospital care across the Carolinas. It did so by placing a particular emphasis on rural and Black residents – an uncommon approach in the Jim Crow-Era South. The Endowment assisted hospitals in expanding and improving existing facilities, obtaining state-of-the art medical technology, and elevating their management practices. In some communities, the Endowment helped to build new hospitals or to convert existing facilities to non-profits. These modernization efforts were explicitly designed to attract higher-quality physicians to the Carolinas (The Duke Endowment 1934; Durdin 1998). By the end of 1942, the Endowment had appropriated over \$3.3 million (\$53 million in 2017 dollars) for hospital expansion and modernization in North Carolina, an investment that significantly improved access to medical facilities and the quality of the medical infrastructure, thereby propelling the practice of medicine in supported communities towards the modern era.¹

¹ While The Duke Endowment operates in both North and South Carolina, we focus our analysis on North Carolina due to data availability as we discuss in Section 2.

Our main empirical approach is a difference-in-differences design that exploits quasi-random variation in Duke support, comparing counties that received a funding commitment from the Endowment to counties without such a commitment, and cohorts born before the initial commitment to those born after the introduction of Duke support. In this effort, we construct a novel data set that combines information from hand-collected archival records, annual reports of The Duke Endowment, individual death certificates, physician directories, and a variety of other sources. We opt for an extensive margin, intent-to-treat specification as our preferred model because treatment was not solely monetary. To convert an appropriation into a payment, the Endowment typically required hospitals to undertake additional steps, such as changing accounting and record keeping practices; ensuring that surgeries were conducted by licensed physicians; allowing de facto inspections by Endowment employees; or forging community partnerships. An intent-to-treat analysis captures the full effect of this multifaceted treatment, allowing for Duke support to both impact health before a single dollar was received and facilitate changes that improved health even in hospitals that never received a payment.

We find that the efforts by Duke to modernize hospitals between 1927 and 1942 caused increases in hospital capacity and the quality of physicians. Duke support increased the number of hospital beds per 1,000 births in not-for-profit hospitals that were eligible for funding by 70.1%, while causing a decline in ineligible for-profit beds of 61.4%. Moreover, supported communities saw an increase in the number of high-quality physicians per 1,000 births of 60.2% and a decrease in the number of poorly-trained physicians of 5.5%, thereby advancing the average state of medical knowledge in locations supported by Duke.

These first-stage effects and mechanisms help explain our core result that Duke financial support for the modernization of hospitals improved health, causing a 7.5% reduction in the infant mortality rate. Treatment effect estimates for Black infants (13.6%) are about three times as large as those for White infants (4.7%). Those exposed to Duke support around the time of birth also enjoyed lasting health effects, as their long-run mortality rates (between the ages 56 and 64) declined by 9.0%. Finally, we document that Duke-supported hospitals were able to use new medicines more effectively, suggesting a complementarity between high-quality hospitals and new medical innovations.

Our findings are robust. Point estimates and statistical significance are not materially affected by the choice of estimator, specification, transformation of the dependent variable, or sample decisions. Our findings are also insensitive to the use of recent techniques that are robust to treatment effect heterogeneity and staggered treatment adoption in a difference-in-differences setting. We further demonstrate that our results are not driven by the two primary threats to identification, differential pre-trends and selection into treatment. Across a number of event study specifications, outcomes in treated and control counties follow parallel trends prior to treatment. Using mortality

data from other Southern states, we further confirm that the results are not driven by selection into treatment, by spillovers, or by the reallocation of medical expertise and investments within the Carolinas.

The program was cost effective. Considering only infant mortality from 1927 to 1942, one life was saved for every \$17,092 (2017 dollars) paid to hospitals, a cost significantly lower than any reasonable estimate for a Value of Statistical Life (VSL). The intervention was also particularly effective in reducing Black infant mortality, resulting in a mortality rate decrease much larger than the estimated decrease for White infants (a reduction of 15.1 vs. 3.1 deaths per 1,000 births) and narrowing the Black-White infant mortality gap by one-quarter. This new evidence that access to high-quality medical care and infrastructure disproportionately reduced Black mortality rates adds to previous work on the Black-White infant mortality gap, some of which finds that it has no relationship with increased healthcare access (Collins and Thomasson 2004; Elder et al. 2016; Anderson et al. 2021a,b).²

This paper makes contributions to several literatures. First, our findings contribute to a series of articles investigating the drivers of the health improvements observed in the early- to mid-twentieth century. Most prior work focuses on the role of public health efforts (e.g., Anderson et al. 2022) or the advent of sulfa drugs (e.g., Thomasson and Treber 2008; Jayachandran et al. 2010). Studies that examine the effects of hospitals or medical care on mortality have primarily concentrated on later time periods (Cutler 2005). For example, Figinski and Troland (2020) show that a rural hospital construction program in Appalachia in the 1950s improved infant health. Our results indicate that even earlier improvements in hospital construction and modernization reduced mortality, likely by attracting better-educated physicians to the area. Thus our findings suggest that improvements in medical care may have played a greater role in reducing mortality in the first half of the twentieth century than has been acknowledged.

Second, we build on the broader literature examining the short- and long-run consequences of prenatal and infant health. We also complement studies examining the determinants of mortality and longevity, in which a common theme is that early life circumstances matter for later-life health (Almond et al. 2018). Despite this large literature, ours is the first study to leverage quasi-experimental variation to show both the short- and long-run mortality consequences of a large hospital-funding program, and to our knowledge it is the first to examine whether hospital improvements in this time period helped to ameliorate long-standing racial gaps in mortality.

²The Duke Endowment’s Indenture of Trust, written in the 1920s, explicitly stated that hospital access should be provided for *both* Black and White individuals. Although Duke-supported hospitals remained segregated and care may not have been equal across race within the same hospital, the Endowment extended hospital access to Black residents in ways that were far from cosmetic. Before the Endowment’s bi-racial mandate, many Black residents had no access to hospital care, while afterwards “all or practically all of the 65 hospital construction projects that the Endowment assisted during its first decade contained accommodations for both races” (Durdin 1998).

A unique feature of our work is that we do not solely focus on applying a well-established intervention to a different vulnerable population. Instead, we analyze a bold experiment aimed at modernizing medical care during a critical phase in its development.

Third, we contribute to the health care literature by demonstrating that high-quality facilities can have complementary effects. We show that improving facilities attracted physicians who graduated from better medical schools. Such an outcome was a clear goal of The Duke Endowment, as the first director of the Endowment’s Hospital Section, Dr. Watson S. Rankin, stated that “[t]he modern hospital... had come to play an essential part not only in the practice of medicine but also in the distribution of medical personnel” and that creating well-supported hospitals would be “the main remedy for the shortage of doctors” (Durdin 1998). Our findings support this view.

An additional complementarity that we examine is the relationship between technological innovation and improvements in physical and human capital, as discussed in Alsan and Goldin (2019). Specifically, we analyze the effect of the introduction of sulfa drugs in 1937, which resulted in significant reductions in infectious disease mortality (Thomasson and Treber 2008; Jayachandran et al. 2010; Lazuka 2020). We demonstrate that Duke’s emphasis on improving healthcare infrastructure and the quality of the medical workforce enabled communities to better leverage sulfa drugs, highlighting that the diffusion of knowledge is necessary for the effective spread and use of medical technologies (Soares 2007). Our findings reveal that in communities with higher baseline infectious disease mortality rates, the impact of The Duke Endowment’s support on infant mortality approximately *tripled* following the advent of sulfa drugs.

Fourth, we contribute to literature on management style, standardization, and firm performance (Bloom and Van Reenen 2007). A distinct aspect of Duke support is that it was a multifaceted bundle of treatments coupled with ample oversight that was not limited to a monetary component. In addition to providing funding for facilities and equipment, the Endowment altered core management practices and standardized hospital accounting procedures.³ It thereby “exerted considerable influence over the direction of hospital care” in supported communities (Gamble 1995). Thus, it is unlikely that monetary assistance alone would have had the same effect. It is possible that the Duke intervention was so successful *because* these many changes served to reinforce one another.

Finally, our work connects with the literature examining the effects of charitable contributions on public health and welfare, including the hookworm eradication campaign sponsored by the Rockefeller Sanitary Commission (Bleakley 2007), and the Rosenwald Rural Schools Initiative (Aaronson and Mazumder 2011; Aaronson et al. 2014). For many years, The Duke Endowment

³Summarizing a letter written by Graham Davis, one-time assistant director of the Endowment, Durdin (1998) notes that “the money they received from the Endowment was probably not as valuable to [the assisted hospitals] as the free service they got from the staff” in the form of consulting, oversight, management expertise, and assistance in converting to uniform record keeping, of whom Duke-supported facilities “were the largest group in the nation employing.”

was the third largest charity in the US and it remains one of the largest charitable foundations in the world (Durdin 1998). Despite its size and success, to our knowledge, virtually no others (outside of historians documenting Endowment activities) have attempted to evaluate the effects of the many programs that it operated.

We view our results as having three important modern policy implications. First, the success of the Endowment's efforts to support hospitals was so pronounced that it "served as both inspiration and model" for the Hill-Burton Act of 1946 (Durdin 1998), a large program designed to stimulate hospital construction and access across the US (Chung et al. 2017). Second, its impact in providing access to the first high-quality medical care facility in a community highlights the importance of attracting physicians to practice in under-served communities and the potential costs of closing rural hospitals today, as removing the last high-quality healthcare option may have a symmetric negative impact (Kozhimannil et al. 2018; Germack et al. 2019; Fischer et al. 2022). Third, our findings are relevant to healthcare investments in developing countries where donations from both private foundations and governments represent a major source of health financing (Yamey et al. 2019).

1 The Duke Endowment and Hospital Care in North Carolina

On December 11, 1924, James Buchanan Duke established The Duke Endowment with initial funding of \$40 million (\$563 million in 2017 dollars) to improve the social and economic welfare of communities in North and South Carolina. In the Indenture of Trust for the Endowment, Duke acknowledged the vital role of hospitals as "indispensable institutions" for "prolonging human life," noting that "[t]he advance in the science of medicine" has made "hospital facilities essential for obtaining the best results in the practice of medicine" (The Duke Endowment 1925). Duke directed that 25.6% of the income from the invested Endowment funds to be apportioned to the modernization, maintenance, and construction of not-for-profit hospitals (The Duke Endowment 1925, p. 10-11). Duke died unexpectedly in October 1925. His will left the Endowment with an additional \$69 million (\$972 million in 2017 dollars), with 90% of the annual income from this contribution specifically directed to hospitals in the Carolinas (Durdin 1998, p. 26).⁴

Another one of Duke's stated objectives in improving hospital facilities was to attract more highly-qualified doctors. In its inaugural report, the Hospital Section noted, "Hospitals hold and attract the better type of physicians... So it is that a local hospital builds up a profession, raises professional standards and serves to improve the practice of medicine not only for patients in the hospital but for patients in the whole county" (The Duke Endowment 1925, 144-45). This objective was realized in North Carolina, where physician density increased between 1923 and 1940 from

⁴Separately in his will, Duke earmarked \$4 million (\$56 million in 2017 dollars) to "build and equip a medical school, hospital, and nurses' home at Duke University" (Durdin 1998, p. 26).

26 to 56 per 1,000 population. The share of births in hospitals likewise increased, rising between 1931 and 1942 from 8.9% to 34.2%, but remained low even compared to the average of Southern states (13.2 % in 1931 and 42% in 1942).

Following Duke’s death and prior to disbursement of any funds from the Endowment, the foundation surveyed the healthcare landscape in the Carolinas with a focus on the quality, size, and density of hospitals, as well as the supply of physicians. Armed with this knowledge, the foundation started appropriating funds in 1927 – the first year we consider a county to have been treated. Panel A of Figure 1 presents the share of North Carolina counties that received Duke appropriations by a specific year. No appropriations were made prior to 1927; however, by 1929, almost 30% of counties had been promised financial support. In subsequent years, the geographic spread slowed, with the share increasing by only an additional 20 percentage points between 1929 and 1942. At the time, the modal treated county-year had only one hospital. We present the spatial distribution of first appropriations by county and year in panel B of Figure 1.

Although Black residents composed 30% of the population in the Carolinas in 1924, they had far more limited access to hospitals than White residents due to institutionalized racial segregation. The Duke Endowment paid “special attention” to racial equality when making funding decisions, and may have helped to narrow long-standing racial disparities in access to health care. In the period we study, The Endowment directed the vast majority of its support in North Carolina— 75.5% of appropriations and 74.0% of payments— to hospitals that accepted patients of both races. White-only hospitals received 15.6% of capital appropriations (16.7% of payments), while Black-only hospitals received 8.9% (9.3% of payments).⁵

2 Data

2.1 Hospital funding

Our paper combines multiple hand-collected data sources into a county-by-year of birth panel. Our measures of treatment are derived from data on appropriations and payments for capital projects reported in the *Annual Reports of the Hospital Section* of The Duke Endowment (See Online Appendix Figure A1 for an image of the raw data from the 1940 annual report). Each entry includes the name of the hospital, its location, as well as the dollar value of appropriations for capital expenditures, actual payments from the Endowment, and the purpose of the funding.

⁵While the Endowment paid “special attention [to] the needs of blacks in white-run biracial and black-run hospitals” (Thomas 2011), Gamble (1997) notes that this “support, came at a price... The assisted facilities would be ‘for blacks’ but would be supervised by whites.” Gamble (1995) cites specific examples, including changes at Lincoln Hospital where The Duke Endowment “limited the access of black physicians to the operating room and insisted that all their work be supervised by white surgeons” and replaced the Black superintendent with a White manager against the will of the surrounding Black community. Gamble (1995) argues that this “arrangement frequently denied black people a voice in determining the direction of institutions which served them” and “reinforced a racial hierarchy within medicine.” She concludes that The Duke Endowment “contributed to the advancement and maintenance of a separate, and not equal, medical world for black physicians, nurses and patients.”

An appropriation for capital expenditures represents a signed commitment from The Duke Endowment to reimburse costs associated with the expansion or improvement of existing hospitals as well as the construction of new facilities. Overall, 89.2% of appropriations were eventually paid. In most cases, the first payments on appropriations were delivered within two years (86.9%) and only 10.8% of appropriations never resulted in a payment (See Online Appendix Figure E1).

Between 1927 to 1942, Duke approved appropriations for 130 projects at 65 hospitals in 48 counties, with 116 projects at 59 hospitals in 44 counties ultimately receiving funding by 1942. Appropriations are categorized by the “purpose” of a project. While these categories are not mutually exclusive (e.g., an addition could also include purchase of equipment), they are broadly informative of the types of activities funded by the Endowment. About 29.2% of projects approved for appropriations were for additions to existing not-for-profit hospitals, while purchases of equipment and of existing facilities (changing a for-profit facility into a not-for-profit) represented 22.3% and 10.0% of approved projects, respectively. Finally, 26.2% of approvals were for the construction of new hospital facilities and 14.6% were for building homes for nurses (Online Appendix Table E1).

2.2 Number of hospitals and hospital beds

Our analysis of hospital capacity is based on data from “Hospital Service in the United States,” published in the *Journal of the American Medical Association* (American Medical Association 1942a; Esteves et al. 2022). These records include the count of hospitals and hospital beds in each county, and are available for 1920, 1925, and annually from 1927 onward. To supplement these data, we include newly digitized hospital listings from the biannual *American medical directory* (AMD) (American Medical Association 1942b). We obtain additional data on hospitals serving Black patients from Pollitt (2017). We focus our analysis specifically on general hospitals, excluding tuberculosis sanitariums and specialty hospitals, which were not the focus of Duke funding. General hospitals are classified based on their control type, with our main focus on not-for-profit – including those designated as church-run, non-profit, or public – and proprietary (e.g., for-profit) hospitals. Not-for-profit status was a prerequisite for Duke support, and proprietary hospitals were required to convert to a non-proprietary status to access funding.

2.3 Number of physicians by race and quality

In our estimates of the physician labor market, we use biannual data from the AMD covering the period 1921-1942, encompassing the universe of licensed physicians practicing in North Carolina. An individual entry includes the physician’s name, race, year of birth, place of practice, medical school, year of medical school graduation, and year of license (See Online Appendix Figure A1 for an extract from the AMD). We aggregate these data to the county-by-AMD-wave level. Following Moehling et al. (2020), we base physician quality on whether the medical school a physician graduated from required a 2-year undergraduate degree at the time of their admission.

As medical school was typically a four-year program, we define a doctor as high quality if they graduated at least four years after the medical school they attended introduced a two-year degree requirement as an admission pre-requisite.⁶ All other doctors are considered low quality.

2.4 Births and mortality

We focus on infant mortality in our analysis because we can accurately gauge exposure to Duke funding during infancy based on information about place and year of birth reported in individual North Carolina death certificates. In the absence of information on place of residence at death, we cannot reliably assign exposure to Duke funding in the case of adult mortality. Likewise, without longitudinal information on place of residence, we cannot accurately assign cumulative exposure to Duke funding over the life cycle.

Infant mortality: We obtain information on place of birth from the universe of individual North Carolina death certificates between 1917 and 1963 (Cook et al. 2014, 2016), and construct a new measure of infant mortality by county of birth based on the aggregation of individual deaths to the county-by-year-of-birth level (IPUMS USA 2015; US Geological Survey 2019; IPUMS USA 2021).⁷ We define infant death as occurring between the day of birth and one year of life, limiting our sample to deaths of individuals born in North Carolina because similar infant mortality data are not available for South Carolina.

Births: We obtain annual data on births by county, year, and race from 1922 to 1948 from volumes of the *Annual Report of the Bureau of Vital Statistics of the North Carolina State Board of Health* (North Carolina Bureau of Vital Statistics 1922). We use the number of births by year and county of occurrence to construct infant mortality rates and to use as weights in regressions with data at the county-by-year level. Data on births are not available prior to 1922.⁸ We combine these data with the count of infant deaths by county and year of birth to construct the infant mortality rate per 1,000 live births.

⁶Moehling et al. (2020) provide much greater discussion about reforms in US medical education. Even prior to the Flexner report in 1910, the American Medical Association recognized the need for sweeping changes in medical education. To be rated high-quality by the AMA, schools had to require two years of undergraduate coursework and thus we adopt the same delineation in our paper.

⁷Data on deaths by place of *residence* are not reported in the aggregate data until 1932 – well into the rollout of Duke funding. James M. Parrott, the Secretary and State Health Officer of North Carolina recognized the problem with reporting aggregate deaths by place of occurrence in the 1933 *Annual Report of the Bureau of Vital Statistics*: “Deaths have been allocated to the place of legal residence of the deceased before death. Prior to 1932, deaths had been recorded by place of occurrence only. This gives a distorted rate for those counties that have medical centers within their borders and for the counties from which their patients are drawn” (North Carolina Bureau of Vital Statistics 1933). Ideally, we would observe information on place of residence at the time of birth but to the best of our knowledge these data do not exist.

⁸The annual reports were not published between 1917 and 1921. We transcribed the data on births by place of occurrence from the North Carolina Bureau of Vital Statistics because publicly available data do not report births separately by race until 1959 (Bailey et al. 2015).

Panel C of Figure 1 depicts North Carolina’s infant mortality rate between 1922 and 1942 for both races, as well as separately for Black and White infants. Until about 1927, the first year with an appropriation from The Duke Endowment, there are stable (if not slightly increasing) mortality rates that are approximately 50% higher for Black as compared to White infants. Starting in 1928, infant mortality for both races begins to decline steadily before dropping precipitously between 1932 and 1933. This drop was not uniform across the state. Panel D of Figure 1 shows that between 1922 and 1942 some counties experienced declines in infant mortality of over 50%, but there were other counties that saw non-trivial increases.

Our main analysis of infant mortality uses county-by-year of birth panel data from 1922 to 1942. As explained above, 1922 is chosen as the first year since data on births by race are not provided for earlier years. The choice of 1942 as the endpoint for the sample is motivated by a number of factors. First, as Figure 1a illustrates, the rollout of counties receiving Duke funding proceeded rapidly upon establishment of the Endowment while the post-1942 expansion (not depicted) saw relatively few additional counties gain access to funding. Nevertheless, we also collected data on Duke Endowment funding for the period up to 1962 and, in Section 6 we document that our infant mortality results are robust to extending the end year of the panel. Second, by ending the sample in 1942, we limit any possible influence that World War II may have on the analysis.⁹ Finally, this choice facilitates consistency between the short-run and long-run analyses.

Long-run mortality: We face two primary challenges in estimating the effect of Duke Endowment funding on long-run mortality. First, given our focus on the 1922 to 1942 birth cohorts, we cannot observe every completed lifespan as some individuals born during these years are still alive. Second, to the best of our knowledge, publicly available aggregate data on mortality for the cohorts of interest do not extend to the present day. Thus, we turn to the public-use version of the Social Security Administration (SSA)’s Numerical Identification Files (Numident) from the National Archives and Records Administration (NARA) that contains records for every Social Security Number (SSN) assigned to individuals with a verified death, or who would have been over 110 years old by December 31, 2007 (Social Security Administration 2007).¹⁰ The data include the exact dates of birth and death, and a place of birth text string. Since place of birth is truncated at the 20th position, we use a crosswalk file that maps the text strings to a geocoded location and associated county of birth (Black et al. 2015). We collapse the Numident at the county- and year-of-birth-by-follow-up-year level, inferring age at death as the difference between year of birth and

⁹After the US entered World War II on December 8, 1941, Fort Bragg in North Carolina was one of the major military installations in the nation. The war effort could have affected both fertility and the health of infants directly through programs targeted at spouses of conscripts such as the Emergency Maternal and Infant Care Program.

¹⁰To the best of our knowledge, the NARA Numident is the only publicly available data source that can be used to study long-run mortality in our context. While the Life-M data include linked mortality and natality records for North Carolina, for our sample cohorts it has more limited coverage and only includes deaths up to 1994.

year of death. We link these data to exposure to Duke Endowment support at the birth-county by year-of-birth level.

Construction of the estimation sample for long-run mortality differs from the short-run analysis in two ways. First, we consider the mortality rates across a range of ages, rather than just the first year of life. This means that an observation is at the birth-year by county-of-birth by follow-up-year level, where a follow up year refers to the year in which we observe long-run mortality for each birth cohort. Stated differently, each observation is the mortality rate by single year of age for a given birth-county and birth-year cohort (Surveillance, Epidemiology, and End Results Program 2022).

For administrative reasons, the public-use Numident is incomplete prior to 1988 (Goldstein et al. 2021), which means that 1932 is the earliest year we are able to include in our long-run analysis.¹¹ Because we want to ensure that our analysis is not confounded by the effects of Medicare eligibility, we do not extend our sample beyond the 1941 birth cohort.¹² To avoid biasing our estimates due to differential survival probabilities by age, we use a balanced panel of follow up ages for each county-by-year birth cohort, which implies that all cohorts are observed exactly nine times at ages 56 to 64 during available follow up years (e.g., the 1932 cohort is observed from 1988 to 1996). We discuss the econometric implications of these sample restrictions in Section 5.

Pneumonia mortality: The North Carolina vital statistics annual reports also include data on pneumonia mortality by county and year. We use this to construct the exposure shares for our analysis of sulfa innovation. Specifically, we compute average pneumonia mortality per 100,000 in the population from 1922 to 1926 – the five years prior to the initial rollout of Duke funding. County population is obtained from the National Historical Geographic Information System (NHGIS) (Manson et al. 2022).

3 Methods

We use a difference-in-differences research design with staggered treatment adoption to estimate the first-stage effects of hospital funding from The Duke Endowment on hospital capacity, and physician supply and quality. Additionally, we examine the impact of hospital modernization on infant and long-run mortality. We consider a county to be treated beginning in the first year that it received an appropriation for capital expenditures and all years thereafter.

First stage: We first consider if Duke support had an effect on the medical sector. We refer to these analyses as the first-stage throughout and we estimate them using a linear two-way fixed

¹¹Deaths at ages younger than 56 for the 1931 and earlier birth cohorts occurred prior to 1988, the earliest year that the public-use Numident is complete. Thus, any reported deaths from earlier cohorts would introduce survival bias.

¹²Changing life expectancy across time provides another justification for restricting the set of birth-year cohorts and follow-up ages. Life expectancy at birth does not surpass age 65 until after 1940 for White individuals and after 1970 for Black individuals (Haines 2008; Arias 2014; Arias et al. 2021). See Online Appendix Figure F1.

effects regression model. We prefer a linear model for the first-stage since it more easily enables comparisons across different mutually exclusive groups (e.g., a comparison of the effect for the number of high-quality doctors vs. the effect for the number of low-quality doctors will also be informative for the total number of physicians). Our first-stage effects are estimated using the following model:

$$Y_{ct}^R = \beta_0 + \beta_1 1(t \geq \text{First appropriation year})_{ct} + \zeta_c + \eta_t + \Theta X_{ct} + \epsilon_{ct} \quad (1)$$

where c indexes county and t indexes year – either the calendar year for hospitals or a wave of the *American Medical Directory* (AMD) for physicians.¹³ For hospitals, our primary outcome of interest, Y_{ct}^R , is the number of hospital beds or hospitals in a county and calendar year. R indexes the hospital type, which can be either not-for-profit (encompassing church-run and public hospitals), proprietary, or the combination of the two categories. For physicians, Y_{ct}^R represents the total number of doctors, or is disaggregated by quality (high vs. low) and race (Black vs. White). For the first stage, we also express the dependent variables as rates per 1,000 live births.¹⁴

The treatment variable is $1(t \geq \text{First appropriation year})_{ct}$, which takes the value of one for the first year that a county received an appropriation from The Duke Endowment and for all years thereafter. The treatment effect of interest is β_1 , which captures the change in the number of hospital beds, hospitals, or doctors of a particular type compared to the untreated counterfactual. This is an intent-to-treat effect, as capital appropriations from The Duke Endowment represented a commitment to provide financial resources that did not account for either the utilization (i.e., actual spending) or the allocation of funding (i.e., what the money was spent on).

A vector of control variables is denoted by X_{ct} : the percent of the population that is illiterate, Black, other non-White race, or residing in urban areas; retail sales per capita, which helps to control for New Deal spending (Fishback et al. 2005); and an indicator for the presence of a county health department (Ferrell and Mead 1936; Hoehn-Velasco 2018, 2021).¹⁵ In a balance test that follows the approach presented by Pei et al. (2019), we do not find a statistically significant relationship between changes in covariates and the treatment, easing concerns that selection or differential trends are driving our effects (See Online Appendix Table G1).¹⁶ Most notably, En-

¹³The AMD was published every second year from 1921, except for a three-year gap between 1931 and 1934.

¹⁴We use counts of births (which are published annually) to avoid potential issues found in other papers that use population estimates that are interpolated between decennial census years (Arthi et al. 2022).

¹⁵Another possible confounder may be the presence of health insurance coverage. However, over the years considered in our analysis very few Americans had any type of health insurance coverage, especially so in the early years of our sample. While state-level data are not available, national data show that only 9% of the US population had coverage for hospital care in 1940 (Health Insurance Institute 1977). During our sample period, this type of coverage (later known as Blue Cross) was sponsored by hospitals. In other words, the existence of hospitals in North Carolina was necessary for people to have access to Blue Cross. Furthermore, these hospital prepayment plans were slower to grow and enroll subscribers until states enacted enabling legislation that exempted them from insurance reserve requirements (Thomasson 2002). North Carolina did not enact such legislation until 1943, after our main sample ends.

¹⁶The p-value on a joint F-test for significance in a regression of Duke support on the covariates is 0.192.

dowment support does not have an effect on our best measure of economic activity, retail sales per capita. This indicates that our findings are driven by changes in healthcare provision rather than general economic improvement resulting from hospital modernization. Nevertheless, to enhance precision and reduce residual variation, we include the vector of control variables described above in our preferred specification. Finally, all specifications include county (ζ) and calendar year or AMD wave (η) fixed effects. We cluster the standard errors (ϵ_{ct}) by county to account for correlated errors within a county. In our preferred specification, we weight observations by the number of births that occur in each county-year.

Event studies: We also present event studies of the following form:

$$Y_{ct}^R = \beta_0 + \sum_{j=-p}^{-2} \beta_j 1(t - \text{First appropriation year}_{ct} = j) + \sum_{j=0}^q \beta_j 1(t - \text{First appropriation year}_{ct} = j) + \zeta_c + \eta_t + \epsilon_{ct} \quad (2)$$

where j indicates the time period relative to the first appropriation year. The coefficients β_{-p} to β_{-2} capture trends prior to the start of the funding while coefficients β_0 to β_q capture the effects of access to Duke Endowment resources.¹⁷ The omitted category is one year prior to the start of the treatment. Our event studies display the coefficients ± 6 for hospitals or ± 5 for physicians, and omit more extreme values which are nonetheless included in the estimation. We use county- and year-of-birth cohort size as weights and avoid additional controls for ease of interpretation and to follow recent recommendations in the literature (Roth et al. 2022). Standard errors (ϵ_{ct}) are clustered by county to account for correlated errors within a county.

Robust estimators: In Figures 2 and 3, we also implement a selection of the recent innovations that address the concerns that the linear two-way fixed effects regression model may be biased in the presence of staggered treatment timing or treatment effect heterogeneity (Goodman-Bacon 2021a; Roth et al. 2022): stacked regression (Cengiz et al. 2019), extended two-way fixed effects (Wooldridge 2021), and the Callaway and Sant’Anna (2021) estimator (see also Online Appendices B and C).¹⁸

Poisson estimation: Our preferred estimation method for the analysis of infant and later-life mortality is Poisson pseudo-maximum likelihood (Correia et al. 2020), under which coefficients of

¹⁷The most extreme negative coefficient, p , in the event studies is 18 for hospitals and 9 for doctors, while the most extreme positive coefficient, q , is 15 for hospitals and 7 for doctors.

¹⁸Our implementation of stacked regression creates treatment-timing group specific data sets that include counties first exposed to Duke Endowment funding in a particular year and “clean control” counties that are never treated in various estimation windows. We stack these data sets and saturate the county- and year-of-birth fixed effects in our DiD regressions and event study specifications with indicators for each of the stacked data sets.

interest can be interpreted as semi-elasticities – the percent change in the outcome due to Duke funding exposure around the time of birth.¹⁹ We prefer Poisson estimation to OLS for the mortality analysis due to the presence of zeros in the dependent variable and because mortality risk increases dramatically across the age ranges we examine. Dependent variables containing zeros are explicitly modeled and handled well by the Poisson (Santos Silva and Tenreyro 2011), avoiding the need to manipulate the data and introduce additional issues (Mullahy and Norton 2022; Chen and Roth 2023).²⁰ Another feature of the Poisson is that it allows for a proportional analysis that ensures the comparability of estimates across groups with different baseline mortality risk. Thus we can easily conduct a joint estimation of the long-run mortality effects without needing to account for the wide spread in baseline mortality risk across the follow-up ages we study.

Short-run mortality: We estimate the following regression:

$$Y_{ct}^R = \exp(\alpha_0 + \alpha_1 1(t \geq \text{First appropriation year})_{ct} + \zeta_c + \eta_t + \Theta X_{ct}) \epsilon_{ct} \quad (3)$$

where c indexes county of birth and t indexes birth year. Our primary outcome of interest Y_{ct}^R is the county-by-birth year count of deaths or the infant mortality rate defined in Section 2.4. We consider both counts and rates as dependent variables to demonstrate that our construction of the denominator in the death rate does not drive our results. We also consider alternative definitions of the mortality variable, e.g., first-day infant mortality rate. We superscript this variable with R to indicate separate outcomes for Black and White persons or outcomes that pool together both racial groups.²¹ The treatment effect of interest is $\exp(\alpha_1) - 1$, which captures the percent change in infant mortality compared to the untreated counterfactual in our difference-in-differences framework. Given that the model is non-linear, the treatment effect is identified under the assumption of parallel trends in the log of the expected potential outcomes (Lechner 2011). For the infant mortality analysis, we use a similar set of robust estimators as in the first-stage analysis, the differences being that we estimate stacked regression and extended two-way fixed effects by Poisson pseudo-maximum likelihood and log the dependent variable for Callaway and Sant’Anna (2021).

In our preferred Poisson specification, we add X_{ct} , a vector of control variables described above, and weight observations by the number of births that occur in each county-year.²² Further-

¹⁹Throughout the paper, “Poisson” refers to the Poisson pseudo-maximum likelihood estimation method, rather than the assumption that the dependent variable has a Poisson distribution.

²⁰Along a similar line, we do not use Poisson estimation for the first stage analysis because the first-stage sample includes counties that have no time variation in hospitals or physicians (e.g., an untreated county could always have 15 beds), the county fixed effect would cause a perfect prediction of this outcome, which would cause these observations to drop out of the estimation. In this case, we prefer the linear model for ease of interpretation due to the ability to retain these counties. The mortality treatment effect estimates are similar when estimated using a linear model.

²¹Hispanic people are not included as a separate racial category in our data. We do not have data on Native Americans, who constituted only 0.5% of the North Carolina population at the time.

²²This method produces identical results to a model that uses the count of deaths as the outcome variable and

more, all specifications include county-of-birth (ζ) and year-of-birth (η) fixed effects. We cluster the standard errors (ε_{ct}) by county of birth to account for correlated errors within a county.

We also present event studies of the following form:

$$Y_{ct}^R = \exp(\alpha_0 + \sum_{j=-18}^{-2} \alpha_j 1(t - \text{First appropriation year}_{ct} = j) + \sum_{j=0}^{15} \alpha_j 1(t - \text{First appropriation year}_{ct} = j) + \zeta_c + \eta_t) \varepsilon_{ct} \quad (4)$$

where j indicates the time period relative to the first appropriation year. The coefficients α_{-18} to α_{-2} capture trends prior to the start of the funding while coefficients α_0 to α_{15} capture the effects of access to Duke Endowment resources. The omitted category is one year prior to the start of the treatment. Our event study plots omit the left-most coefficients for relative time periods -18 to -7 , and the right-most coefficients for relative time periods 7 to 15 . The dropoff in the number of treated counties by event time beyond ± 6 motivates the choice of these cutoffs (See Online Appendix Figure I2). We weight the estimates by county- and year-of-birth cohort size and do not include additional controls. Standard errors (ε_{ct}) are clustered by county of birth.

4 Results

4.1 First stage: Hospitals

We first investigate the extent to which The Duke Endowment achieved James Buchanan Duke's vision to increase access to modern hospitals for the residents of North Carolina. We analyze the impact of support from The Duke Endowment on two key aspects of hospital capacity measured at the county-by-year level: total hospital beds and the number of hospitals. Given that The Duke Endowment explicitly restricted its funds to hospitals run on a not-for-profit basis, we separately analyze the effects for both proprietary (for-profit) and not-for-profit hospitals. This approach allows us to assess potential compositional changes in the hospital industry stemming from The Duke Endowment's support.

Figure 2 plots trends in hospital bed capacity in counties that received capital appropriations from The Duke Endowment by 1942 in comparison to those that did not (top row), and for treated counties, the ratio of beds per year to the number of beds in the year prior to treatment (middle row). Beds in non-profit hospitals increased dramatically after the county received Duke funding, while beds in proprietary hospitals decreased, and the overall count increased as well.

Table 1 presents results based on Equation 1 for the effects of Duke financial support on hospital capacity. Our simplest analyses, displayed in columns 1 to 3, report the effects on counts of beds (Panel A) and hospitals (Panel B), while columns 4 to 6 show the effects on beds and hospitals

includes an exposure variable equal to the number of births.

per 1,000 births. Across all specifications, results indicate that The Duke Endowment increased hospital capacity. In our preferred specification (reported in column 6) that focuses on beds per 1,000 births and includes both weights and controls, we find that Duke support increased the number of beds by 25 per 1,000 births, or 49.7% when calculated at the sample mean (See Online Appendix Tables N1 and N2 for summary statistics). These increases were concentrated among beds in not-for-profit hospitals, which increased by 70.1%, compared to a decrease in for-profit beds of 61.4%. Event studies in Figure 2 validate these findings and generally confirm an absence of pre-trends. We extend these analyses to 1950 in Section L of the appendix and find that effects are persistent across the years included in this sample with a longer time frame as well.

Support from The Duke Endowment also resulted in an increase in the number of hospitals in North Carolina. While less precisely estimated across all specifications than the effects on beds, results reported in Table 1 suggest that Duke assistance resulted in an increase in not-for-profit hospitals by 0.33 per 1,000 births per treated county, or 36.3%. In contrast, proprietary hospitals decreased by 42.4%. At the extensive margin, only 40 of North Carolina's 100 counties had a general hospital in 1922. By 1942, 65 counties had at least one.

4.2 First stage: Physician labor market

Next, we consider whether the effects of support from The Duke Endowment extended beyond hospitals and into the physician labor market. Even though the funding was not directly tied to hiring physicians, those who ran the foundation believed that improving facilities would attract higher-quality doctors. The top row of Figure 3 shows the number of pooled, high-quality, and low-quality physicians per 1,000 live births across treatment status and calendar time. It is clear from this figure that the number of high-quality physicians increased in *both* treated and control counties and that the increase was greater in treated counties following the rollout of support from the Endowment. Similarly, the number of low-quality physicians decreased in both treated and control counties, with a somewhat greater decrease in treated counties.

Table 2 presents the physician labor supply results. We document the effects on the number of doctors (columns 1 to 3) and doctors per 1,000 births (columns 4 to 6) measured at the county-by-AMD wave level for all physicians (Panel A), as well as for Black (Panel B) and White (Panel C) physicians. Within each panel, we also measure the effects of Duke support on high- and low-quality physicians. Column 1 represents the simplest specification, where weights and controls are not included and the dependent variable is the count of physicians. Our preferred specification in column 6 expressed as a rate includes weights and controls, and it is equivalent to the specification we later use to examine the impact of Duke support on the infant mortality rate.

In the pooled sample, the results indicate that The Duke Endowment realized its goal of attracting better-qualified physicians to the hospitals it supported. Table 2 shows a statistically significant

positive effect of Duke Endowment support on the total labor supply of physicians in a county, and a shift in the composition of physicians – an increase in higher-quality physicians and a decrease in lower-quality ones. The results, reported in our preferred specification in column 6, suggest that the number of doctors per 1,000 live births increased nearly 11.1%, with high-quality doctors increasing by 60.2% and low-quality physicians decreasing by 5.5%, although the latter is less precisely estimated. This change in physician composition likely had meaningful effects on the state of medical knowledge in an area. In our sample, “higher quality” physicians were licensed, on average, 21 years later than “lower quality” physicians. This means that for every low-quality physician replaced by a newer one, the state of medical knowledge advanced by more than two decades. When estimating the impact of Duke support by the race of physician, we observe a 245.7% increase in the number of high-quality Black physicians (2.3 per 1,000 births), and of high-quality White physicians by 28.0% (2.8 per 1,000 births). Conversely, low-quality Black physicians decreased by 50.7%, but the number of low-quality White physicians per 1,000 births was less affected.

We also explore alternative measures of physician quality, considering factors such as whether the physician’s medical school ever received an A or A+ rating from the American Medical Association (AMA), its existence and approval status by the AMA in 1942, or if it had closed or been absorbed by another medical school between 1922 and 1942, and find similar results (See Online Appendix Tables C2 and C3). While most physicians were generalists in the early-twentieth century, we find that Duke funding led to an increase in surgeons and specialists, as defined by various measures of specialization, including AMA membership, fellow status, and society memberships (See Online Appendix Table C4). We also find suggestive evidence from decennial US census data that Duke funding increased employment in the medical and hospital industry among nurses, attendants, and clerical workers (See Online Appendix Figure C3). These figures provide evidence that skilled hospital employment, including the clerical workers needed to implement the uniform accounting and record keeping systems championed by The Duke Endowment, likely increased alongside Duke support.

4.3 Infant mortality

We begin by presenting raw infant mortality data by calendar year and treatment status. The first row of Figure 4 depicts infant mortality rates per 1,000 live births in North Carolina between 1922 and 1942 separately for counties that ever received Duke Endowment support by 1942 (solid pink line) and those that did not (dashed black line). We present results pooled across racial groups (column a) and separately for Black (column b) and White infants (column c). There are three descriptive patterns of interest in these figures. First, before the initial appropriation was granted in 1927, the infant mortality rates trended in a similar way in both treated and untreated counties, suggesting that our identifying assumption of parallel trends is valid. Second, the infant mortality

rate is higher in the pre-Endowment period for eventually treated counties, which is consistent with the idea that the Endowment targeted worse off and underserved communities. Third, after 1927, the infant mortality rate in treated counties converged to that of untreated counties, suggesting that the Endowment could have played a causal role in driving the observed data patterns.

The middle row of Figure 4 plots average infant mortality for treated counties in event time. Pre-treatment data are displayed using a solid pink line and post-treatment data are plotted by a dashed pink line. The dashed black line shows a best-fit line estimated using only pre-treatment data. This trend is continued into the post-treatment period, providing a straightforward basis for comparisons of potential outcomes. These figures show a clear trend break following the first capital appropriation from the Endowment. In each figure, the deviation is substantial and negative, indicating that Duke support likely reduced infant mortality.

Our difference-in-differences results from estimating Equation 3, reported in Table 3, are consistent with these descriptive patterns. In columns 1 to 3, the dependent variable is the number of infant deaths and in columns 4 to 6 it is the infant mortality rate per 1,000 live births. Each coefficient in panels A to C is estimated by Poisson pseudo-maximum likelihood and comes from a separate regression that includes county-of-birth and year-of-birth fixed effects. Columns 1 and 4 present baseline two-way fixed effects models without weights or controls, columns 2 and 5 weight observations by the number of county-by-year births, and columns 3 and 6 include weights and the control variables described in Section 3. Panel A presents pooled estimates while panels B and C estimate effects separately for Black and White infants, respectively.²³ In the last row, we present p-values for the difference in coefficients presented in panels B and C that are estimated using a model that is fully interacted with race.

We find that exposure to Duke Endowment support around the time of birth reduced infant mortality by 6.1% to 8.1% (panel A). The effects are larger for Black infants compared with White infants. In our preferred specification (column 6), the estimated coefficients suggest that Duke support led to a 7.5% reduction in overall infant mortality, a 13.6% reduction in Black infant mortality, and a 4.7% reduction in White infant mortality (p-value: 0.11).²⁴ The difference between the race-specific estimates is statistically significant (with a p-value of 0.02). The effects on pooled infant mortality are similar in size across project types: new hospitals, additions to hospitals, the purchase of equipment, and the purchase of existing facilities – typically proprietary hospitals slated

²³There are fewer observations in panels B and C because in 5 counties we do not observe Black births in all county-by-year cells. We drop these counties from the analysis split by race but include them in the pooled analysis in panel A. Pooled results are almost identical when run on the sample considered in panels B and C. We discuss the adjustments to birth or death counts in greater detail in Online Appendix H.

²⁴The estimates from our preferred specification are among the most conservative across a range of models that we consider (see Figure 5). In many of these specifications, and across alternatives reported in Online Appendix D, we find larger and more precisely estimated effects, including statistically significant effects for White infant mortality. See Section 6 for further discussion.

for conversion to a non-profit (see Online Appendix Figure E2). While we find large effects for mortality on the day of birth, capturing the impact of healthcare provided before, during, and immediately after delivery, we also find effects on infant mortality from day 1 to 365 infant mortality. Thus, our results are not fully explained by healthcare provided at birth, suggesting that support from The Duke Endowment likely improved neonatal health care as well (See Online Appendix Table G4).²⁵

The average treatment effects documented in Table 3 are further validated by event studies based on estimating Equation 4. These results are presented in the bottom row of Figure 4. In each case, before any funding was allocated to a county, there are no significant differences in infant mortality between treatment and control groups, supporting our assumption of parallel trends. In contrast, in the post-treatment periods we do see marked declines in mortality starting from the second year after the initial appropriation.²⁶

Magnitudes: As context for our estimates, we consider the overall changes in infant mortality over the study period. The pooled infant mortality rate in North Carolina declined by 35.9% (from 72.6 to 46.6 deaths per 1,000 live births) from the time period just before Duke’s intervention (1922-1926) to the end of the study period (1938-1942). Similar reductions of 37.9% and 34.8% were observed for Black and White infant mortality, respectively. Our treatment effect estimates imply that 21% of the reduction in pooled infant mortality could be attributed to Duke support. These back-of-the-envelope calculations differ considerably by race, as Duke support contributed to 35.9% of the reduction in Black infant mortality but only 13.6% of the reduction for White infants.

We can also compare our effect sizes to other estimates in the literature, especially those examining a variety of health-related interventions from the US in the first half of the twentieth century. For example, the Sheppard-Towner Act reduced infant mortality by up to 1.9 deaths per 1,000 births, explaining 21% of the infant mortality rate decline from 1924 to 1929, with much of the

²⁵Our results are also insensitive to excluding deaths in the first week or in the first month of infancy. As we follow published infant mortality statistics in restricting attention to live births, we show in Online Appendix G that potential measurement error in recording stillbirths does not drive our results for mortality on the day of birth.

²⁶In Online Appendix G we further document effects on fertility measured by the number of births per 1,000 women (Table G3) and on maternal mortality (Table G2). Point estimates for fertility neither imply major changes in childbearing due to improved health care infrastructure, nor suggest a cohort-composition changes concern that one might have when interpreting our main specifications that use the number of births to create both the infant mortality rate and to weight observations. In any case, our specifications with death counts as an outcome are unaffected by concerns related to the number of births. On the other hand, in Table G2 we find evidence that maternal mortality declined by roughly 20% following a Duke capital appropriation. Unlike our infant mortality data, these data are not constructed from death certificates, but rather they are reported at the county-year level in North Carolina annual vital statistics reports. The data are only reported beginning in 1932, and thus the sample includes roughly half the number of observations as in the main specification. Due to the limited number of observations and lack of microdata, we cannot conduct the same battery of robustness test that we do for infant mortality. We view these results as supportive evidence that Duke funding improved health for other individuals than infants.

aggregate effect driven by decreases in mortality among Black infants (Moehling and Thomasson 2014). On the other hand, midwifery licensing reforms had a modest effect on average infant mortality and no detectable relationship with White infant mortality. For non-White infants, the coefficients were larger at 9.6% and statistically significant (Anderson et al. 2020).

Our estimated effects are consistent in size with those observed in more recent healthcare interventions. For example, Goodman-Bacon (2018) examined the impact of Medicaid introduction in the 1960s and finds that a 1 percentage point change in the initial Aid to Families with Dependent Children (AFDC) rate in a state resulted in a 1.4% reduction in infant mortality, implying an ATT of about 30% for non-White infants. Similarly, a 30 percentage point increase in Medicaid eligibility for pregnant women between 1979 and 1992 decreased infant mortality by 8.5% (Currie and Gruber 1996).

Intensive margin: Using data on the exact amount of money appropriated and disbursed by The Duke Endowment, we also investigate the intensive margin effects of financial support on infant mortality and conduct a cost-benefit analysis. We re-estimate Equation 3 by substituting cumulative appropriations or cumulative payments at the county-by-year level in place of a dummy treatment variable. The results of this analysis are presented in Table 4. Columns 1 and 2 show effects per \$1 million (2017 \$) of appropriations and payments, respectively, while column 3 shows the conversion rate between a dollar appropriated (i.e., a commitment to pay) and a dollar paid. We find that almost 90% of the promised appropriations were disbursed. When we consider the actual money spent by the foundation, we find reductions in the infant mortality rate in the range of 7.6% (White infants) to 8.9% (Black infants) per \$1 million.

Based on the amount appropriated and spent, one life was saved for every \$19,110 appropriated or \$17,092 spent (in 2017 dollars), implying that the Endowment’s intervention passes any cost-benefit analysis using a reasonable estimate for a VSL.²⁷ Most other papers in the literature lack the detailed cost data needed to provide similar estimates, with the notable exceptions of Currie and Gruber (1996) and Goodman-Bacon (2018), who find that an infant death could be averted at a cost of about \$1.8 million and \$171,000 (in 2017 dollars), respectively, suggesting that the investment in hospital infrastructure during the 1920s to 1940s might have been more cost-effective than Medicaid’s introduction in the 1960s or its later expansions in the 1980s and 1990s.

4.4 Interaction with medical innovation

We now ask if the effects of Duke support were magnified (complements) or diminished (substitutes) by medical innovation in the form of sulfa drugs – the first modern treatment that could effectively control and fight bacterial infections. To the extent that sulfa drugs reduce mortality

²⁷For example, Ashenfelter and Greenstone (2004) estimate VSL at \$2.8 million while in a review article Viscusi (2018) finds a median VSL in US at \$10.3 million.

(Jayachandran et al. 2010), we would expect the two inputs to be complements. On the other hand, easy access to effective outpatient treatments such as antibiotics might lead to a shift in healthcare provision, rendering hospital access – the primary goal of Duke support – less important, thus leading to the two inputs being substitutes.

We first estimate a shift-share model of sulfa drugs on infant mortality. Since pneumonia was highly treatable with sulfa, we use baseline pneumonia mortality as the share factor. Specifically, the variable $\text{Pneumonia mortality}_c$ is the average county-level pneumonia mortality for the years 1922 to 1926. The shift factor is an indicator variable Post sulfa_t that takes the value of 1 for the years 1937 and beyond when sulfa drugs were available. Results from this analysis suggest that sulfa drugs reduced pooled infant mortality by 5.2% (see Online Appendix Table G6). To assess whether the two inputs – Duke support and sulfa drugs – are complements, substitutes, or independent treatments, we interact our roll-out (Duke support) and shift-share (sulfa drugs) difference-in-differences designs and estimate the following equation:

$$\begin{aligned}
Y_{ct}^R = & \exp(\gamma_0 + \gamma_1 1(t \geq \text{First appropriation year})_{ct}) \\
& + \gamma_2 1(t \geq \text{First appropriation year})_{ct} \times \text{Post sulfa}_t \\
& + \gamma_3 1(t \geq \text{First appropriation year})_{ct} \times \text{Pneumonia mortality}_c \\
& + \gamma_4 \text{Pneumonia mortality}_c \times \text{Post sulfa}_t \\
& + \gamma_5 1(t \geq \text{First appropriation year})_{ct} \times \text{Pneumonia mortality}_c \times \text{Post sulfa}_t \\
& + \zeta_c + \eta_t + \Theta \mathbf{X}_{ct}) \epsilon_{ct}
\end{aligned} \tag{5}$$

where Y_{ct}^R , ζ_c , η_t , \mathbf{X}_{ct} , and $1(t \geq \text{First appropriation year})_{ct}$ are defined as in Equation 3, while $\text{Pneumonia mortality}_c$ and Post sulfa_t are defined as described above. The main parameter of interest in this equation is γ_5 which describes how the effects of receiving Duke funding change after sulfa drugs become available, accounting for variation in baseline pneumonia mortality across counties. The auxiliary parameters, γ_1 to γ_4 allow us to compute Duke funding effects separately for the pre- and post-sulfa periods (shift factor) and separately for counties with high- and low-pneumonia mortality (share factor). We cluster the standard errors (ϵ_{ct}) by county of birth to account for correlated errors within a county.

Table 5 reports linear combinations of these coefficients for ease of interpretation. Panel A presents an estimate of the interaction between Duke funding and the introduction of sulfa drugs, comparing counties at the 25th percentile of baseline pneumonia mortality to those at the 75th percentile. Results indicate that gains from Duke support increased with the availability of medical innovation, and did so to a greater extent in counties with high baseline pneumonia mortality, implying that these two inputs are complements. Our results are in line with Alsan and Goldin (2019), who find compounding effects of clean water and sewerage infrastructure.

In panels B and C, we break down the interaction coefficient to examine the effects of Duke

support separately in areas with high baseline pneumonia mortality and those without, as well as before and after the introduction of sulfa drugs.

We observe no statistically significant effects of Duke funding in counties with low baseline pneumonia mortality, either before or after the introduction of sulfa drugs (rows 2 and 4 in Panel B). Furthermore, the difference in the Duke effect between the pre- and post-sulfa periods in these counties is small and statistically insignificant (row 2 of Panel C).

However, in counties with high baseline pneumonia mortality, most of the estimated coefficients are statistically significant both before and after sulfa drugs become available (rows 1 and 3 of Panel B). Moreover, these estimated coefficients are larger in the post-period. This reflects the complementary nature of hospital modernization with medical innovation, which in this case may be explained by the fact that The Duke Endowment attracted highly-qualified physicians to the area who were better able to leverage the new technology.

5 Long-run mortality

Given the broad literature suggesting that early-life interventions can have long-run impacts, we next examine if the health gains from Duke support translated into reductions in long-run mortality for those exposed to funding around the time of their birth. On one hand, if reductions in early-life mortality are positively associated with health capital, and by extension human capital and income, then we would also expect to find a decrease in mortality at older ages. On the other hand, if vulnerable individuals induced by Duke funding to survive past the first year of life likewise survived past age 55, then we would expect to find null effects or even increases in long-run mortality.²⁸

We implement the same difference-in-differences framework from Section 4 based on a comparison of cohorts exposed to Duke financial support around the time of birth to those without such support to estimate effects on long-run mortality using the following equation:

$$Y_{ctf}^R = \exp(\beta_0 + \beta_1 1(t \geq \text{First appropriation year})_{ct} + \zeta_c \cdot \nu_a + \eta_t \cdot \nu_a + \Theta \mathbf{X}_{ctf}) \epsilon_{ctf} \quad (6)$$

which is similar to Equation 3 except: (i) we stack birth county-by-year observations over follow-up ages f and (ii) due to the incompleteness of the Numident prior to 1988 (as described in Section 2) we rely on the 1932 to 1941 birth cohorts, which can all be observed for nine follow up years between ages 56 and 64. Y_{ctf}^R is the number of deaths for those born in county c and in birth cohort t and observed in follow-up year f . All specifications include birth county by follow-up age ($\zeta_c \cdot \nu_a$) and birth year by follow-up age ($\eta_t \cdot \nu_a$) fixed effects, along with a vector of controls,

²⁸All our long-run analyses are conditional on survival to at least age 56 for reasons described in Section 2.

\mathbf{X}_{cft} .²⁹ We cluster standard errors at the birth county level and weight each observation by the number of births in each birth county-by-year cohort.

We present the long-run results in Table 6, structured similarly to Table 3, but with the number of later-life deaths as the dependent variable.³⁰ We present pooled estimates in panel A, with separate effects for Black and White adults shown in panels B and C, respectively. These coefficients represent the impact of treatment around the time of birth on mortality between ages 56 and 64, conditional on survival to age 56. We find consistently negative and statistically significant coefficients. In our preferred specification (column 3), the average treatment effect is a 9.0% reduction in long-run mortality. This effect size is slightly larger than our preferred estimate from Table 3, but it is sufficiently close for us to conclude that many of the health benefits gained during infancy appear to have persisted into adulthood. We present and discuss event study estimates corresponding to these results in Section F.1 of the Online Appendix.

The long-run effects on Black and White mortality are similar across specifications, in contrast to the larger effects observed on Black infant mortality in Section 4. This difference may arise from the fact that some of the most vulnerable Black infants, whose survival beyond the first year of life could have been attributed to the intervention, did not live long enough to be included in our long-run sample, which is consistent with known disparities in life expectancy at birth between Black and White babies.³¹ Alternatively, the marginal effects of Duke support may simply not differ by race beyond age 55.

Our long-run results by race differ from Goodman-Bacon (2021b), who observed that early-age Medicaid eligibility reduced non-AIDS adult mortality before age 37 by 14.5% for White individuals and 8.7% for non-White individuals.³² However, our pooled long-run estimates are similar in magnitude to the long-run impact of access to in-hospital deliveries (Fischer et al. 2021) and much larger than the effects of a nurse home visiting program (Hjort et al. 2017). We conclude that Duke support improved the long-run welfare of those treated around the time of birth. Yet, we are almost certainly understating the population-wide impact of Duke support on long-run mortality since it may have also benefited groups who only had access to it *after* infancy and Duke

²⁹Since we observe each county-by-year of birth cohort at ages 56 to 64 during follow up, we allow baseline controls to flexibly vary with follow-up age. These additional control variables are included in the vector \mathbf{X}_{cft} .

³⁰We prefer to use death counts in this analysis since we do not have data on the surviving number of people from each county-year of birth cohort, which would be ideal for computing death rates. In Online Appendix Table F1, we find that results are substantively unchanged when we use the death rate per 1,000 live births in each county-by-year of birth cohort as the outcome.

³¹For example, in 1900, Black and White life expectancy at birth was 33 and 48 years while by 2017 it increased to 61 and 69 years for these two groups, respectively (National Center for Health Statistics 2017). Data for 1900 to 1950 is more scarce but Ewbank (1987) suggests that near the end of our sample period in 1939 to 1940, life expectancy at birth in North Carolina for Black and White infants was 54 and 64 years, respectively.

³²Interestingly, using the same variation, Goodman-Bacon (2018) finds reductions in infant and child mortality only for non-White children.

access may have improved health at later ages for those who already access to Duke supported hospitals at birth.

Permanence of effect: While we find evidence of a long-term reduction in mortality for individuals treated at birth, a question remains: how enduring were the effects of Duke support on infant mortality and hospital capacity? In terms of potential outcomes, it is possible that treated areas could have eventually reached parity with untreated areas in the absence of the treatment, implying temporary effects of Duke support. Section L of the Online Appendix explores this issue by extending the sample until 1950 for the hospital sector analysis and 1962 for the infant mortality analysis. Our findings provide clear evidence that the effects of Duke support are persistent and do not diminish over time. Consequently, our results should be interpreted as a permanent changes in infant mortality and hospital capacity between treated and untreated counties.

6 Robustness Checks

Point estimates and statistical significance are robust to the choice of estimator, specification, transformation of the dependent variable, choice of the control group, and other sample restrictions such as dropping small counties or changing the endpoints of the sample.³³ We provide a visual overview of many of these checks for our infant mortality results in Figure 5.

Robust estimators: Our event studies in Figures 2 to 4 are insensitive to the use of stacked regression (Cengiz et al. 2019), extended two-way fixed effects (Wooldridge 2021), and the Callaway and Sant’Anna (2021) estimators that are designed to address concerns about treatment effect heterogeneity and bias in a difference-in-differences setting with staggered treatment adoption. The results are similar across all of the estimators and, as before, we do not see evidence of pre-trends in our preferred specifications (see Online Appendix Tables B1, C1, and D1 for static difference-in-differences estimates).

The validity of our event studies is further supported by the Goodman-Bacon (2021a) decomposition, which reveals that 2x2 DiD comparisons of treated versus untreated units receive the lion’s share of the weight in the overall two-way fixed effects DiD estimate (see Online Appendix Table I1 and Figure I1). Our main infant mortality results are also invariant to specifying the dependent variable in logs (Online Appendix Table D2), while stacked Poisson event studies are insensitive to adjusting the width of the event-time window from ± 3 to ± 6 and including control variables (Online Appendix Figure D1), providing support for parallel pre-trends.

The effects on White infant mortality are comparable to or larger than those estimated using our preferred specification (Table 3), ranging from -4.3% to -18.4% (Online Appendix Table D1), and

³³When treatment is staggered, DiD estimates place greater weight on treatment groups that are treated closer to the middle of the panel and may be sensitive to panel length. Even though over half of the Duke rollout occurred within the first two years of program launch, the extended rollout period could potentially bias our estimates due to dynamic treatment effects that vary across treatment timing groups (Goodman-Bacon 2021a; Baker et al. 2022).

are generally statistically significant, which suggests that our preferred estimates may be especially conservative. At the same time, estimates for Black infants are always statistically significant and almost universally much larger than estimates for White infants.

Using counties in other Southern states ineligible for Duke funding as controls: We address concerns about selection into treatment and spillover effects with a variety of analyses that add other non-Carolina Southern counties. For example, North Carolina counties in the control group may have been ill-prepared to take advantage of Duke funding or simply may not have had any not-for-profit hospitals that could be supported. In addition, higher-quality physicians could have moved from control to treated counties or the state government could have redirected resources from counties that did not get Duke support to those that did (or vice versa). In the presence of such effects, our difference-in-differences estimates may overstate the benefits of the intervention. Thus, we define alternate sets of control-group counties that should not have been affected by any spillovers or selection.

Up to the 1930s, infant mortality rates were relatively similar across ever-treated North Carolina counties and other Southern counties that were ineligible for Duke funding because they were outside of the Carolinas (See Online Appendix Figure J1). However, as Duke support expanded over the remainder of the decade, North Carolina experienced a large decline in infant mortality for both White and Black infants that was not observed in other Southern states. Our results are robust to alternative DiD (Online Appendix Table F3) and event study (Online Appendix Figure J2) specifications that use these non-Carolina Southern counties as controls, although these estimates suggest much more similar effects by race.

We also use propensity score matching to construct an alternative set of control counties from the Southern states and to conduct a placebo analysis. First, we only select controls counties from other Southern states that are observably similar to treated North Carolina counties. Naturally, these control counties are ineligible for Duke founding. These results are presented in Online Appendix Tables M2 to M4 and are similar to those in Online Appendix Table F3. Second, our placebo exercise uses only non-Carolina Southern counties that we divide into those that look the most and the least like our true treated counties. We then assign treatment to the former group and estimate our preferred infant mortality regressions. As expected, results presented in Online Appendix Table M1 yield small, inconsistently signed, and almost never statistically significant coefficients.

Instrumental variables: As another layer of evidence showing that our findings are not driven by selection, we consider a version of our intensive-margin treatment effects in an instrumental variables setting. The instrument interacts two exogenous sources of variation. The first source is temporal and leverages the fact that as income from the Endowment's investments increased,

so did the ability of the Endowment to appropriate and spend funds. The second source is cross-sectional and exploits the fact that North Carolina counties with a not-for-profit hospital at the time the Endowment was founded had an advantage over other counties because they were eligible for capital appropriations to improve existing hospitals.

Compared to our preferred intensive margin results, the instrumental variables framework yields larger point estimates that are statistically significant at conventional levels (See Online Appendix Table J1). We report both Anderson and Rubin (1949) 95% confidence sets and tF 95% confidence intervals (Lee et al. 2022) for each IV estimate, demonstrating that the strength of our first stage is not an issue. Results suggest that for every one million dollars that The Duke Endowment appropriated in a county, the infant mortality rate declined between 14.6 and 17.4%. For payments, the effect sizes are between 16.2 and 17.4% per one million dollars. If selection into treatment, targeting of funds, or trends in counties with not-for-profit hospitals were driving our results, we would expect the instrumental variable results to attenuate these treatment effect estimates (See Online Appendix J for additional details). The finding of a larger, negative effect mitigates these concerns and provides additional evidence in favor of our overall argument.

Randomization inference: Next, we explore whether the non-random selection of counties that received Duke support or the non-random timing of treatment may have distorted our main estimates for infant mortality. In Online Appendix Figure K1, we present the results from a randomization inference exercise based on the specification from column 5 of Table 3). In this procedure, we randomly assign 48 North Carolina counties to the treatment path of the 48 counties that actually received treatment. As expected under the assumption of no selection bias, the average placebo estimate is symmetrical and centered around zero.

We also use this exercise to construct two-sided randomization inference based p-values for our preferred point estimates.³⁴ Our results remain statistically significant, indicating that they are not driven by spurious time trends in infant mortality across all counties or by spillovers across treated versus untreated counties in North Carolina.

Long-run mortality Our long-run mortality results are unaffected if we collapse the data to the cohort-level and use cumulative mortality at ages 56 to 64 as the outcome rather than including separate observations for mortality at each follow-up age (Online Appendix Table F2). They are robust to including non-Carolina counties that were ineligible for Duke funding as controls (Online Appendix Table F3) and adding South Carolina to the treatment group (Online Appendix Table F4).

³⁴We do so by comparing the absolute value of each treatment effect estimate to the absolute value of pseudo-treatment effect estimates drawn from 10,000 different randomly drawn treatment assignments. The p-value is the percent of random draws that are larger in absolute value than the estimated treatment effect.

7 Conclusions

The barriers that limit access to effective healthcare in the US today – including a lack of healthcare providers and facilities in rural and impoverished areas, financial constraints, unequal access based on race and socioeconomic status, inadequate insurance coverage, and structural obstacles such as transportation difficulties – are not unique to the modern era. These challenges were equally prevalent almost 100 years ago in 1924, when James Buchanan Duke established The Duke Endowment with the intention of alleviating such burdens in the Carolinas.

In this paper, we leverage exogenous variation in support provided by the foundation – including the building of new hospitals, expansions and upgrades of existing facilities, modernization of equipment, and improved managerial practices – to study its consequences on infant and long-run mortality, with a particular focus on understanding racial disparities in health outcomes.

Our findings reveal that The Duke Endowment’s support had a substantial impact on the healthcare infrastructure of North Carolina. It led to an expansion in the number of not-for-profit hospitals and an increase in available beds. Furthermore, it significantly increased the number of practicing physicians in treated counties, effectively replacing less-qualified doctors with highly-trained counterparts. Furthermore, we show that counties supported by Duke disproportionately benefited from the invention of sulfa drugs, providing some of the first causal evidence that hospital funding and medical innovation are complements rather than substitutes.

As a result of healthcare infrastructure improvements facilitated by The Duke Endowment’s support, infant mortality decreased by 7.5%. This impact was more pronounced for Black infants, who experienced a 13.6% reduction in mortality compared to 4.7% decrease for White infants. These gains for Black infants succeeded in narrowing the pre-intervention racial gap in infant mortality by one-quarter. Furthermore, we provide evidence that the health benefits of support from The Duke Endowment persisted until at least modern retirement age for those exposed around the time of birth, reducing the average mortality between the ages 56 and 64 by 9.0%.

We calculate that each infant life was saved at a remarkably low cost of \$17,092 (2017 dollars). However, it is important to note that in today’s context, where baseline health and healthcare access are considerably higher, a similar intervention would likely yield a smaller impact at a higher cost. Nevertheless, given the ongoing challenges in the provision of healthcare to rural areas, these results highlight the potential for investments in hospitals to still yield significant and cost-effective impacts.

Although the intervention we study occurred in the first-half of the twentieth century in the US, we view our findings as not only contributing to expanding basic knowledge about the development of US healthcare, but also to historical and contemporaneous public policy. Despite the viewpoint that modern medicine did not materially affect health outcomes in the US until the 1950s, our

results are *proof of concept* that investments in for-the-times modern hospital infrastructure were indeed productive in both the short- and long-run. Furthermore, what was considered a modern hospital in the US during the 1930s and 1940s pales in comparison to the currently available technology, but it may not be that dissimilar to what is offered in less-developed nations and especially in under-served and rural regions of these countries. To the extent that our findings generalize to such settings, they may guide policymakers in improving infant and adult health in a cost-effective way.

The Duke Endowment's financial support for healthcare in the Carolinas also *served as both inspiration and model* for the Hill-Burton Act of 1946 (Durdin 1998), which was a federal program that subsidized large-scale construction of hospitals in under-served locations throughout the US to ensure that each state had an average of 4.5 hospital beds per 1,000 residents. Over the life of the program, the federal government spent over \$3.7 billion on that goal (Chung et al. 2017), but very little is known about its effects on health or human capital outcomes of individuals that it touched. While Duke support was the exception rather than the norm during the first-half of the twentieth century, and thus likely could not have contributed to the aggregate trends in mortality, it appears plausible that a non-trivial fraction of the massive decline in infant mortality in the US between 1945 and 1975 (from 44 to 18 deaths per 1,000 live births) might be attributable to this intervention given that the Hill-Burton program was modelled on The Duke Endowment. We view this as both a policy-relevant and promising agenda for future research, which could also inform questions about the scale-up of smaller *trial programs* to the national level.

Finally, it is important to point out that we focus on a very specific research question and conduct analysis that is limited in part by data availability. Since we find large and positive health effects that persist in the long-run across the Carolinas, it is possible that other economic outcomes of the treated individuals, such as human capital or wages, also improved. Furthermore, since the support from The Duke Endowment was not targeted at pregnant women or infants, and because it induced permanent changes in access to modern hospitals in supported communities, it plausibly benefited other groups beyond those considered in this paper. This means that our estimates are likely to be a conservative lower bound for the total welfare improvements. We view the questions about the cumulative effects of the program, effects on other outcomes, and impacts on other age groups as important future extensions of this work.

References

- Aaronson, Daniel, Fabian Lange, and Bhashkar Mazumder (2014) “Fertility Transitions Along the Extensive and Intensive Margins,” *American Economic Review*, Vol. 104, No. 11, pp. 3701–3724.
- Aaronson, Daniel and Bhashkar Mazumder (2011) “The Effect of Rosenwald Schools on Black Achievement,” *Journal of Political Economy*, Vol. 119, No. 5, pp. 821–888.
- Almond, Douglas, Janet Currie, and Valentina Duque (2018) “Childhood Circumstances and Adult Outcomes: Act II,” *Journal of Economic Literature*, Vol. 56, No. 4, pp. 1360–1446.
- Alsan, Marcella and Claudia Goldin (2019) “Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880-1920,” *Journal of Political Economy*, Vol. 127, No. 2, pp. 586–638.
- American Medical Association (1942a) “Hospital Service in the United States, 1920-1942,” in *Journal of the American Medical Association*, Chicago, IL: American Medical Association.
- (1942b) *American medical directory, 1906-1942*, Chicago, IL: American Medical Association.
- Anderson, D. Mark, Ryan Brown, Kerwin Kofi Charles, and Daniel I. Rees (2020) “Occupational Licensing and Maternal Health: Evidence from Early Midwifery Laws,” *Journal of Political Economy*, Vol. 128, No. 11, pp. 4337–4383.
- Anderson, D. Mark, Kerwin Kofi Charles, and Daniel I. Rees (2021a) “The Federal Effort to Desegregate Southern Hospitals and the Black-White Infant Mortality Gap,” NBER Working Paper No. 27970.
- (2022) “Reexamining the Contribution of Public Health Efforts to the Decline in Urban Mortality,” *American Economic Journal: Applied Economics*, Vol. 14, No. 2, pp. 126–157.
- Anderson, D. Mark, Kerwin Kofi Charles, Daniel I. Rees, and Tianyi Wang (2021b) “Water Purification Efforts and the Black-white Infant Mortality Gap, 1906–1938,” *Journal of Urban Economics*, Vol. 122, p. 103329.
- Anderson, T. W. and Herman Rubin (1949) “Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations,” *The Annals of Mathematical Statistics*, Vol. 20, No. 1, pp. 46 – 63.
- Arias, E. (2014) “United States life tables, 2010,” *National vital statistics reports*, Vol. 63, No. 7.
- Arias, E., B. Tejada-Vera, F. Ahmad, and K.D. Kochanek (2021) “Provisional life expectancy estimates for 2020,” *Vital Statistics Rapid Release*, No. 15, pp. 50–61.
- Arthi, Vellore, Brian Beach, and W. Walker Hanlon (2022) “Recessions, Mortality, and Migration Bias: Evidence from the Lancashire Cotton Famine,” *American Economic Journal: Applied Economics*, Vol. 14, No. 2, pp. 228–55.
- Ashenfelter, Orley and Michael Greenstone (2004) “Using Mandated Speed Limits to Measure the Value of a Statistical Life,” *Journal of Political Economy*, Vol. 112, No. S1, pp. S226–S267.
- Bailey, Martha, Karen Clay, Price Fishback, Michael Haines, Shawn Kantor, Edson Severnini, and Anna Wentz (2015) *U.S. County-Level Natality and Mortality Data, 1915-2007*: Inter-university Consortium for Political and Social Research [distributor].
- Baker, Andrew., David F. Larker, and Charles C.Y. Wang (2022) “How Much Should We Trust Staggered Difference-in-Differences Estimates?” *Journal of Financial Economics*, Vol. 144, No. 2, pp. 370–395.
- Black, Dan A., Seth G. Sanders, Evan J. Taylor, and Lowell J. Taylor (2015) “The Impact of the Great Migration on Mortality of African Americans: Evidence from the Deep South,” *American Economic Review*, Vol. 105, No. 2, pp. 477–503.
- Bleakley, Hoyt (2007) “Disease and Development: Evidence from Hookworm Eradication in the American South,” *Quarterly Journal of Economics*, Vol. 122, No. 1, pp. 73–117.
- Bloom, N. and J. Van Reenen (2007) “Measuring and Explaining Management Practices Across Firms and Countries,” *The Quarterly Journal of Economics*, Vol. 122, No. 4, pp. 1351–1408.
- Callaway, Brantly and Pedro H.C. Sant’Anna (2021) “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, Vol. 225, No. 2, pp. 200–230.
- Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019) “The Effect of Minimum Wages on Low-Wage Jobs,” *The Quarterly Journal of Economics*, Vol. 134, No. 3, pp. 1405–1454.
- Chen, Jiafeng and Jonathan Roth (2023) “Logs with zeros? Some problems and solutions,” arXiv:2212.06080.
- Chung, Andrea Park, Martin Gaynor, and Seth Richards-Shubik (2017) “Subsidies and Structure: The Lasting Impact of the Hill-Burton Program on the Hospital Industry,” *The Review of Economics and Statistics*, Vol. 99, No. 5, pp. 926–943.

- CMS (2022) “NHE Fact Sheet | CMS,” U.S. Centers for Medicare & Medicaid Services.
- Collins, William J. and Melissa A. Thomasson (2004) “The Declining Contribution of Socioeconomic Disparities to the Racial Gap in Infant Mortality Rates, 1920–1970,” *Southern Economic Journal*, Vol. 70, No. 4, pp. 746–776.
- Cook, Lisa D., Trevon D. Logan, and John M. Parman (2014) “Distinctively Black Names in the American Past,” *Explorations in Economic History*, Vol. 53, pp. 64–82.
- (2016) “The Mortality Consequences of Distinctively Black Names,” *Explorations in Economic History*, Vol. 59, pp. 114–125.
- Correia, Sergio, Paulo Guimarães, and Tom Zylkin (2020) “Fast Poisson Estimation with High-Dimensional Fixed Effects,” *The Stata Journal*, Vol. 20, No. 1, pp. 95–115.
- Costa, Dora L. (2015) “Health and the Economy in the United States from 1750 to the Present,” *Journal of Economic Literature*, Vol. 53, No. 3, pp. 503–70.
- Craig, Lee (2006) “Consumer expenditures,” in Susan B. Carter, Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright eds. *Historical Statistics of United States: Millennial Edition*, vol. 3, Cambridge: Cambridge University Press.
- Currie, Janet and Jonathan Gruber (1996) “Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women,” *Journal of Political Economy*, Vol. 104, No. 6, pp. 1263–1296.
- Cutler, David M. (2005) *Your Money or Your Life: Strong Medicine for America’s Health Care System*, Oxford New York: Oxford University Press.
- Durbin, Robert F. (1998) *Lasting Legacy to the Carolinas*, Durham, NC: Duke University Press.
- Elder, Todd E., John H. Goddeeris, and Steven J. Haider (2016) “Racial and Ethnic Infant Mortality Gaps and the Role of Socio-Economic Status,” *Labour Economics*, Vol. 43, pp. 42–54.
- Esteves, Rui, Kris James Mitchener, Peter Nencka, and Melissa A Thomasson (2022) “Do Pandemics Change Healthcare? Evidence from the Great Influenza,” Working Paper 30643, National Bureau of Economic Research.
- Ewbank, Douglas (1987) “History of Black Mortality and Health before 1940,” *Milbank Quarterly*, Vol. 65, No. 1, pp. 100–128.
- Ferrell, J. Atkinson and P. A. Mead (1936) *History of county health organizations in the United States, 1908–33*, Washington, DC: U.S. Government Printing Office.
- Figinski, Theodore F. and Erin Troland (2020) “Health Insurance and Hospital Supply: Evidence from 1950s Coal Country,” *Finance and Economics Discussion Series*, Vol. 2020–03, No. 033.
- Fischer, Martin, Martin Karlsson, and Nikolaos Prodrromidis (2021) “The Long-term Effects of Hospital Deliveries,” IZA Discussion Paper No. 14562.
- Fischer, Stephanie, Heather Royer, and Corey White (2022) “Health Care Centralization: The Health Impacts of Obstetric Unit Closures in the US,” NBER Working Paper No. 30141.
- Fishback, Price V., William C. Horrall, and Shawn Kantor (2005) “Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales During the Great Depression,” *The Journal of Economic History*, Vol. 65, No. 1, p. 36–71.
- Gamble, Vanessa Northington (1995) *Making a Place for Ourselves: The Black Hospital Movement, 1920–1945*, New York: Oxford University Press.
- (1997) “Black Autonomy versus White Control: Black Hospitals and the Dilemmas of White Philanthropy, 1920–1940,” *Minerva*, Vol. 35, No. 3, pp. 247–267.
- Germack, Hayley Drew, Ryan Kandrack, and Grant Martsolf (2019) “When Rural Hospitals Close, the Physician Workforce Goes,” *Health Affairs*, Vol. 38, No. 12, pp. 2086–2094.
- Goldstein, Joshua R., Monica Alexander, Casey Breen, Andrea Miranda González, Felipe Menares, Maria Osborne, Mallika Snyder, and Ugur Yildirim (2021) *CenSoc Mortality File: Version 2.0*, Berkeley, CA: University of California.
- Goodman-Bacon, Andrew (2018) “Public Insurance and Mortality: Evidence from Medicaid Implementation,” *Journal of Political Economy*, Vol. 126, No. 1, pp. 216–262.
- (2021a) “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, Vol. 225, No. 2, pp. 254–277.
- (2021b) “The Long-run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health and Labor Market Outcomes,” *American Economic Review*, Vol. 111, No. 8, pp. 2550–

2593.

- Haines, Michael (2008) "Fertility and Mortality in the United States," in Robert Whaples ed. *EH.Net Encyclopedia*: Economic History Association.
- Health Insurance Institute (1977) *Source Book of Health Insurance Data, 1976-1977*: Health Insurance Institute.
- Hjort, Jonas, Mikkel Sølvsten, and Miriam Wüst (2017) "Universal Investment in Infants and Long-run Health: Evidence from Denmark's 1937 Home Visiting Program," *American Economic Journal: Applied Economics*, Vol. 9, No. 4, pp. 78–104.
- Hoehn-Velasco, Lauren (2018) "Explaining Declines in US Rural Mortality, 1910-1933: The Role of County Health Departments," *Explorations in Economic History*, Vol. 70, pp. 42–72.
- (2021) "The Long-term Impact of Preventative Public Health Programs," *Economic Journal*, Vol. 131, No. 634, pp. 797–826.
- IPUMS USA (2015) "ICPSR County Codes," <https://usa.ipums.org/usa/volii/ICPSR.shtml>, Accessed August 28, 2023.
- (2021) "PLACENHG Code List," https://usa.ipums.org/usa/resources/volii/ipums_placenhg_code_list.xlsx, Accessed September 5, 2023.
- Jayachandran, Seema, Adriana Lleras-Muney, and Kimberly Smith (2010) "Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs," *American Economic Journal: Applied Economics*, Vol. 2, No. 2, pp. 118–146.
- Kozhimannil, Katy, Peiyin Hung, Carrie Henning-Smith, Michelle Casey, and Shailendra Prasad (2018) "Association Between Loss of Hospital-Based Obstetric Services and Birth Outcomes in Rural Counties in the United States," *JAMA*, Vol. 319, No. 12, pp. 1239–1247.
- Lazuka, Volha (2020) "Infant Health and Later-life Labor Market Outcomes: Evidence from the Introduction of Sulfa Antibiotics in Sweden," *Journal of Human Resources*, Vol. 55, No. 2, pp. 660–698.
- Lechner, Michael (2011) "The Estimation of Causal Effects by Difference-in-Difference Methods," *Foundations and Trends in Econometrics*, Vol. 4, No. 3, pp. 165–224.
- Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack Porter (2022) "Valid t-Ratio Inference for IV," *American Economic Review*, Vol. 112, No. 10, pp. 3260–90.
- Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles (2022) *IPUMS National Historical Geographic Information System: Version 17.0 [Dataset]*, Minneapolis, MN: IPUMS.
- Moehling, Carolyn M., Gregory T. Niemesh, Melissa A. Thomasson, and Jaret Treber (2020) "Medical Education Reforms and the Origins of the Rural Physician Shortage," *Cliometrica*, Vol. 14, No. 2, pp. 181–225.
- Moehling, Carolyn and Melissa Thomasson (2014) "Saving Babies: The Impact of Public Education Programs on Infant Mortality," *Demography*, Vol. 51, No. 2, pp. 367–386.
- Mullahy, John and Edward C Norton (2022) "Why Transform Y? A Critical Assessment of Dependent-Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes," Working Paper 30735, National Bureau of Economic Research.
- National Center for Health Statistics (2017) "Health, United States, 2017: Table 015," Hyattsville, MD.
- North Carolina Bureau of Vital Statistics (1922) *Annual report of the Bureau of Vital Statistics of the North Carolina State Board of Health, 1922-1948*, Raleigh, NC: Bureau of Vital Statistics of the North Carolina State Board of Health.
- (1933) *Annual report of the Bureau of Vital Statistics of the North Carolina State Board of Health, 1933*, Raleigh, NC: Bureau of Vital Statistics of the North Carolina State Board of Health.
- Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt (2019) "Poorly Measured Confounders Are More Useful on the Left than on the Right," *Journal of Business & Economic Statistics*, Vol. 37, No. 2, pp. 205–216.
- Pollitt, Phoebe Ann (2017) *African American Hospitals in North Carolina: 39 Institutional Histories, 1880-1967*, Jefferson, NC: McFarland.
- Roth, Jonathan, Pedro Sant'Anna, Alyssa Bilinski, and John Poe (2022) "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," Working Paper.
- Santos Silva, J.M.C. and Silvana Tenreiro (2011) "Further Simulation Evidence on the Performance of the Poisson Pseudo-maximum Likelihood Estimator," *Economics Letters*, Vol. 112, No. 2, pp. 220–222.
- Soares, Rodrigo R. (2007) "On the Determinants of Mortality Reductions in the Developing World," *Popu-*

- lation and Development Review*, Vol. 33, No. 2, pp. 247–287.
- Social Security Administration (2007) *Numerical Identification Files (NUMIDENT)*, created, 1936 - 2007, documenting the period 1936 - 2007 [dataset], Washington, DC: Social Security Administration.
- Starr, Paul (2017) *The Social Transformation of American Medicine*, New York: Basic Books, updated edition.
- Surveillance, Epidemiology, and End Results Program (2022) “County-Level Population Files - Single-year Age Groups,” <https://seer.cancer.gov/popdata/download.html>, Accessed June 29, 2022.
- The Duke Endowment (1925) *Annual Report of the Hospital Section of the Duke Endowment for the year 1925*: The Duke Endowment.
- (1925b) *Yearbook of the Duke Endowment*: The Duke Endowment.
- (1934) *Annual Report of the Hospital Section of the Duke Endowment for the year 1934*: The Duke Endowment.
- Thomas, Karen Kruse (2011) *Deluxe Jim Crow: Civil Rights and American Health Policy, 1935-1954*, Athens: University of Georgia Press.
- Thomasson, Melissa A (2002) “From sickness to health: the twentieth-century development of US health insurance,” *Explorations in Economic History*, Vol. 39, No. 3, pp. 233–253.
- Thomasson, Melissa and Jaret Treber (2008) “From Home to Hospital: The Evolution of Childbirth in the United States, 1928-1940,” *Explorations in Economic History*, Vol. 45, pp. 76–99.
- US Geological Survey (2019) *Geographic Names Information System (GNIS)*: US Geological Survey.
- Viscusi, Kip (2018) “Best Estimate Selection Bias in the Value of a Statistical Life,” *Journal of Benefit-Cost Analysis*, Vol. 9, No. 2, pp. 205–246.
- Wooldridge, Jeffrey M. (2021) “Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators,” Working Paper.
- Yamey, Gavin, Dean Jamison, Odd Hanssen, and Agnes Soucat (2019) “Financing Global Common Good for Health: When the World is a Country,” *Health Systems and Reform*, Vol. 5, No. 3, pp. 334–349.

8 Tables

Table 1. First-stage effects of Duke support on hospitals and beds: Extensive margin intent-to-treat estimates

	$Y_{ct}^R = \text{Beds or Hospitals}$			$Y_{ct}^R = \text{Beds or Hospitals per 1,000 births}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Beds</i>						
Total	35.46*** (9.91)	44.72*** (13.36)	32.63*** (10.79)	31.11*** (6.13)	27.48*** (6.62)	24.57*** (5.49)
Non-profit/church/public	39.06*** (10.10)	47.71*** (14.10)	35.97*** (11.21)	35.83*** (6.26)	30.46*** (6.90)	28.17*** (5.86)
Proprietary	-4.63* (2.38)	-5.19 (3.24)	-5.72* (3.11)	-4.97* (2.63)	-4.35* (2.51)	-4.89* (2.59)
<i>B. Hospitals</i>						
Total	0.18* (0.10)	0.11 (0.12)	0.10 (0.11)	0.31* (0.18)	0.15 (0.12)	0.20* (0.11)
Non-profit/church/public	0.31*** (0.10)	0.21 (0.14)	0.22* (0.12)	0.53*** (0.16)	0.25** (0.11)	0.33*** (0.12)
Proprietary	-0.14* (0.07)	-0.14* (0.08)	-0.17** (0.08)	-0.20 (0.12)	-0.12 (0.08)	-0.14* (0.08)
Observations	2,100	2,100	2,100	2,100	2,100	2,100
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate two-way fixed effects regression. In columns 1 to 3, the dependent variable is the number of hospital beds (Panel A) or the number of hospitals (Panel B) in a county and year. Within Panel A, the dependent variable is the number of beds in hospitals of any type (top row); in non-profit, church-owned, or public hospitals (middle row); or in proprietary hospitals (bottom row). Panel B presents the same classification across rows with the number of hospitals of each type as the dependent variable. Columns 4 to 6 express the dependent variables as rates per 1,000 live births and follows the same structure as columns 1 to 3. Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. In columns 2 to 3 and 5 to 6 the weights are the number of births in a county and year. Standard errors are clustered at the county level.

Table 2. First-stage effects of Duke support on number of doctors: Extensive margin intent-to-treat estimates

	$Y_{ct}^R = \text{Doctors}$			$Y_{ct}^R = \text{Doctors per 1,000 live births}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
All	4.63** (2.08)	8.06** (3.26)	5.64** (2.47)	1.28 (1.29)	3.40** (1.44)	2.83** (1.23)
High quality	7.87*** (2.32)	12.22*** (3.45)	9.02*** (2.58)	3.05** (1.22)	4.80*** (1.44)	3.88*** (1.23)
Low quality	-3.13*** (1.00)	-4.16*** (1.34)	-3.35*** (1.11)	-1.70** (0.73)	-1.40* (0.73)	-1.02 (0.69)
Observations	1,100	1,100	1,100	1,100	1,100	1,100
<i>B. Black</i>						
All	0.41* (0.21)	0.79* (0.44)	0.57* (0.31)	0.02 (0.90)	1.17* (0.65)	0.80 (0.57)
High quality	0.74*** (0.24)	1.60*** (0.44)	1.34*** (0.37)	1.36*** (0.47)	2.53*** (0.67)	2.31*** (0.59)
Low quality	-0.33** (0.15)	-0.81*** (0.27)	-0.77*** (0.22)	-1.34 (0.86)	-1.37*** (0.51)	-1.51*** (0.55)
Observations	1,078	1,078	1,078	1,078	1,078	1,078
<i>C. White</i>						
All	4.21** (1.93)	6.77** (2.80)	4.38** (1.98)	-0.52 (2.40)	3.60* (2.09)	2.64 (1.72)
High quality	6.97*** (2.10)	9.87*** (2.96)	6.71*** (2.00)	1.58 (2.20)	4.25** (1.99)	2.79* (1.61)
Low quality	-2.65*** (0.94)	-3.11** (1.36)	-2.31** (1.07)	-1.91* (1.14)	-0.67 (1.04)	-0.11 (0.94)
Observations	1,078	1,078	1,078	1,078	1,078	1,078
County FE	Yes	Yes	Yes	Yes	Yes	Yes
AMD Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate two-way fixed effects regression. In columns 1 to 3, the dependent variable is the number of doctors (Panel A), the number of Black doctors (Panel B), or the number of White doctors (Panel C) in a county and publication year of the *American Medical Directory* (AMD wave). During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. Within each panel, the dependent variable is the number of doctors (top row); the number of high-quality doctors (middle row); or the number of low-quality doctors (bottom row). A high-quality doctor is one who graduated from a medical school at least 4 years after it introduced a two-year college degree prerequisite for admission. All other doctors are considered low quality. Columns 4 to 6 express the dependent variables as rates per 1,000 live births and follows the same structure as columns 1 to 3. Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. In columns 2 to 3 and 5 to 6 the weights are the average number of births in each county for the years in the AMD wave. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942 for any AMD wave. Standard errors are clustered at the county level.

Table 3. Effect of Duke support on infant deaths and infant mortality rate: Extensive margin intent-to-treat estimates

	$Y_{ct}^R = \text{Infant deaths}$			$Y_{ct}^R = \text{Infant mortality rate}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled infant deaths or infant mortality rate</i>						
Percent effect from Duke (=1)	-6.12** (2.34)	-6.05** (2.56)	-6.51*** (2.47)	-8.09*** (2.51)	-8.11*** (2.36)	-7.49*** (2.30)
Observations	2,100	2,100	2,100	2,100	2,100	2,100
<i>B. Black infant deaths or infant mortality rate</i>						
Percent effect from Duke (=1)	-9.30*** (3.52)	-8.50** (3.54)	-11.37*** (3.29)	-7.97 (7.59)	-12.15*** (3.46)	-13.61*** (3.19)
Observations	1,995	1,995	1,995	1,995	1,995	1,995
<i>C. White infant deaths or infant mortality rate</i>						
Percent effect from Duke (=1)	-4.44 (2.94)	-4.25 (3.73)	-4.14 (3.02)	-8.00** (3.49)	-5.95** (2.98)	-4.71 (2.93)
Observations	1,995	1,995	1,995	1,995	1,995	1,995
P-value for difference by race	0.25	0.40	0.09	1.00	0.13	0.02
County of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. In columns 1 to 3, the dependent variable is the number of infant deaths in a county and year. In columns 4 to 6, it is the infant mortality rate per 1,000 live births. Each coefficient represents the percent reduction in infant mortality due to receiving a capital appropriation from The Duke Endowment. A coefficient of -10 would mean that infant mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. In columns 2 to 3 and 5 to 6 the weights are the number of births in a county and year. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942. The bottom row presents a p-value from the interaction of race with our treatment variable from a model that fully interacts all variables with race. Standard errors are estimated using the delta method and are clustered at the county level.

Table 4. Effect of Duke support on infant mortality rate: Intensive margin estimates

	$Y_{ct}^R = \text{Infant mortality rate}$		$Y_{ct}^R = \text{Cumulative payments}$
	Appropriations (1)	Payments (2)	Appropriations (3)
<i>A. Pooled infant mortality rate</i>			
Percent effect from \$1 million of Duke support	-7.16*** (0.97)	-7.71*** (1.18)	
Cumulative appropriations			0.87*** (0.04)
Observations	2,100	2,100	2,100
<i>B. Black infant mortality rate</i>			
Percent effect from \$1 million of Duke support	-9.05*** (1.66)	-8.91*** (1.56)	
Cumulative appropriations			0.87*** (0.03)
Observations	1,995	1,995	1,995
<i>C. White infant mortality rate</i>			
Percent effect from \$1 million of Duke support	-6.51*** (1.20)	-7.64*** (1.63)	
Cumulative appropriations			0.86*** (0.04)
Observations	1,995	1,995	1,995
P-value for difference by race	0.34	0.72	0.35
County of birth FE	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Weights	Yes	Yes	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimates from econometric specifications based on column 6 of Table 3. The first two columns test for an intensive margin effect, with the independent variable of interest being cumulative appropriations or cumulative payments to hospitals rather than a binary variable for ever having had an appropriation from The Duke Endowment. The dependent variable in columns 1 and 2 is the infant mortality rate per 1,000 live births. The third column looks at the relationship between cumulative appropriations and cumulative payments. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942. The bottom row presents a p-value from the interaction of race with our treatment variable from a model that fully interacts all variables with race. Standard errors are estimated using the delta method and are clustered at the county level.

Table 5. Interaction of rollout of Duke support and discovery of sulfa drugs

	Y_{ct}^R = Infant mortality rate		
	Pooled	Black	White
	(1)	(2)	(3)
<i>A. Interaction of Duke rollout \times sulfa shift-share DiD</i>			
Post-pre sulfa Pneumonia _{.75} – Post-pre sulfa Pneumonia _{.25} (γ_5)	–15.09*** (4.94)	–12.57* (6.54)	–16.20*** (5.32)
<i>B. Duke vs. no Duke</i>			
Post-Sulfa, Pneumonia _{.75} $\gamma_1 + \gamma_2 + \eta_{.75} \times (\gamma_3 + \gamma_5)$	–15.77*** (3.96)	–19.44*** (5.58)	–14.82*** (4.07)
Post-Sulfa, Pneumonia _{.25} $\gamma_1 + \gamma_2 + \eta_{.25} \times (\gamma_3 + \gamma_5)$	1.51 (7.05)	–4.07 (9.29)	2.47 (7.43)
Pre-Sulfa, Pneumonia _{.75} $\gamma_1 + \eta_{.75} \times \gamma_3$	–5.57** (2.37)	–11.60*** (2.75)	–2.72 (3.34)
Pre-Sulfa, Pneumonia _{.25} $\gamma_1 + \eta_{.25} \times \gamma_3$	–3.37 (3.29)	–7.96 (4.81)	–1.93 (4.15)
<i>C. (Duke vs. no Duke) \times (Pre vs. post sulfa) or (Duke vs. no Duke) \times (Pneumonia IQR)</i>			
Pneumonia _{.75} , Post-pre Sulfa $\gamma_2 + \eta_{.75} \times \gamma_5$	–10.81** (4.18)	–8.87 (5.52)	–12.44*** (4.51)
Pneumonia _{.25} , Post-pre Sulfa $\gamma_2 + \eta_{.25} \times \gamma_5$	5.05 (7.07)	4.23 (7.90)	4.49 (8.20)
Post-Sulfa, Pneumonia _{IQR} $(\eta_{.75} - \eta_{.25}) \times (\gamma_3 + \gamma_5)$	–17.02*** (5.77)	–16.02** (7.76)	–16.88*** (5.63)
Pre-Sulfa, Pneumonia _{IQR} $(\eta_{.75} - \eta_{.25}) \times \gamma_3$	–2.27 (2.88)	–3.95 (3.59)	–0.81 (3.08)
Observations	2,100	1,995	2,100

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimates in each column come from separate Poisson pseudo-maximum likelihood regressions. Each panel reports exponentiated linear combinations of coefficients from Equation 5. Because we report exponentiated linear combinations of the underlying coefficients, linear combinations of this non-linear transform may not be equal to the reported coefficient due to Jensen's inequality (i.e., $\exp(a + b) \leq \exp(a) + \exp(b)$). Column 2 drops the five counties that ever have zero Black births between 1922 and 1942. The dependent variable is the number of infant deaths per 1,000 live births in county c and year t birth cohort. Duke exposure is a binary variable that takes the value of one in all years starting in the first year that a county received a capital appropriation from the Endowment. Sulfa exposure (shift) is measured by an indicator that equals to one for all years from 1937 onward. The share factor is the average pneumonia mortality per 100,000 individuals in a county between 1922 and 1926. Pneumonia_{IQR}($\eta_{.75} - \eta_{.25}$) denotes the difference between the 75th and 25th percentiles of baseline pneumonia mortality. All regressions include county and year fixed effects. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Observations are weighted by the number of births in a county and year. Standard errors are estimated using the delta method and are clustered at the county level.

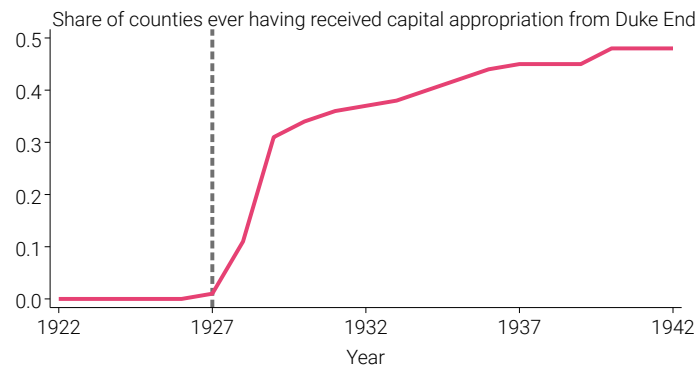
Table 6. Effect of Duke support around time of birth on long-run mortality at ages 56 to 64

	$Y_{ct}^R = \text{Long-run deaths}$		
	(1)	(2)	(3)
<i>A. Pooled long-run deaths</i>			
Percent effect from Duke (=1)	-7.66** (3.08)	-10.07*** (2.66)	-8.99*** (2.81)
Observations	9,000	9,000	9,000
<i>B. Black long-run deaths</i>			
Percent effect from Duke (=1)	-8.04** (3.57)	-8.77*** (2.80)	-7.58** (3.54)
Observations	8,150	8,150	8,150
<i>C. White long-run deaths</i>			
Percent effect from Duke (=1)	-6.93* (3.83)	-10.56*** (2.92)	-9.53*** (2.86)
Observations	8,630	8,630	8,630
P-value for difference by race	0.82	0.61	0.63
County of birth X Age FE	Yes	Yes	Yes
Year of birth X Age FE	Yes	Yes	Yes
Weights	No	Yes	Yes
Controls	No	No	Yes

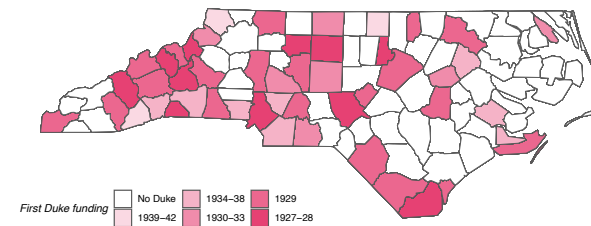
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The unit of observation is a birth county by birth year by follow-up age triplet. Birth cohorts are restricted to 1932 to 1941. Deaths are restricted to ages 56 to 64 and years 1988 to 2005. The dependent variable is the number of age-specific deaths. Each coefficient represents the percent reduction in long-run mortality due to receiving a capital appropriation from The Duke Endowment around the time of birth. A coefficient of -10 would mean that long-run mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Control variables in column 3 include flexible interactions of age fixed effects with % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. In columns 2 to 3, the weights are county-by-year birth cohort size. All specifications include county-of-birth fixed effects interacted with follow-up age fixed effects and year-of-birth fixed effects interacted with follow-up age fixed effects. Panels B and C drop counties that ever have zero race-specific births between 1932 and 1941. Observations differ across the two samples as observations are dropped that are perfectly separated by either county-of-birth by follow-up age fixed effects or year-of-birth by follow-up age fixed effects. The bottom row presents a p-value from the interaction of race with our treatment variable from a model that fully interacts all variables with race. Standard errors are estimated using the delta method and are clustered at the county level.

9 Figures

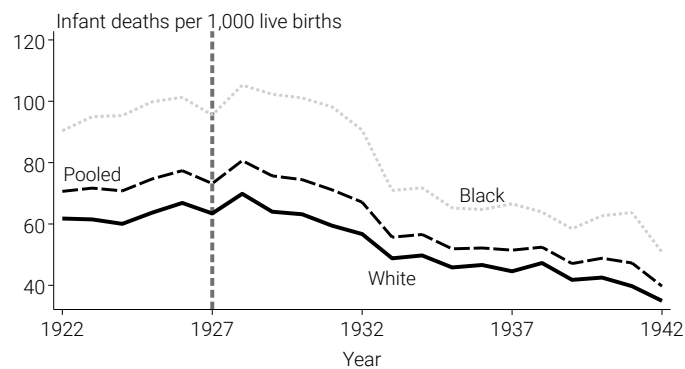
Figure 1. Rollout of Duke support and changes in infant mortality rate, 1922-1942



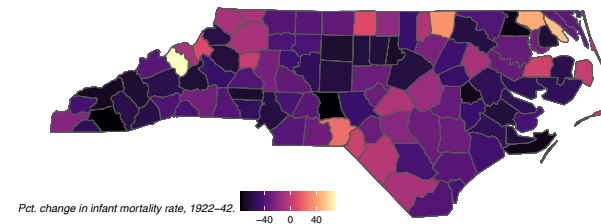
(a) Share of counties treated by a given year



(c) Duration of exposure to Duke support by county



(b) Infant mortality rate by year and race

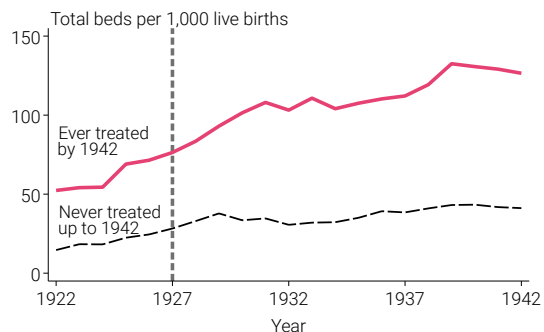


(d) Change in infant mortality rate, 1922-1942

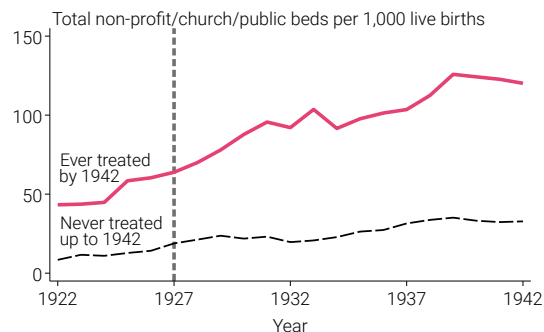
Notes: Figure 1a plots the fraction of ever-treated counties in North Carolina by calendar year. Figure 1c shows a map of county boundaries for the 100 counties in the state of North Carolina. The color gradient provides a visualization of the time variation in the rollout of funding from The Duke Endowment. Counties are assigned to one of five groups based on the initial year in which they received funding (dark to light red): 1927-28, 1929, 1930-33, 1934-38, 1939-42. Darker counties were exposed to funding for a longer period of time. Counties that did not receive funding from The Duke Endowment during the sample period from 1922 to 1942 are colored white. Treatment is measured by appropriations for capital expenditures from The Duke Endowment. Treatment timing is based on the *Annual Report of the Hospital Section* for the years 1927 to 1942, published by The Duke Endowment. Figure 1b plots the average annual infant mortality rate per 1,000 live births in North Carolina between 1922 to 1942, weighted by county population, for Black infants (gray dotted line), White infants (solid black line), as well as Black and White infants pooled together (dashed black line). Figure 1d plots the percent change in the infant mortality rate per 1,000 live births in each North Carolina county between 1922 and 1942. Darker colors imply declines in infant mortality while brighter colors imply increases in infant mortality.

Figure 2. Hospital beds at the county-level by treatment status, ownership, and event time.

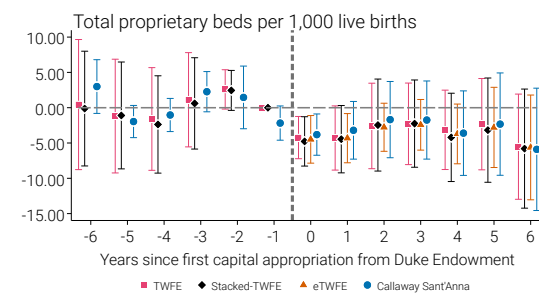
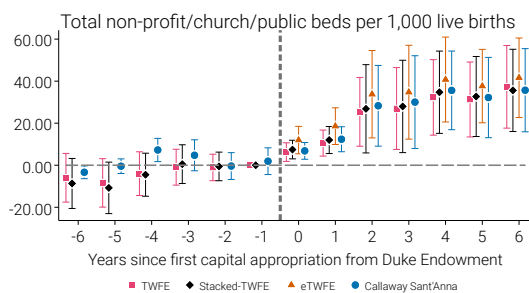
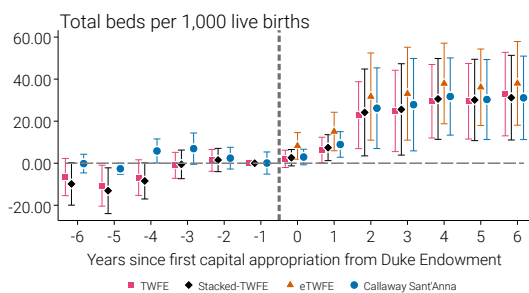
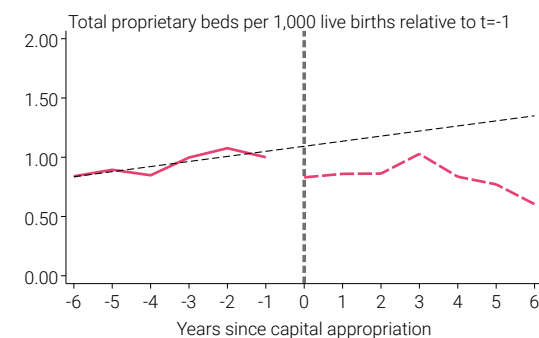
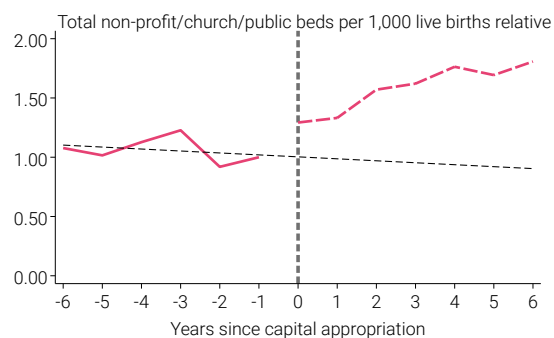
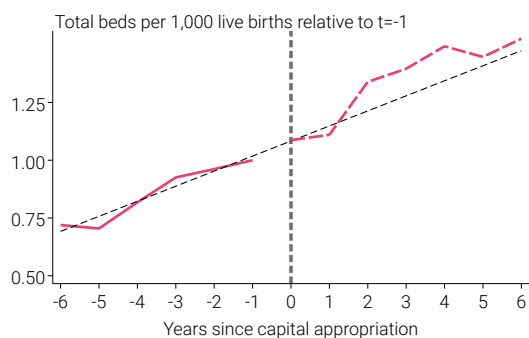
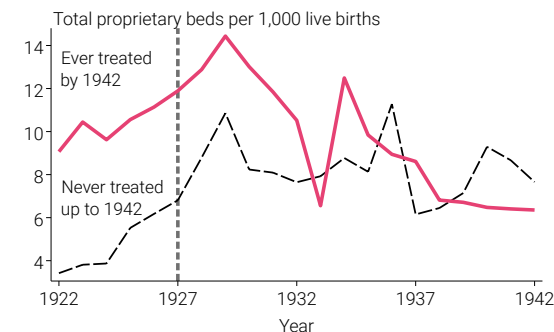
(a) All general beds



(b) Non-profit/public/church general beds



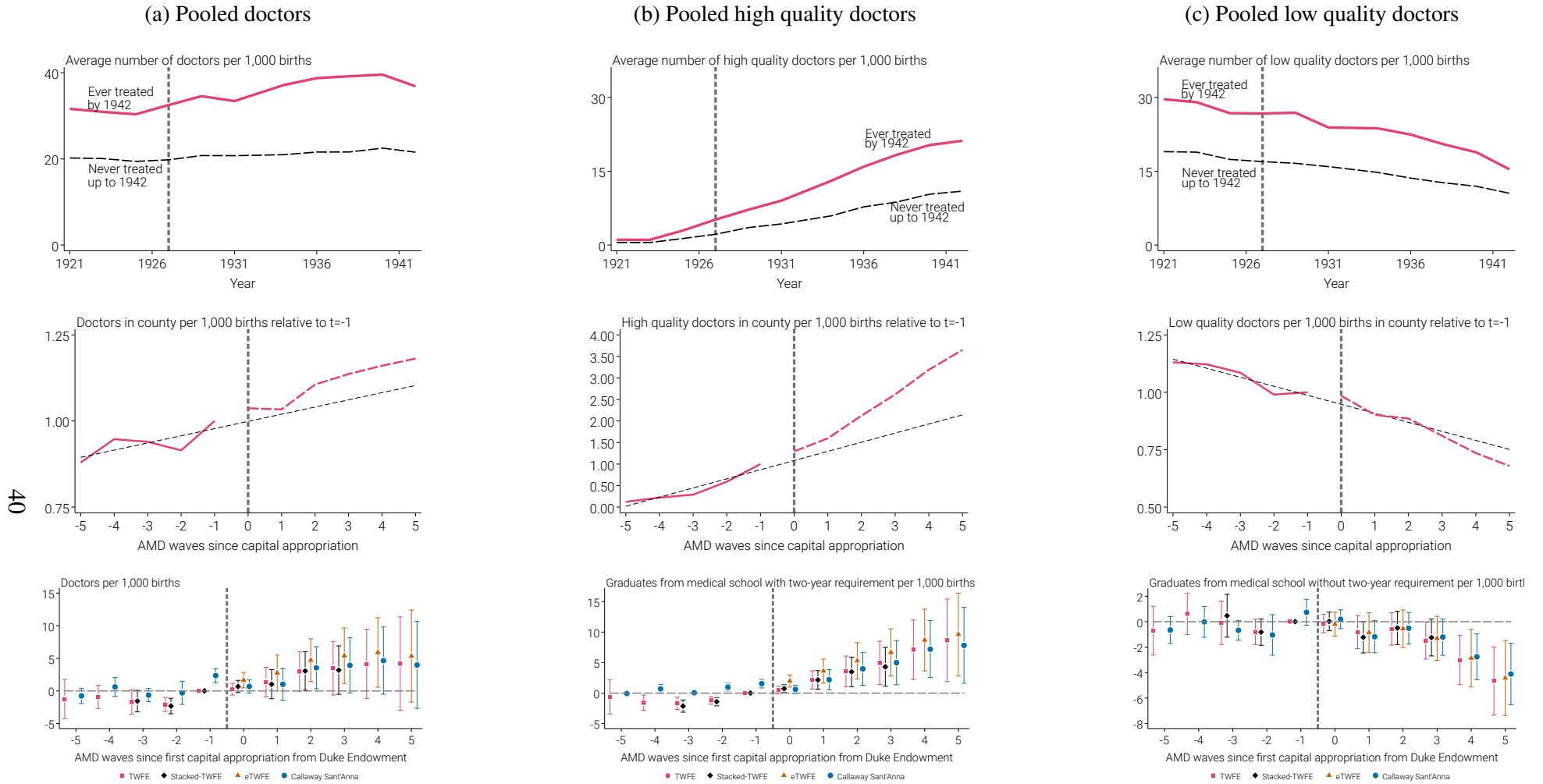
(c) Proprietary general beds



39

Notes: Column (a) plots results for the number of beds per 1,000 live births in all general hospital, column (b) plots the number of beds per 1,000 live births in non-profit, public, or church-run general hospitals, and column (c) plots the number of beds per 1,000 live births in proprietary general hospitals. The top row of plots shows the average annual number of general hospital beds in a county by Duke treatment status. Counties “Ever treated by 1942” first received Duke funding during the sample period, between 1927 and 1942, while counties “Never treated up to 1942” did not. The middle row plots the number of general hospital beds relative to the year prior to treatment in counties treated by 1942. The black dashed line is a best-fit line estimated using only pre-treatment data for treated counties. This trend is continued into post-treatment event time to serve as comparison to actual outcomes. The last row plots event study estimates of coefficient values and 95% confidence intervals for the lead and lag indicator variables for relative time periods from $t = -6$ to $t = 6$ around the first year that a county received an appropriation for capital expenditures from The Duke Endowment. More extreme relative time periods are estimated but not shown in the figures. The omitted category is -1 year before initial treatment. An observational unit is a county by year cell. Each plot shows four event study estimators: two-way fixed effects, stacked regression, extended two-way fixed effects (Wooldridge 2021), and Callaway and Sant’Anna (2021). All regressions include county and year fixed effects. Regressions are weighted by county-by-year of birth cohort size. Standard errors are clustered by county.

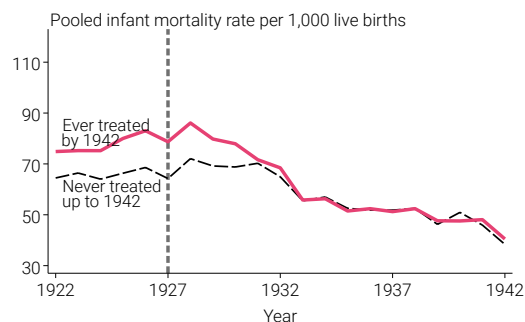
Figure 3. Physician supply by treatment status, quality, and event time.



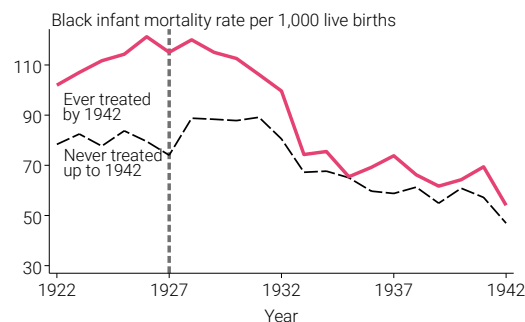
Notes: Column (a) plots results for the number of doctors, column (b) plots the number of high-quality doctors, and column (c) plots the number of low-quality doctors. All columns pool doctors by race. A high-quality doctor is one who graduated from a medical school at least 4 years after it introduced a two-year college degree prerequisite for admission. All other doctors are considered low quality. The top row of plots shows the average annual number of doctors in a county by Duke treatment status. Counties "Ever treated by 1942" first received Duke funding during the sample period, between 1927 and 1942, while counties "Never treated up to 1942" did not. The middle row plots the number of doctors relative to the year prior to treatment in counties treated by 1942. Event time represents the number of *American Medical Directory* (AMD) waves since the first year that a county received a capital appropriation from the Endowment. During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. The black dashed line is a best-fit line estimated using only pre-treatment data for treated counties. This trend is continued into post-treatment event time to serve as comparison to actual outcomes. The last row plots two-way fixed effects event study estimates of coefficient values and 95% confidence intervals for the lead and lag indicator variables for relative time periods from $t = -5$ to $t = 5$ around the first AMD wave after a county received an appropriation for capital expenditures from The Duke Endowment (or from $t = -3$ to $t = 3$ in the case of the stacked event study). More extreme relative time periods are estimated but not shown in the figures. The omitted category is -1 AMD wave before initial treatment. An observational unit is a county-by-AMD wave. Each plot shows four event study estimators: two-way fixed effects, stacked regression, the extended two-way fixed effects (Wooldridge 2021), and Callaway and Sant'Anna (2021). All regressions include county and AMD-wave fixed effects. Regressions are weighted by county of birth cohort size averaged over the two or three years of the AMD wave. Standard errors are clustered by county.

Figure 4. Event study estimates and infant mortality rates by treatment status, race, and event time.

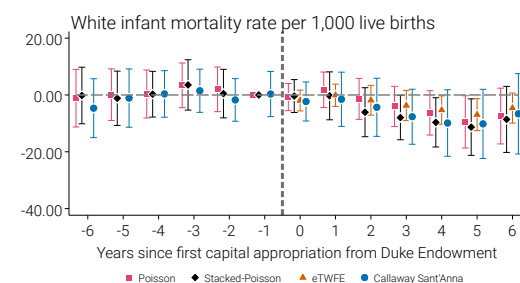
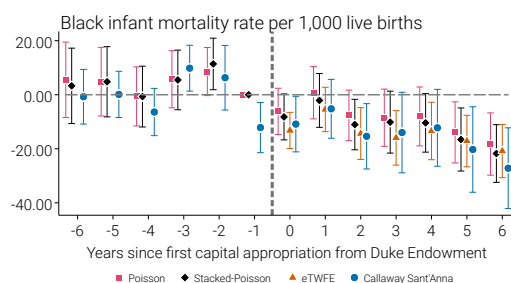
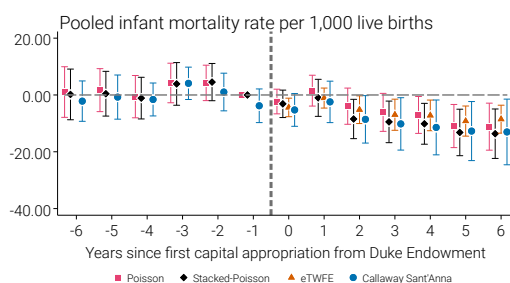
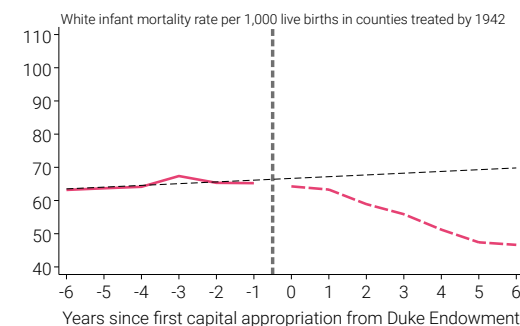
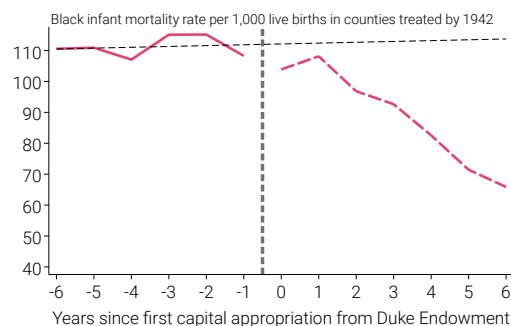
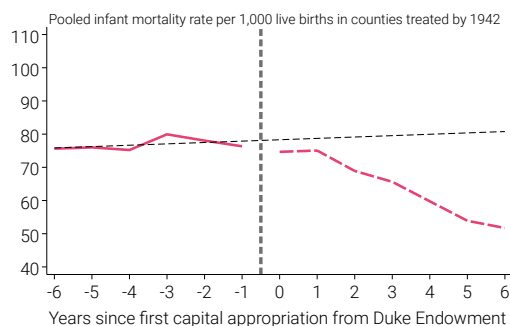
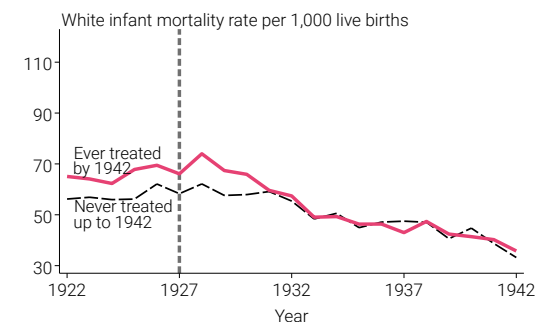
(a) Pooled infant mortality rate



(b) Black infant mortality rate



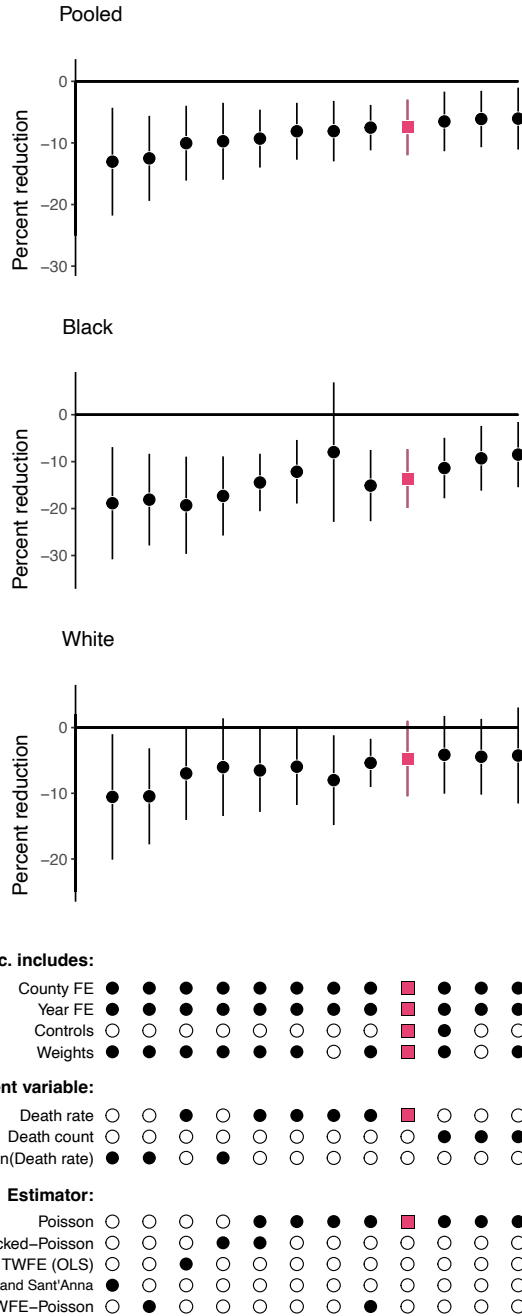
(c) White infant mortality rate



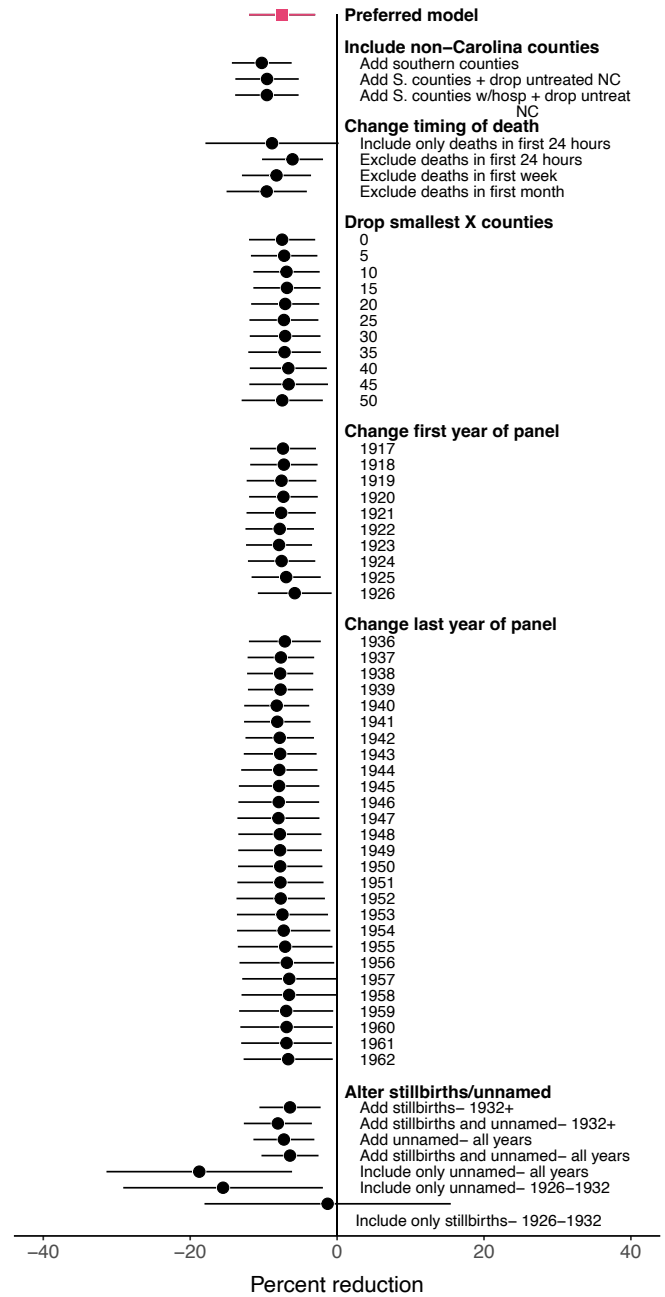
Notes: The top row of plots shows the annual infant mortality rate per 1,000 live births by Duke treatment status for Blacks and Whites pooled together, Blacks only, and Whites only, respectively. Counties “Ever treated by 1942” first received Duke funding during the sample period, between 1927 and 1942, while counties “Never treated up to 1942” did not. The middle row plots the infant mortality rate per 1,000 live births in counties treated by 1942 in event time relative to the first year of receiving a capital appropriation from the Endowment for the same three respective groups as in the top row. The black dashed line is a best-fit line estimated using only pre-treatment data for treated counties. This trend is continued into post-treatment event time to serve as comparison to actual outcomes. The last row plots Poisson pseudo maximum likelihood regression estimates of exponentiated coefficient values and 95% confidence intervals for the lead and lag indicator variables for time periods from $t = -6$ to $t = 6$ around the first year that a county received an appropriation for capital expenditures from The Duke Endowment. Our event study plots omit the left-most coefficients for relative time periods between -18 and -7 and right-most coefficients for relative time periods between 7 and 15 . The omitted category is -1 year before initial treatment. An observational unit is a county-by-year birth cohort. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. The dependent variables are the infant mortality rate per 1,000 live birth in birth county c and year t : pooled by race, for Blacks only, and for Whites only, respectively. All regressions include county and year fixed effects. Regressions are weighted by county-by-year of birth cohort size. Standard errors are estimated using the delta method and are clustered by county of birth.

Figure 5. A visual representation of key robustness checks

(a) Varying specification, dependent variable, and estimator



(b) Varying sample for pooled infant mortality rate and preferred model



Notes: Point estimates and 95-percent confidence interval for the coefficient on Duke support. See Table 3 and Online Appendix Tables D1 and D2 for a description of the specifications plotted in panel A. See Online Appendix Tables F3, G4, and G5 for specifications in panel B. Specifications that change the first or last year of the panel use births from Bailey et al. (2015) in constructing the dependent variable because North Carolina Vital Statistics data are only available from 1922 to 1948. Estimates in panel B are based on our preferred specification from column 6 of Table 3. Standard errors are estimated using the delta method and are clustered at the county level.

Online Appendix

The Gift of a Lifetime: The Hospital, Modern Medicine, and Mortality

Alex Hollingsworth, Krzysztof Karbownik, Melissa A. Thomasson, and Anthony Wray

A Sample images of data sources

Figure A1. Sample images

(a) The Duke Endowment's *Annual Report of the Hospital Section*

CONSTRUCTION, EQUIPMENT AND PURCHASE APPROPRIATIONS AND PAYMENTS							
	Location	Unpaid Appropriations Balance Dec. 31, 1939	Appropriated 1940	Payments 1940	Unpaid Appropriations Balance Dec. 31, 1940	Purpose	Approximate Cost of Projects Completed in 1940
16 PROJECTS.....		111,750	146,100	186,850	71,000		882,400
12 NORTH CAROLINA PROJECTS.....		82,250	98,600	146,350	34,500		803,200
Ashe County Memorial Hospital.....	Jefferson.....		3,000		3,000	Equipment	
Cabarrus County Hospital.....	Concord.....	20,000	20,000	40,000		Addition	177,200
Columbus County Hospital.....	Whiteville.....		2,500		2,500	Home for Nurses	
Community Hospital.....	Roxboro.....		12,000	12,000		Purchase	29,000
Grace Hospital.....	Morganton.....		5,000	5,000		Equipment	15,500
Mountain Sanitarium.....	Fletcher.....		8,000	5,000	3,000	Home for Nurses	
Presbyterian Hospital.....	Charlotte.....	60,000	16,500	76,500		New Plant	566,000
Randolph Hospital.....	Asheboro.....		4,000	2,000	2,000	Equipment	
Rowan Memorial Hospital.....	Salisbury.....		3,600	3,600		Addition	11,000
Rutherford Hospital.....	Rutherfordton.....	2,250		2,250		Addition	4,500
Shelby Hospital.....	Shelby.....		6,000		6,000	Addition	
Transylvania Community Hospital.....	Brevard.....		18,000		18,000	New Plant	

(b) *American Medical Directory*

DURHAM, N. C.		AMERICAN MEDICAL DIRECTORY		FOREST CITY	
SMITH, DAVID TILLERSON, b'98; Md.7,'22; '131, N.B.'23; Ⓢ A36, D11.15, F1.2; Prof. Bact. and Assoc. Prof. Med., N.C.7—Dover Rd., Hope Valley; office, Duke Hospital; 9-5 Solomon, Wm. Weldon (col.), b'04; D.C.3,'33; Intern—Lincoln Hospital		Owens, Zack Doxey, b'04; Md.1,'30; '130—311 W. Church St.; office, Medical Arts Bldg		FARMVILLE, 2,056, PITT Joyner, Claudius Cameron, b'72; Pa.2,'99; '199 Morrill, David S., b'76; Md.4,'97; '197 Stevens, Alexander Hamilton Jr., b'05; Ga.1,'32; '133	
SPEED, JOS. ANDERSON, b'88; Pa.2,'14; '114—321 E. Main St.; office, 212 W. Main St		PETERS, WM. ANTHONY, b'90; Va.4,'15; '115; Ⓢ S—104 S. Road St.; office, 106 S. Road St.; 9-10		WILLIS, WM. MOORE, b'90; Va.4,'14; '115; Ⓢ	
SPEKTER, LOUISE, b'08; N.Y.45,'33; Intern—Duke Hospital		SAWYER, WALTER WESLEY, b'80; Md.1,'03; '1'03; Ⓢ OALR*—5 N. Elliott St		FAYETTEVILLE, 13,049, CUMBERLAND ALLGOOD, REESE ALEXANDER, b'89; Md.1,'12; '115; Ⓢ Ob	
SPIKES, NORMAN OWEN, b'00; Pa.2,'24; '124—601 N. Hyde Park Ave.; office, 331 W. Main St		Styron, Nathan Henry, Jr. (col.) b'82; N.C.3,'08; '1'08		ASHCRAFT, JOHN ELLIS, b'60; N.Y.5,'87; '1'87—Pershing Heights	
Sprunt, Douglas Hamilton, b'00; Conn.1,'27; Assoc. Prof. Path., N.C.7—Dover Rd., Hope Valley; office, Duke Hospital		WALKER, HERBERT DILLON, b'77; Md.1,'02; '1'02; Ⓢ—306 E. Main St.		AVERITT, KIRBY GLADSTONE, b'70; Md.4,'93; '1'93; Ⓢ Ob—R.D.5	
STANFORD, LOIS BROOKE FOOTE, b'96; Pa.1,'21; '125; Ⓢ CP—1411 Alabama Ave.; office, 123 W. Main St.; 9-12, 2-4		WARD, IVIE ALPHONSO, b'79; N.C.1,'07; '1'07; OALR—225 N. Road St.; office, 308 E. Main St.; 9-12, 2-5		Devane, Wm. P. (col.), b'82; Tenn.14,'15 Tenn.7,'18; '1'20—434 Gillespie St.; office, 135 Person St	
STANFORD, WM. RANEY, b'92; Pa.1,'19; '1'19; Ⓢ I* D9—1411 Alabama Ave.; office, 123 W. Main St.; 9-30-12:30, 2-4:30		WHITE, WM. HENRY CLAY, b'98; Va.1,'22; '1'29; Ⓢ S also Member Ill. State Med. Soc.—118 E. Burgess St.; office, 5 N. Elliott St.; 10-12, 2-4		Foster, Malcolm Tennyson, b'02; Ga.5,'27; '1'27—County Health Department	
STROWD, WM. AMICK, b'82; N.C.1,'09; '1'09—901 N. Mangum St.; office, 415½ E. Chapel		WILLIAMS, CLAUDE BURGESS, b'77; Va.6,'03; '1'03—101 E. Burgess St.; office, 5 Elliott St.; 9-11, 2-5		GAINES, JOHN WHITE, b'94; Pa.2,'17; '1'17; U—815 Arsenal Ave.; office, Hoy St.; 9-1, 2-4	
		ELIZABETHTOWN, 765, BLADEN		HIGHSMITH, JACOB FRANK, JR., b'01; Pa.1,'27; '1'27; Ⓢ—Highsmith Hospital	
				HIGHSMITH, JACOB F., b'68; Pa.2,'89; '1'89; Ⓢ S A28	

Notes: Example images of The Duke Endowment's *Annual Report of the Hospital Section*, the source for our measure of exposure to Duke support (top) and the *American Medical Directory*, the source for the number of physicians by county and year (bottom). Photo credits: Authors.

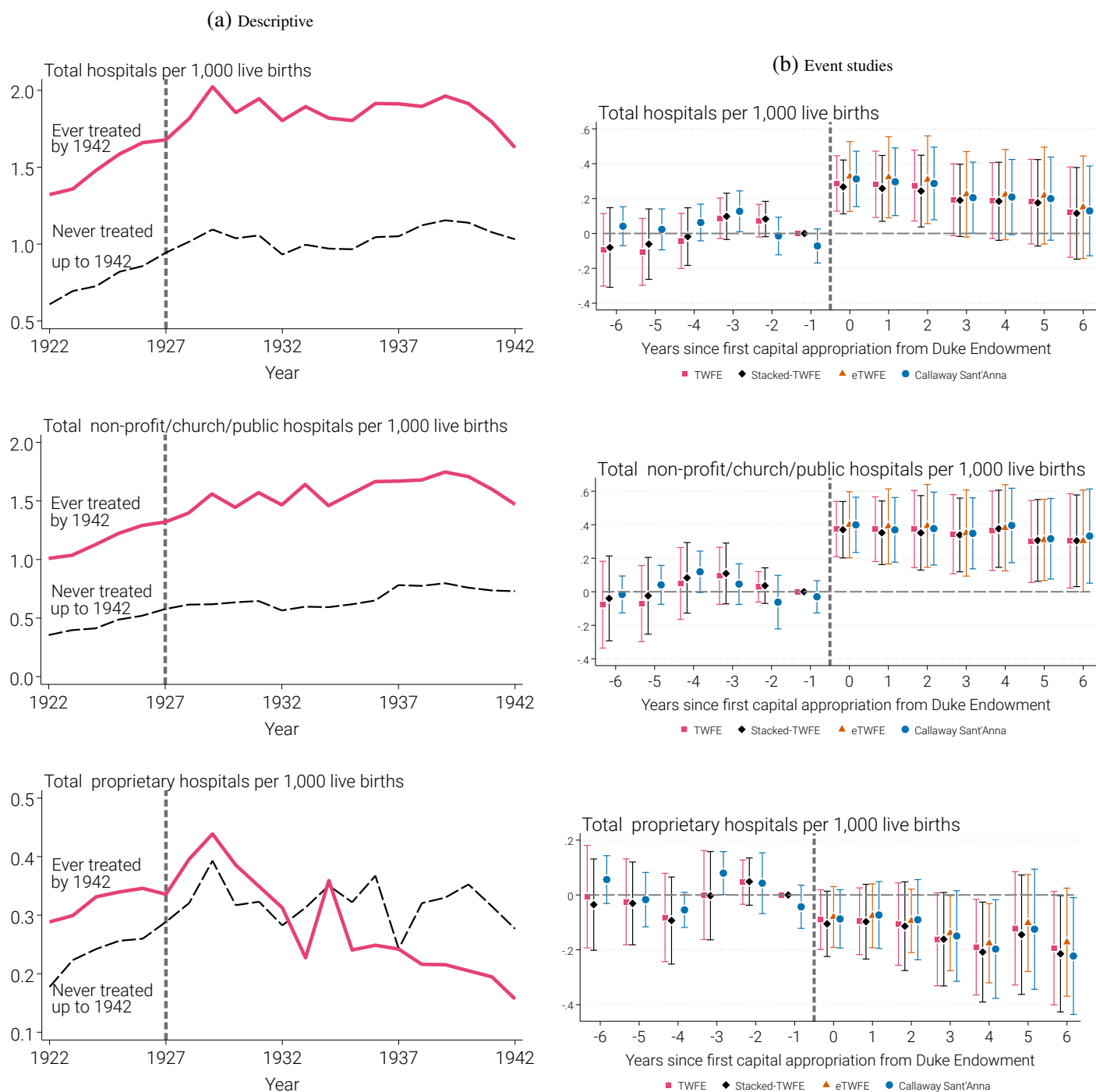
B First stage for hospitals and hospital beds: Additional results

Table B1. First stage for hospitals and hospital beds: Robustness to other estimators and treatment definition

	$Y_{ct}^R = \text{Beds or Hospitals}$				$Y_{ct}^R = \text{Beds or Hospitals per 1000 births}$			
	(1) TWFE	(2) Stacked-TWFE	(3) eTWFE	(4) CS	(5) TWFE	(6) Stacked-TWFE	(7) eTWFE	(8) CS
<i>A. Beds (Duke treatment: All appropriations)</i>								
Total	44.72*** (13.36)	40.96*** (13.80)	72.86*** (15.73)	49.23*** (16.98)	27.48*** (6.62)	26.82*** (7.94)	38.13*** (9.30)	31.19*** (9.01)
Non-profit/church/public	47.71*** (14.10)	43.54*** (14.16)	78.39*** (14.98)	55.09*** (16.61)	30.46*** (6.90)	29.41*** (8.08)	42.04*** (9.05)	34.61*** (8.87)
Proprietary	-5.19 (3.24)	-4.48 (2.95)	-8.02** (3.12)	-7.35*** (2.47)	-4.35* (2.51)	-3.79 (2.42)	-5.45** (2.12)	-4.04 (2.81)
<i>B. Hospitals (Duke treatment: Exclude homes for nurses)</i>								
Total	0.20 (0.14)	0.16 (0.15)	0.18 (0.14)	0.14 (0.13)	0.20* (0.12)	0.20* (0.12)	0.16 (0.15)	0.16 (0.12)
Non-profit/church/public	0.33** (0.16)	0.31* (0.16)	0.34** (0.14)	0.35** (0.14)	0.32*** (0.11)	0.31*** (0.12)	0.29** (0.14)	0.29** (0.13)
Proprietary	-0.17* (0.09)	-0.18** (0.09)	-0.20** (0.08)	-0.24** (0.11)	-0.13* (0.08)	-0.13* (0.08)	-0.14* (0.08)	-0.14 (0.09)
Observations	2,100	7,371	2,100	2,100	2,100	7,371	2,100	2,100
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No	No	No

Notes: Each coefficient comes from a separate regression estimated by two-way fixed effects (columns 1 and 5), stacked regression (columns 2 and 6), extended two-way fixed effects by Wooldridge (2021) (columns 3 and 7), or Callaway and Sant'Anna (2021) (columns 4 and 8). In columns 1 to 4, the dependent variable is the number of hospital beds (Panel A) or the number of hospitals (Panel B) in a county and year. Within Panel A, the dependent variable is the number of beds in hospitals of any type (top row); in non-profit, church-owned, or public hospitals (middle row); or in proprietary hospitals (bottom row). Panel B presents the same classification across rows with the number of hospitals of each type as the dependent variable. Columns 5 to 8 express the dependent variables as rates per 1,000 live births and follows the same structure as columns 1 to 4. Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. In panel A the treatment includes all capital appropriations while panel B excludes homes for nurses. The weights are the number of births in a county and year. Standard errors are clustered at the county level.

Figure B1. Hospitals at the county-level by treatment status, ownership, and event time.



Notes: Column (a) plots shows the average annual number of hospitals in a county by Duke treatment status. Column (b) presents event-study estimates and 95% confidence intervals for the lead and lag indicator variables for relative time periods from $t = -6$ to $t = 6$ around the first year that a county received an appropriation for capital expenditures from The Duke Endowment. More extreme relative time periods are estimated but not shown in the figures. The omitted category is -1 year before initial treatment. An observational unit is a county by year cell. Each plot shows four event study estimators: two-way fixed effects, stacked regression, extended two-way fixed effects (Wooldridge 2021), and Callaway and Sant'Anna (2021). All regressions include county and year fixed effects. Regressions are weighted by county-by-year of birth cohort size. Standard errors are clustered by county. The top row of plots shows the average annual number of general hospitals in a county by Duke treatment status. Appropriations for nurse homes are excluded from the definition of treatment. Counties "Ever treated by 1942" first received an appropriation for Duke funding during the sample period, between 1927 and 1942, while counties "Never treated up to 1942" did not. The middle row shows not-for-profits. The final row shows proprietary hospitals.

C First stage results for doctors and other health care professionals

Table C1. Effects on number of doctors: Robustness to alternate estimators

	$Y_{ct}^R = \text{Doctors}$				$Y_{ct}^R = \text{Doctors per 1000 births}$			
	(1) TWFE	(2) Stacked-TWFE	(3) eTWFE	(4) CS	(5) TWFE	(6) Stacked-TWFE	(7) eTWFE	(8) CS
<i>A. Pooled</i>								
All	8.06** (3.26)	5.75** (2.39)	11.47*** (4.30)	6.96** (3.55)	3.40** (1.44)	3.30*** (1.22)	4.17** (2.02)	2.84 (1.86)
High quality	12.22*** (3.45)	8.92*** (2.27)	17.19*** (4.43)	13.20*** (3.80)	4.80*** (1.44)	3.87*** (1.09)	6.43*** (1.92)	4.73*** (1.69)
Low quality	-4.16*** (1.34)	-3.16*** (1.10)	-5.77*** (1.79)	-6.38*** (1.96)	-1.40* (0.73)	-0.62 (0.53)	-2.29** (0.90)	-2.00*** (0.65)
Observations	1,100	2,128	1,100	1,100	1,100	2,128	1,100	1,100
<i>B. Black</i>								
All	0.79* (0.44)	0.81* (0.45)	0.99** (0.47)	0.72* (0.40)	1.17* (0.65)	1.43** (0.72)	1.24* (0.71)	1.34* (0.77)
High quality	1.60*** (0.44)	1.45*** (0.41)	1.90*** (0.52)	1.57*** (0.53)	2.53*** (0.67)	2.33*** (0.66)	2.87*** (0.79)	2.60*** (0.84)
Low quality	-0.81*** (0.27)	-0.64** (0.25)	-0.91*** (0.30)	-0.85*** (0.25)	-1.37*** (0.51)	-0.90* (0.49)	-1.63*** (0.56)	-1.26** (0.51)
Observations	1,078	2,058	1,078	1,078	1,078	2,058	1,078	1,078
<i>C. White</i>								
All	6.77** (2.80)	4.80** (1.97)	10.12*** (3.84)	5.70* (3.18)	3.60* (2.09)	3.51** (1.72)	4.62 (3.07)	2.51 (2.77)
High quality	9.87*** (2.96)	7.20*** (1.90)	14.73*** (3.96)	10.98*** (3.27)	4.25** (1.99)	3.26** (1.55)	6.33** (2.73)	4.00* (2.40)
Low quality	-3.11** (1.36)	-2.41** (1.14)	-4.68** (1.92)	-5.43*** (2.09)	-0.67 (1.04)	0.17 (0.82)	-1.77 (1.36)	-1.68* (1.01)
Observations	1,078	2,058	1,078	1,078	1,078	2,058	1,078	1,078
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AMD Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No	No	No

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression estimated by two-way fixed effects (columns 1 and 5), stacked regression (columns 2 and 6), extended two-way fixed effects by Wooldridge (2021) (columns 3 and 7), or Callaway and Sant'Anna (2021) (columns 4 and 8). In columns 1 to 4, the dependent variable is the number of doctors (Panel A), the number of Black doctors (Panel B), or the number of White doctors (Panel C) in a county and publication year of the *American Medical Directory* (AMD wave). During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. Within each panel, the dependent variable is the number of doctors (top row); the number of high-quality doctors (middle row); or the number of low-quality doctors (bottom row). A high-quality doctor is one who graduated from a medical school at least 4 years after it introduced a two-year college degree prerequisite for admission. All other doctors are considered low quality. Columns 5 to 8 express the dependent variables as rates per 1,000 live births and follows the same structure as columns 1 to 4. Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Weights are the average number of births in each county for the years in the AMD wave. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942 for any AMD wave. Standard errors are clustered at the county level.

Table C2. Effects on doctors: Robustness to other measures of high quality

	$Y_{ct}^R = \text{Doctors}$	$Y_{ct}^R = \text{Doctors per 1000 births}$
	(1)	(2)
<i>A. Pooled - High Quality</i>		
Graduates from medical school with two-year requirement	9.02*** (2.58)	3.88*** (1.23)
Graduates from medical school ever with A/A+ AMA rating	6.29*** (1.96)	3.02*** (0.96)
Graduates from medical school that exists and is approved in 1942	7.48*** (2.55)	3.27*** (1.23)
Graduates from medical school that remains open	7.77*** (2.66)	3.43*** (1.28)
Observations	1,100	1,100
<i>B. Black - High Quality</i>		
Graduates from medical school with two-year requirement	1.34*** (0.37)	2.31*** (0.59)
Graduates from medical school ever with A/A+ AMA rating	1.33*** (0.36)	2.27*** (0.64)
Graduates from medical school that exists and is approved in 1942	1.33*** (0.36)	2.27*** (0.64)
Graduates from medical school that remains open	1.27*** (0.35)	2.23*** (0.65)
Observations	1,078	1,078
<i>C. White - High Quality</i>		
Graduates from medical school with two-year requirement	6.71*** (2.00)	2.79* (1.61)
Graduates from medical school ever with A/A+ AMA rating	4.70*** (1.50)	2.42* (1.34)
Graduates from medical school that exists and is approved in 1942	5.52*** (1.93)	2.21 (1.63)
Graduates from medical school that remains open	5.80*** (2.03)	2.41 (1.71)
Observations	1,078	1,078

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression estimated by two-way fixed effects. In column 1, the dependent variable is the number of doctors (Panel A), the number of Black doctors (Panel B), or the number of White doctors (Panel C) in a county and publication year of the *American Medical Directory* (AMD wave). During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. Column 2 expresses the dependent variables as rates per 1,000 live births and follows the same structure as column 1. Within each panel, the dependent variable is a different measure of high-quality doctors: the number of doctors who graduated from a medical school at least 4 years after it introduced a two-year college degree prerequisite for admission, our main measure of high quality (top row); the number of doctors who graduated from a medical school with an A or A+ rating from the American Medical Association, (second row); the number of doctors who graduated from a medical school that existed and was approved in 1942 (third row); or the number of doctors who graduated from a medical school that did not close and was not absorbed by another school during the sample period (bottom row). Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Weights are the average number of births in each county for the years in the AMD wave. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942 for any AMD wave. Standard errors are clustered at the county level.

Table C3. Effects on doctors: Robustness to other measures of low quality

	$Y_{ct}^R = \text{Doctors}$	$Y_{ct}^R = \text{Doctors per 1000 births}$
	(1)	(2)
<i>A. Pooled - Low Quality</i>		
Graduates from medical school without two-year requirement	-3.35*** (1.11)	-1.02 (0.69)
Graduates from medical school without A/A+ AMA rating	-0.65 (0.85)	-0.19 (0.50)
Did not graduate from medical school that exists and is approved in 1942	-1.84*** (0.64)	-0.44 (0.47)
Graduates from medical school that closed	-2.13*** (0.63)	-0.60 (0.44)
Observations	1,100	1,100
<i>B. Black - Low Quality</i>		
Graduates from medical school without two-year requirement	-0.77*** (0.22)	-1.51*** (0.55)
Graduates from medical school without A/A+ AMA rating	-0.75*** (0.20)	-1.47*** (0.46)
Did not graduate from medical school that exists and is approved in 1942	-0.75*** (0.20)	-1.47*** (0.46)
Graduates from medical school that closed	-0.70*** (0.20)	-1.43*** (0.46)
Observations	1,078	1,078
<i>C. White - Low Quality</i>		
Graduates from medical school without two-year requirement	-2.31** (1.07)	-0.11 (0.94)
Graduates from medical school without A/A+ AMA rating	-0.33 (0.75)	0.23 (0.65)
Did not graduate from medical school that exists and is approved in 1942	-1.14* (0.58)	0.43 (0.58)
Graduates from medical school that closed	-1.42** (0.55)	0.23 (0.54)
Observations	1,078	1,078

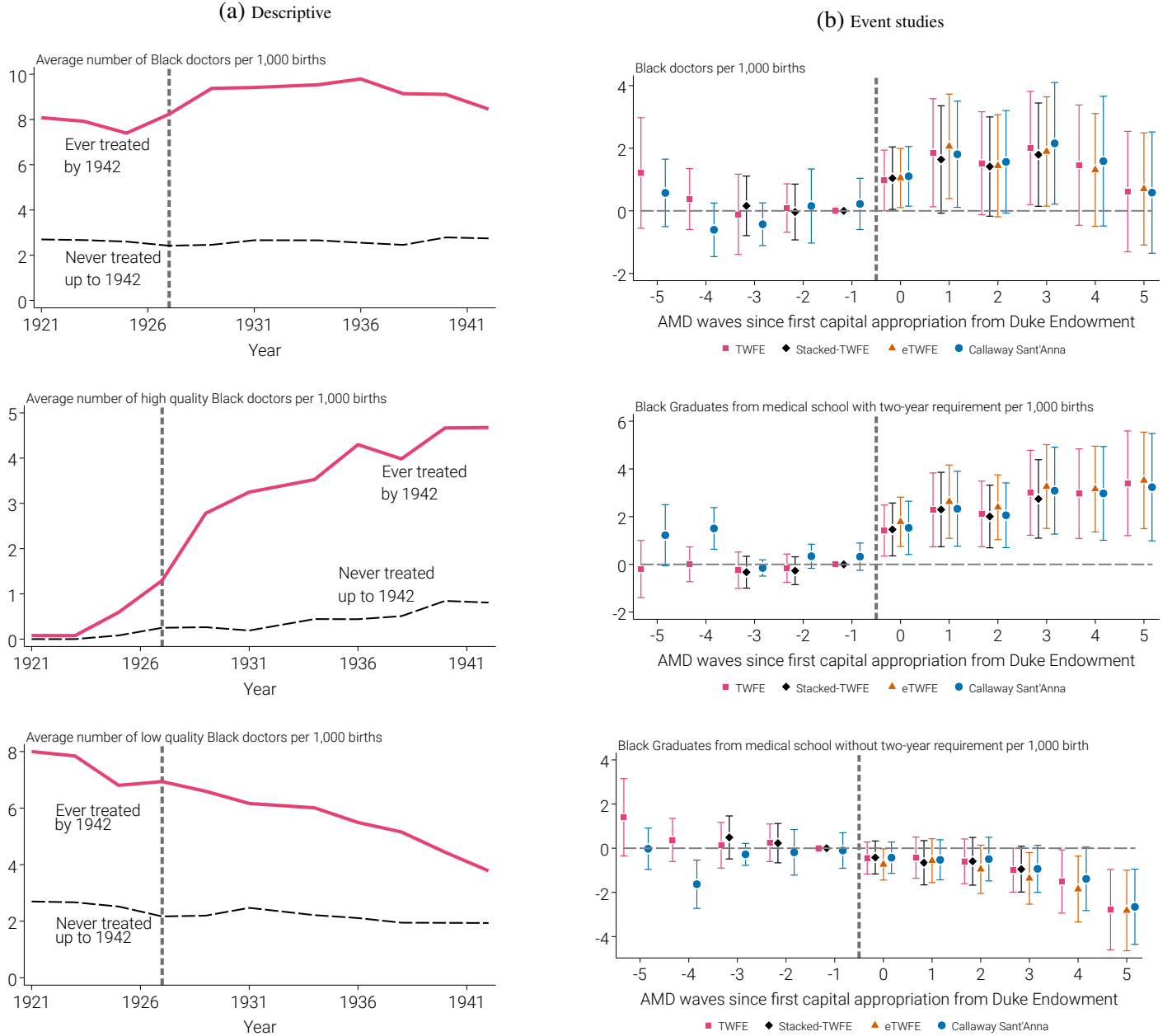
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression estimated by two-way fixed effects. In column 1, the dependent variable is the number of doctors (Panel A), the number of Black doctors (Panel B), or the number of White doctors (Panel C) in a county and publication year of the *American Medical Directory* (AMD wave). During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. Column 2 expresses the dependent variables as rates per 1,000 live births and follows the same structure as column 1. Within each panel, the dependent variable is a different measure of low-quality doctors: the number of doctors who graduated from a medical school that did not introduce a two-year college degree prerequisite for admission at least 4 years prior to graduation, our main measure of low quality (top row); the number of doctors who graduated from a medical school without an A or A+ rating from the American Medical Association, (second row); the number of doctors who graduated from a medical school that did not exist or was not approved in 1942 (third row); or the number of doctors who graduated from a medical school that closed or was by another school during the sample period (bottom row). Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Weights are the average number of births in each county for the years in the AMD wave. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942 for any AMD wave. Standard errors are clustered at the county level.

Table C4. Effects on doctors: Additional heterogeneity analyses

	$Y_{ct}^R = \text{Doctors}$			$Y_{ct}^R = \text{Doctors per 1,000 births}$		
	(1)	(2)	(3)	(4)	(5)	(6)
Surgeons	0.82*** (0.19)	1.11*** (0.33)	0.96*** (0.23)	0.86*** (0.19)	0.72*** (0.17)	0.75*** (0.17)
Specialists	3.34*** (1.04)	5.63*** (1.86)	3.70*** (1.13)	1.99*** (0.55)	2.49*** (0.68)	1.97*** (0.54)
AMA Fellows	1.96** (0.77)	3.09** (1.33)	1.62* (0.83)	1.00* (0.56)	1.12* (0.59)	0.72 (0.48)
AMA Members	2.42** (1.15)	4.78** (2.04)	2.93** (1.33)	1.28 (0.81)	1.91** (0.89)	1.38* (0.74)
Doctors from N.C. medical school	-0.06 (0.53)	0.01 (0.72)	0.01 (0.64)	-0.35 (0.36)	-0.15 (0.41)	-0.15 (0.37)
Doctors under 40	2.11 (1.57)	3.18 (2.31)	2.51 (1.76)	0.92 (1.07)	1.51 (1.26)	1.29 (0.95)
Doctors licensed after Flexner report	7.00*** (1.89)	11.24*** (3.22)	7.72*** (2.03)	2.54** (0.98)	4.26*** (1.11)	3.36*** (0.94)
Doctors licensed before Flexner report	-2.89*** (0.78)	-3.92*** (1.15)	-2.71*** (0.82)	-1.21** (0.50)	-1.17** (0.50)	-0.75 (0.50)
Observations	1,100	1,100	1,100	1,100	1,100	1,100
County FE	Yes	Yes	Yes	Yes	Yes	Yes
AMD Wave FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

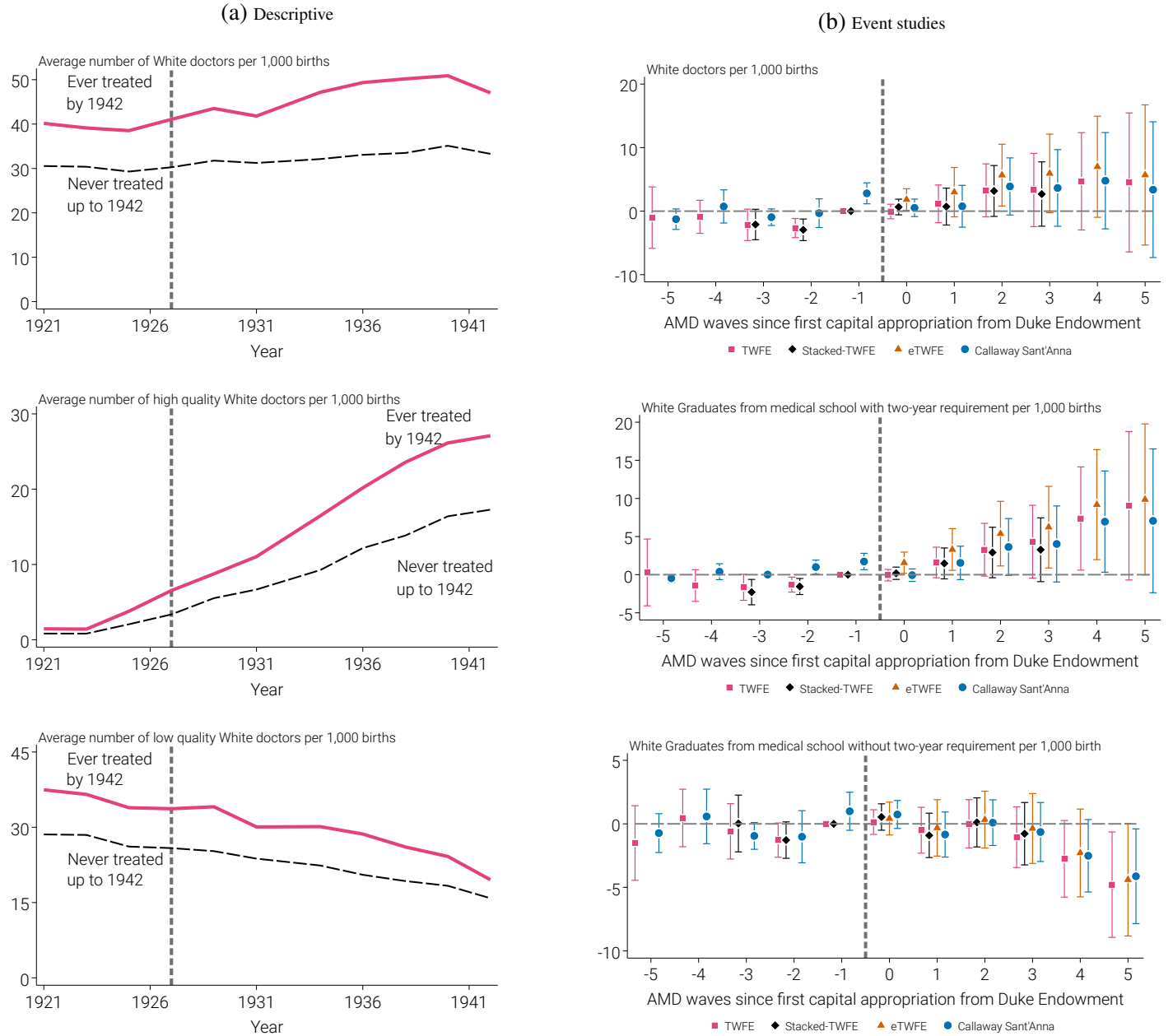
Notes: Each coefficient comes from a separate regression estimated by two-way fixed effects. The unit of observation is a county and publication year of the *American Medical Directory* (AMD wave). During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. Across the rows, the dependent variable is a measure of doctors belonging to a particular subset: surgeons, specialists, Fellows of the American Medical Association (AMA), AMA members, doctors who graduated from North Carolina medical schools, doctors under age 40, and doctors licensed after/before the publication of the Flexner Report in 1910. In columns 1 to 3, the dependent variable is the number of doctors in the subset, while columns 4 to 6 expresses it as a rate per 1,000 live births. Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Weights are the average number of births in each county for the years in the AMD wave. Standard errors are clustered at the county level.

Figure C1. Black physician results by treatment status, quality, and event time.



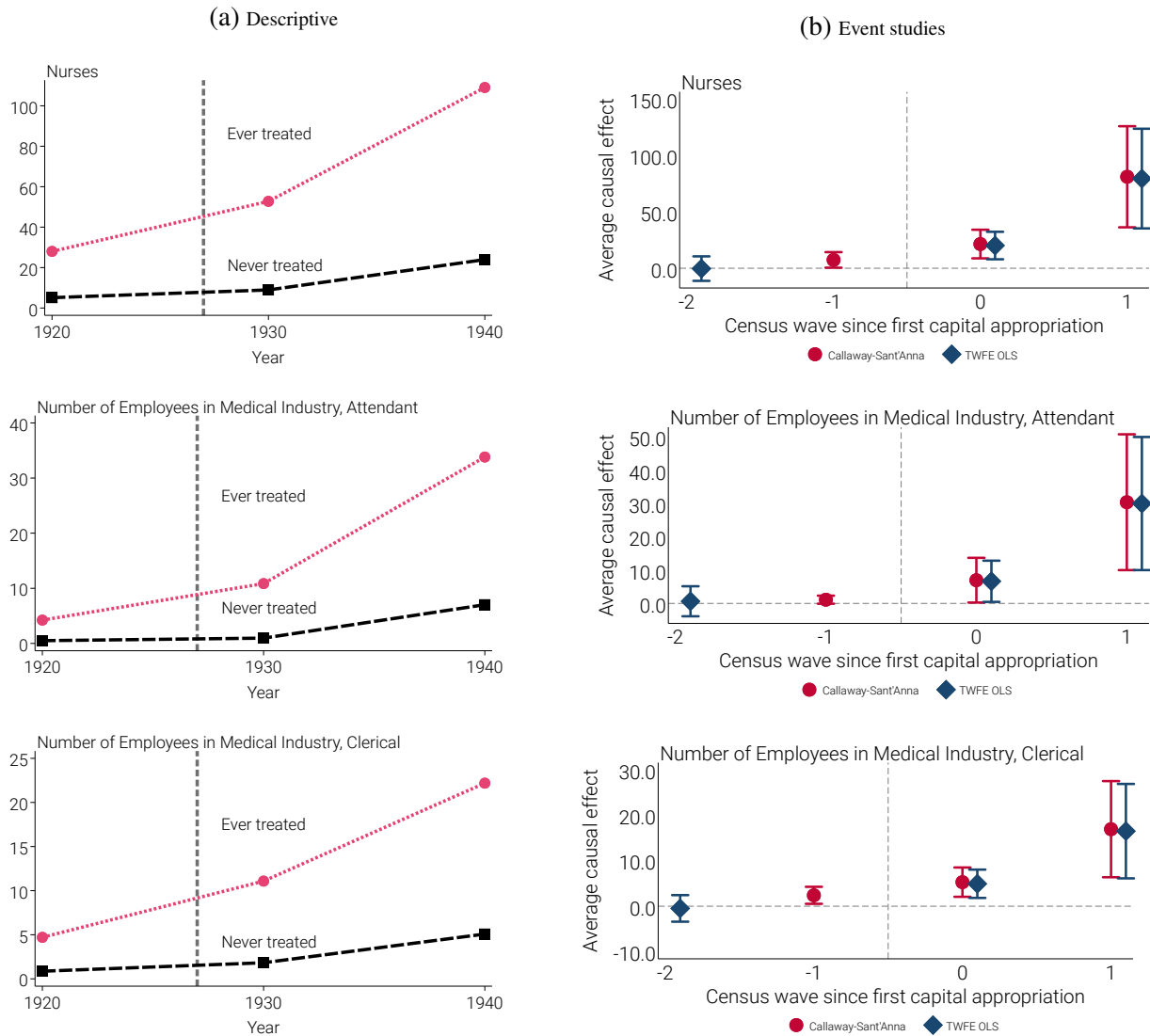
Notes: Column (a) plots shows the average annual number of doctors in a county by Duke treatment status. Column (b) plots event study estimates of coefficient values and 95% confidence intervals for the lead and lag indicator variables for relative time periods from $t = -5$ to $t = 5$ around the first AMD wave after a county received an appropriation for capital expenditures from The Duke Endowment (or from $t = -3$ to $t = 3$ in the case of the stacked event study). The first row presents the number of Black doctors per 1,000 births, the second row presents the number of high-quality Black doctors per 1,000 births, and the last row presents the number of low-quality Black doctors per 1,000 births. A high-quality doctor is one who graduated from a medical school at least 4 years after it introduced a two-year college degree prerequisite for admission. All other doctors are considered low quality. Counties “Ever treated by 1942” first received Duke funding during the sample period, between 1927 and 1942, while counties “Never treated up to 1942” did not. Event time represents the number of *American Medical Directory* (AMD) waves since the first year that a county received a capital appropriation from the Endowment. During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. More extreme relative time periods are estimated but not shown in the figures. The omitted category is -1, the AMD wave before initial treatment. Each plot shows four event study estimators: two-way fixed effects, stacked regression, extended two-way fixed effects (Wooldridge 2021), and Callaway and Sant’Anna (2021). An observational unit is a county-by-AMD wave. All regressions include county and AMD-wave fixed effects. Regressions are weighted by county of birth cohort size averaged over the two or three years of the AMD wave. Standard errors are clustered by county.

Figure C2. White physician results by treatment status, quality, and event time.



Notes: Column (a) plots shows the average annual number of doctors in a county by Duke treatment status. Column (b) plots event study estimates of coefficient values and 95% confidence intervals for the lead and lag indicator variables for relative time periods from $t = -5$ to $t = 5$ around the first AMD wave after a county received an appropriation for capital expenditures from The Duke Endowment (or from $t = -3$ to $t = 3$ in the case of the stacked event study). The first row presents the number of White doctors per 1,000 births, the second row presents the number of high-quality White doctors per 1,000 births, and the last row presents the number of low-quality White doctors per 1,000 births. A high-quality doctor is one who graduated from a medical school at least 4 years after it introduced a two-year college degree prerequisite for admission. All other doctors are considered low quality. Counties “Ever treated by 1942” first received Duke funding during the sample period, between 1927 and 1942, while counties “Never treated up to 1942” did not. Event time represents the number of *American Medical Directory* (AMD) waves since the first year that a county received a capital appropriation from the Endowment. During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938, 1940, and 1942. More extreme relative time periods are estimated but not shown in the figures. The omitted category is -1, the AMD wave before initial treatment. Each plot shows four event study estimators: two-way fixed effects, stacked regression, extended two-way fixed effects (Wooldridge 2021), and Callaway and Sant’Anna (2021). An observational unit is a county-by-AMD wave. All regressions include county and AMD-wave fixed effects. Regressions are weighted by county of birth cohort size averaged over the two or three years of the AMD wave. Standard errors are clustered by county.

Figure C3. Employment of other health care professionals



Notes: The figures in column (a) plot the average number of nurses (top row), as well as attendants (middle row) and clerical workers (bottom row) employed in the medical industry in 1920, 1930, and 1940, separately for North Carolina counties that were treated by Duke support by 1940 and those that were not. Column (b) displays event studies for the same set of outcomes, in which each unit of event time is a decade (the time between census waves). Event studies are estimated by two-way fixed effects or Callaway and Sant'Anna (2021). Data on the number of nurses and medical professionals come from aggregating individual records in the complete count censuses (Ruggles et al. 2023).

D Infant mortality: Additional results

Table D1. Robustness of infant mortality results to alternate estimators

	$Y_{ct}^R = \text{Infant deaths}$				$Y_{ct}^R = \text{Infant mortality rate}$			
	(1) Poisson	(2) Stacked-Poisson	(3) eTWFE-Poisson	(4) Callaway Sant'Anna	(5) Poisson	(6) Stacked-Poisson	(7) eTWFE-Poisson	(8) Callaway Sant'Anna
<i>A. Pooled</i>								
Percent effect from Duke (=1)	-6.05** (2.56)	-8.56*** (2.55)	-11.60*** (4.23)	-21.85*** (5.29)	-8.11*** (2.36)	-9.30*** (2.40)	-7.51*** (1.88)	-10.04*** (3.10)
Observations	2,100	6,643	2,100	2,100	2,100	6,643	2,100	2,100
<i>B. Black</i>								
Percent effect from Duke (=1)	-8.50** (3.54)	-10.36** (3.96)	-6.05** (2.73)	-25.23*** (7.25)	-12.15*** (3.46)	-14.44*** (3.12)	-15.11*** (3.87)	-19.29*** (5.29)
Observations	1,995	6,266	1,995	1,995	1,995	6,266	1,995	1,995
<i>C. White</i>								
Percent effect from Duke (=1)	-4.25 (3.73)	-7.40** (3.33)	-7.37*** (2.61)	-18.36*** (5.04)	-5.95** (2.98)	-6.52** (3.22)	-5.38*** (1.87)	-6.98* (3.61)
Observations	1,995	6,266	1,995	1,995	1,995	6,266	1,995	1,995
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No	No	No

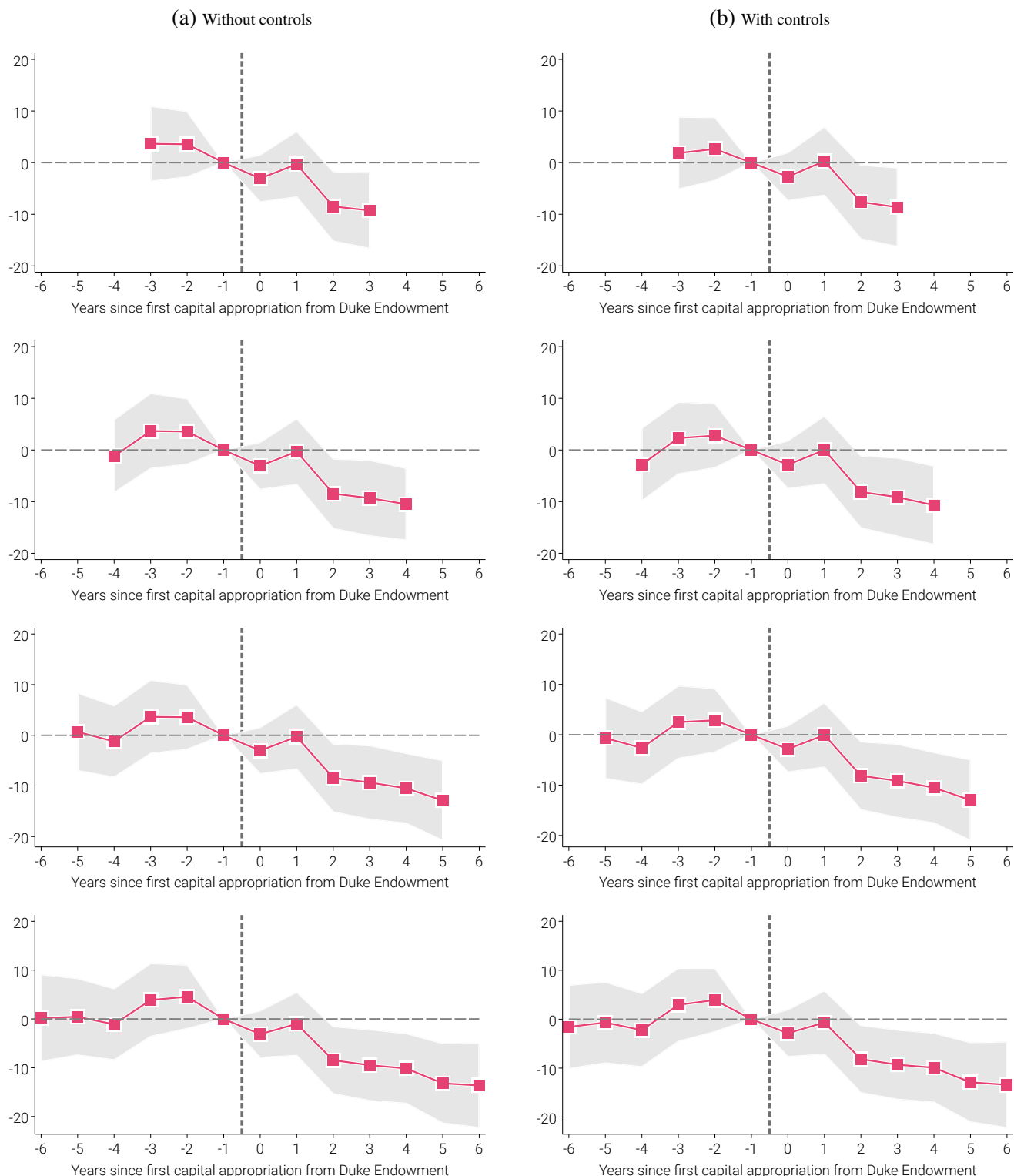
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression estimated by Poisson (columns 1 and 5), stacked Poisson with a balanced panel (columns 2 and 6), the extended two-way fixed effects estimator by Wooldridge (2021) (columns 3 and 7), or the Callaway and Sant'Anna (2021) estimator (columns 4 and 8). In columns 1 to 4, the dependent variable is the number of infant deaths (Panel A), the number of Black infant deaths (Panel B), or the number of White infant deaths (Panel C) in a county-by-year of birth cohort. Columns 5 to 8 express the dependent variables as rates per 1,000 live births and follow the same structure as columns 1 to 4. Each coefficient represents the percent change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. Coefficients are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Standard errors are calculated using the delta method. The weights are the number of births in a county and year. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942. Standard errors are clustered at the county level.

Table D2. Robustness of infant mortality results to log specification

	$Y_{ct}^R = \ln(\text{Infant mortality rate})$			
	(1)	(2)	(3)	(4)
	TWFE	Stacked-TWFE	eTWFE	CS
<i>A. Pooled</i>				
Percent effect from Duke (=1)	-0.09*** (0.03)	-0.10*** (0.03)	-0.13*** (0.04)	-0.13*** (0.04)
Observations	2,100	6,643	2,100	2,100
<i>B. Black</i>				
Percent effect from Duke (=1)	-0.14*** (0.05)	-0.17*** (0.04)	-0.18*** (0.05)	-0.19*** (0.06)
Observations	1,935	6,127	1,926	1,908
<i>C. White</i>				
Percent effect from Duke (=1)	-0.06* (0.03)	-0.06 (0.04)	-0.10*** (0.04)	-0.11** (0.05)
Observations	2,096	6,622	2,096	2,095
County FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes
Controls	No	No	No	No

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression estimated by two-way fixed effects (column 1), stacked two-way fixed effects (column 2), the extended two-way fixed effects estimator by Wooldridge (2021) (column 3), or the Callaway and Sant'Anna (2021) estimator (column 4). The dependent variable is the natural log of infant mortality rate per 1,000 live births (Panel A), the natural log of Black infant mortality rate per 1,000 live births (Panel B), or the natural log White infant mortality rate per 1,000 live births (Panel C) in a county-by-year of birth cohort. Each coefficient represents the change in the outcome variable due to receiving a capital appropriation from The Duke Endowment. The weights are the number of births in a county and year. Panels B and C drop observations with zero race-specific births or infant deaths while the log transformation drops county-year observations with zero deaths. Standard errors are clustered at the county level.

Figure D1. Stacked Poisson event studies for effects of Duke support on pooled infant mortality rate



Notes: Extensive margin intent-to-treat estimates. Each figure presents event studies from a separate stacked regression including county and year fixed effects. The dependent variable is the pooled infant mortality rate per 1,000 live births. Each stack includes treated counties from a single treatment timing group and control counties that are not treated within the event time window of $\pm \kappa$. Across the rows of the figure, κ varies from 3 to 6 leads and lags. All samples include 11 stacks. We exclude the 1940 timing groups since forming a complete stack for this group would require using data outside our main sample period from 1922 to 1942. Control variables in column (b) include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Standard errors are estimated using the delta method and are clustered at the county level. The shaded areas are 95% confidence intervals based on these standard errors.

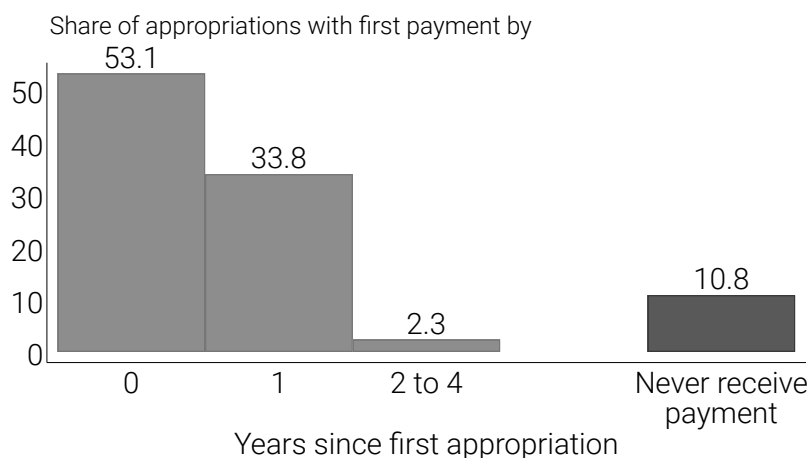
E Duke funding: Heterogeneity by project type

Table E1. Appropriation and payment details by project type

	Mean	S.D.	Min.	Max.	N
<i>All projects</i>					
Appropriations, millions	0.44	0.68	0.01	4.66	130
Payments, millions	0.40	0.65	0.00	4.31	130
<i>All projects, excluding homes for nurses</i>					
Appropriations, millions	0.49	0.72	0.01	4.66	111
Payments, millions	0.44	0.69	0.00	4.31	111
<i>New hospitals or plants</i>					
Appropriations, millions	0.69	0.60	0.04	2.87	34
Payments, millions	0.58	0.70	0.00	2.87	34
<i>Additions</i>					
Appropriations, millions	0.42	0.83	0.01	4.66	38
Payments, millions	0.38	0.79	0.00	4.31	38
<i>Equipment</i>					
Appropriations, millions	0.23	0.71	0.01	3.68	29
Payments, millions	0.21	0.52	0.00	2.58	29
<i>Purchases of existing facilities</i>					
Appropriations, millions	0.68	0.47	0.12	1.79	13
Payments, millions	0.67	0.49	0.00	1.79	13
<i>Homes for nurses</i>					
Appropriations, millions	0.16	0.10	0.03	0.36	19
Payments, millions	0.16	0.11	0.00	0.36	19

Notes: Summary statistics for appropriations and payments from The Duke Endowment in millions of 2017 dollars. The sample includes all appropriations for hospitals in North Carolina initiated between 1927 and 1942 and all payments made on these appropriations up to and including 1962. The unit of observation is a unique appropriation identifier. An appropriation identifier links all appropriations and payments made on those appropriations starting from the initial appropriation for a hospital until all active appropriations have been paid off. Any subsequent appropriations for a given hospital are assigned a separate appropriation identifier. Some appropriations or payments may apply to more than one project type.

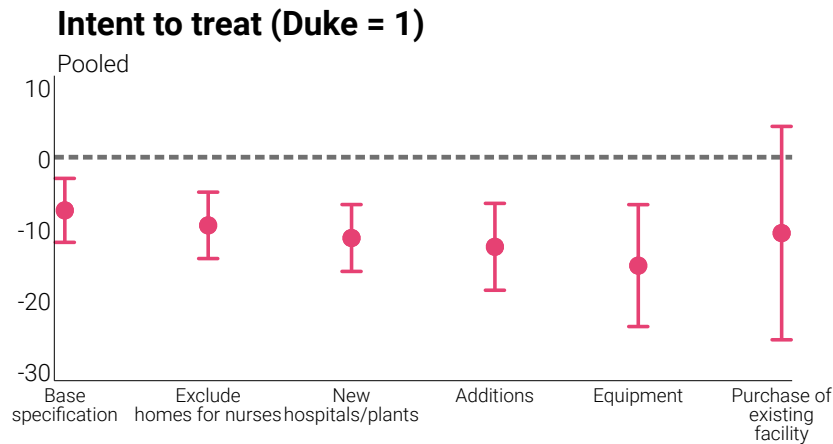
Figure E1. Time from appropriation to payment



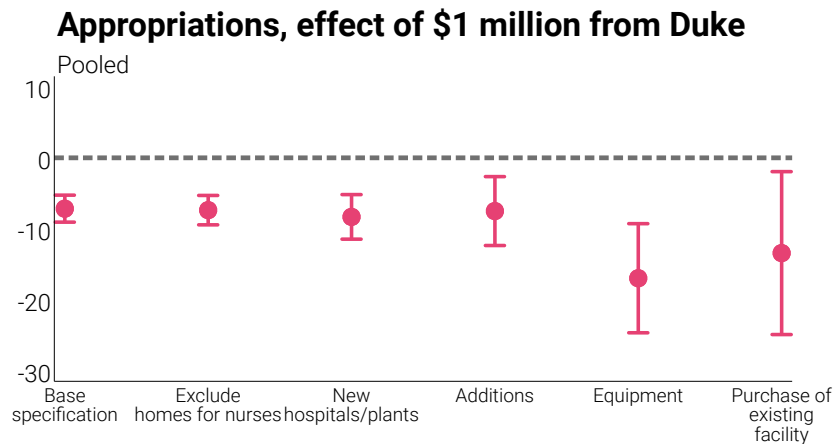
Notes: This figure plots the share of unique appropriations by the number of years after the initial appropriation when the first payment was received by the hospital.

Figure E2. Differential effects of Duke support by project type: Extensive and intensive margin estimates

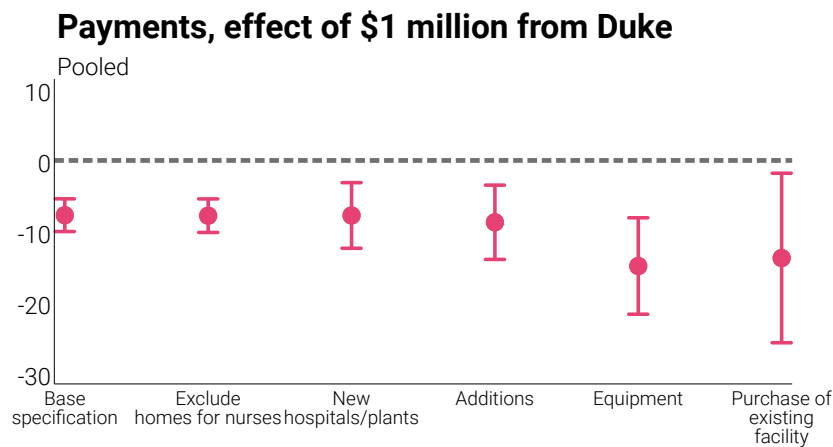
(a) Intent to treat (binary)



(b) Appropriation amount (Effect of 1 \$ million)



(c) Payment amount (Effect of 1 \$ million)



Notes: Each point estimate comes from a separate regression and represents the percent reduction in infant mortality due to Duke support. Treatment is defined as category-specific support. Each category-specific sample drops counties that only received funding for other project categories. Panel A reports an intent to treat analysis with a binary treatment of having received a capital appropriation. A coefficient of -10 would mean that infant mortality declines by 10%. Coefficients reported in the figure are transformed in the following way: $100 \times (\exp(\beta) - 1)$. In panel B, treatment is defined as appropriations to the county while in panel C it is defined as actual payments. Monetary amounts are converted to millions of 2017 dollars. All regressions include county and year fixed effects. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. The weights are the number of births in a county and year. Standard errors are estimated using the delta method and are clustered at the county-level.

F Long-run analysis: Additional details, robustness checks, and event studies

F.1 Event studies

Before reviewing event study results for long-run mortality, it is helpful to recall that our long-run analysis estimates the effect of the *same* treatment as the short-run analysis – Duke support *around the time of birth*. In all event studies, event time is defined as the year of birth minus the year of first capital appropriation from The Duke Endowment. For example, for a county that received its first capital appropriation in 1935, the 1940 birth cohort would receive an event-time value of 5. Thus a positive event time indicates that treatment occurred in the year of birth *or before*. Likewise, -5 in event time value represents a cohort born in a county that received its first capital appropriation five years after the birth year (i.e., when the cohort was five years old).

We report three event study specifications: one that is directly comparable to the short-run analyses (examining the effect in event time in the six years before and after treatment); one that extends the time before birth during which a county could have received support to ten years; and another that prevents compositional changes in the treated counties from muddying the interpretation of the event study estimates. We discuss each of these event studies in turn below.

First, in Figure F2, we consider event study estimates that are directly comparable to the event studies for the short-run analyses. We follow Goodman-Bacon (2021) by making our reference age far from the year of birth so that we can better understand if there are dynamic treatment effects at other points in childhood. We opt for -6 to be our reference year because we have far fewer years of treated observations contributing to these event study analyses than in Goodman-Bacon (2021), as displayed in Figure F3. Thus, event-time coefficient $k < 0$ can be interpreted as the effect of first receiving a capital appropriation at age k relative to the effect of receipt at age six, while event-time coefficient $k \geq 0$ can be interpreted as the effect of receiving a capital appropriation around birth relative to the effect of receipt at age six.

Figure F2 shows a somewhat noisy, but clear pattern, demonstrating that long-run mortality is lower for those who had a capital appropriation in the first year of their life (event time of -1) or earlier (positive event time). The fact that there are no trends in the *positive* event-time coefficients is comforting to us as this indicates that there were no trends differentially affecting those county-birth-year cohorts that received an appropriations before their year of birth. For example, medical technology could have been consistently improving in treated counties, thus gradually improving life expectancy at birth for each birth-year cohort. If this were the case, then we would expect to find larger long-run mortality reductions for cohorts born ten years after an appropriation than for cohorts born three years after an appropriation.

In Figure F4, we consider a second event study that extends the analysis to include cohorts born in a county up to 10 years after the first capital appropriation received by the county. This extension introduces minimal additional imbalance since there are a large number of treated birth cohorts contributing to identify these event time estimates. This extended analysis solidifies the result from the event study in Figure F2 by showing clearly that long-run mortality is lower – and not trending differentially – for those who were born in counties that had received a capital appropriation before or during the year of birth.

While these event studies largely confirm our overall findings, we find them potentially challenging to interpret due to the imbalance in observations across event times displayed in Figure F3. Some of the noise, or possibly some of the effect, could be driven by differences across event times in the set of counties contributing to the estimation of each coefficient. Thus, in Figure F5, we consider a final event study that restricts the set of treated counties to those that are observable from event times -2 to +6 and find clear evidence of a differential effect of Duke support on long-run mortality around the time of birth.

Table F1. Effect of Duke support around time of long-run birth on mortality at ages 56 to 64, adding death rates to main table

	$Y_{ct}^R = \text{Long-run deaths}$			$Y_{ct}^R = \text{Long-run mortality rate}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled long-run deaths or long-run mortality rate</i>						
Percent effect from Duke (=1)	-7.66** (3.08)	-10.07*** (2.66)	-8.99*** (2.81)	-5.11 (5.19)	-8.38* (4.35)	-7.59* (4.48)
Observations	9,000	9,000	9,000	9,000	9,000	9,000
<i>B. Black long-run deaths or long-run mortality rate</i>						
Percent effect from Duke (=1)	-8.04** (3.57)	-8.77*** (2.80)	-7.58** (3.54)	-12.52* (6.52)	-9.25*** (2.91)	-6.59* (3.61)
Observations	8,150	8,150	8,150	8,043	8,043	8,043
<i>C. White long-run deaths or long-run mortality rate</i>						
Percent effect from Duke (=1)	-6.93* (3.83)	-10.56*** (2.92)	-9.53*** (2.86)	-0.57 (6.50)	-7.91 (5.22)	-8.05 (5.08)
Observations	8,630	8,630	8,630	8,630	8,630	8,630
P-value for difference by race	0.82	0.61	0.63	0.25	0.79	0.77
County of birth X Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth X Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The unit of observation is a birth county by birth year by follow-up age triplet. Birth cohorts are restricted to 1932 to 1941. Deaths are restricted to ages 56 to 64 and years 1988 to 2005. In columns 1 to 3, the dependent variable is the number of age-specific deaths. In columns 4 to 6, it is the death rate per 1,000 population in a county-by-year of birth cohort (Surveillance, Epidemiology, and End Results Program 2022). Each coefficient represents the percent reduction in later-life mortality due to receiving a capital appropriation from The Duke Endowment around the time of birth. A coefficient of -10 would mean that later-life mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Control variables in columns 3 and 6 include flexible interactions of age fixed effects with % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. In columns 2 to 3 and 5 to 6, the weights are county-by-year birth cohort size. Panels B and C drop counties that include cohorts with zero births in any year between 1932 and 1941. Observations differ across the two samples as observations that are perfectly separated by either county-of-birth by follow-up age fixed effects or year-of-birth by follow-up age fixed effects are dropped. The bottom row presents a p-value from the interaction of race with our treatment variable from a model that fully interacts all variables with race. All specifications include county-of-birth fixed effects interacted with follow-up age fixed effects and year-of-birth fixed effects interacted with follow-up age fixed effects. Standard errors are estimated using the delta method and are clustered at the county level.

Table F2. Long-run mortality robustness: Cumulative mortality by county and year of birth

	$Y_{ct}^R = \text{Long-run deaths}$		
	(1)	(2)	(3)
<i>A. Pooled long-run deaths</i>			
Percent effect from Duke (=1)	-7.64** (3.09)	-10.04*** (2.67)	-8.92*** (2.82)
Observations	1,000	1,000	1,000
<i>B. Black long-run deaths</i>			
Percent effect from Duke (=1)	-8.08** (3.56)	-8.52*** (2.75)	-7.09* (3.63)
Observations	950	950	950
<i>C. White long-run deaths</i>			
Percent effect from Duke (=1)	-6.94* (3.81)	-10.57*** (2.90)	-9.47*** (2.85)
Observations	960	960	960
P-value for difference by race	0.81	0.54	0.56
County of birth	Yes	Yes	Yes
Year of birth	Yes	Yes	Yes
Weights	No	Yes	Yes
Controls	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The unit of observation is a birth county by birth year. Birth cohorts are restricted to 1932 to 1941. Deaths are restricted to ages 56 to 64 and years 1988 to 2005. The dependent variable is the cumulative number of deaths. The coefficient *Treated* represents the percent reduction in long-run mortality at ages 56 to 64 due to receiving a capital appropriation from The Duke Endowment around the time of birth. A coefficient of -10 would mean that long-run mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Control variables in column 3 include flexible interactions of age fixed effects with % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. In columns 2 to 3, the weights are county-by-year birth cohort size. Panels B and C drop counties that include cohorts with zero births in any year between 1932 and 1941. All specifications include county-of-birth fixed effects interacted with follow-up age fixed effects and year-of-birth fixed effects interacted with follow-up age fixed effects. Observations differ across the two samples as observations that are perfectly separated by either county-of-birth by follow-up age fixed effects or year-of-birth by follow-up age fixed effects are dropped. The bottom row presents a p-value from the interaction of race with our treatment variable from a model that fully interacts all variables with race. Standard errors are estimated using the delta method and are clustered at the county level.

Table F3. Infant and long-run mortality: Adding other Southern states to the samples

	Y_{ct}^R = Infant mortality rate			Y_{ct}^R = Long-run deaths		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
Percent effect from Duke (=1)	-10.25*** (2.08)	-9.54*** (2.20)	-9.56*** (2.20)	-12.46*** (2.99)	-13.13*** (3.11)	-13.13*** (3.11)
Observations	3,801	2,813	2,797	16,794	12,114	12,114
<i>B. Black</i>						
Percent effect from Duke (=1)	-11.60*** (2.78)	-8.53*** (2.69)	-8.56*** (2.68)	-10.22** (4.30)	-13.12*** (4.66)	-13.12*** (4.67)
Observations	3,462	2,531	2,522	16,306	11,996	11,983
<i>C. White</i>						
Percent effect from Duke (=1)	-10.55*** (2.55)	-10.97*** (2.72)	-11.00*** (2.72)	-12.12*** (3.33)	-12.33*** (3.35)	-12.29*** (3.36)
Observations	3,476	2,545	2,533	17,264	12,864	12,756
P-value for difference by race	0.74	0.44	0.44	0.71	0.88	0.88
County of birth FE	Yes	Yes	Yes	No	No	No
County of birth X Age FE	No	No	No	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	No	No	No
Year of birth X Age FE	No	No	No	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Exclude untreated NC	No	Yes	Yes	No	Yes	Yes
Exclude without non-profit hosp.	No	No	Yes	No	No	Yes

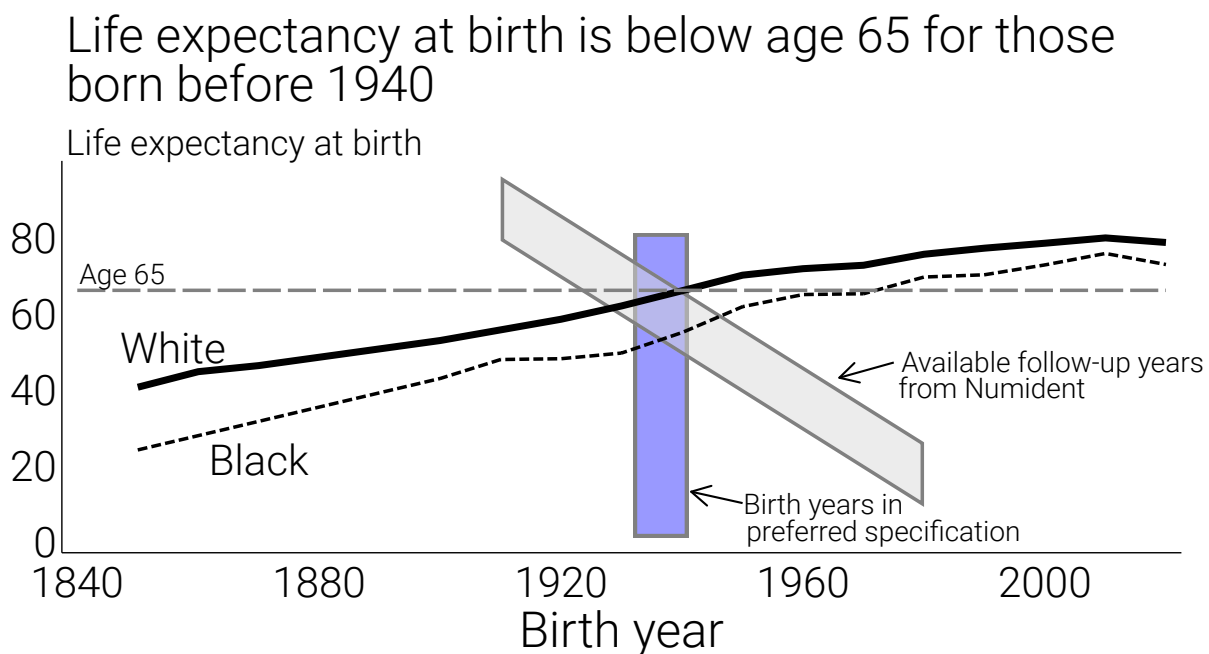
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. In columns 1 to 3, the unit of observation is a county-by-year of birth cell. In columns 4 to 6, the unit of observation is a birth county by birth year by follow-up age triplet. Birth cohorts are restricted to 1932 to 1941. Deaths are restricted to ages 56 to 64 and years 1988 to 2005. In columns 1 to 3, the dependent variable is the number of infant mortality deaths. In columns 4 to 6, it is the number of later-life deaths. Each coefficient represents the percent reduction in infant (columns 1 to 3) and long-run (columns 4 to 6) mortality due to receiving a capital appropriation from The Duke Endowment around the time of birth. A coefficient of -10 would mean that mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Baseline control variables in columns 1 to 3 include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Columns 4 to 6 also include flexible interactions of the baseline controls with age fixed effects. The weights are county-by-year birth cohort size. Panels B and C drop counties that include cohorts with zero births in any year between 1932 and 1941. Observations differ across the two samples as observations that are perfectly separated by any fixed effect are dropped. All infant mortality specifications include county of birth and year of birth fixed effects. All long-run mortality specifications include county-of-birth fixed effects interacted with follow-up age fixed effects and year-of-birth fixed effects interacted with follow-up age fixed effects. The bottom row presents a p-value from the interaction of race with our treatment variable from a model that fully interacts all variables with race. Standard errors are estimated using the delta method and are clustered at the county level.

Table F4. Long-run mortality: Including South Carolina and other Southern states

	Only data from North and South Carolina			Add data from other southern states		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
Percent effect from Duke (=1)	-6.38** (2.47)	-8.87*** (2.09)	-8.76*** (2.21)	-11.64*** (2.30)	-12.09*** (2.38)	-12.08*** (2.39)
Observations	13,140	13,140	13,140	23,031	15,741	15,597
<i>B. Black</i>						
Percent effect from Duke (=1)	-5.79 (3.55)	-7.57** (3.51)	-9.22** (3.98)	-10.94*** (3.56)	-11.92*** (3.87)	-11.92*** (3.87)
Observations	11,890	11,890	11,890	20,927	14,127	14,111
<i>C. White</i>						
Percent effect from Duke (=1)	-7.10** (3.18)	-9.67*** (2.45)	-9.49*** (2.52)	-12.45*** (2.90)	-13.02*** (2.92)	-13.01*** (2.93)
Observations	11,940	11,940	11,940	21,447	14,637	14,565
P-value for difference by race	0.78	0.61	0.95	0.74	0.82	0.82
County of birth X Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth X Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	Yes	Yes	Yes
Include non-Carolina	No	No	No	Yes	Yes	Yes
Exclude untreated Carolina	No	No	No	No	Yes	Yes
Exclude without non-profit hosp.	No	No	No	No	No	Yes

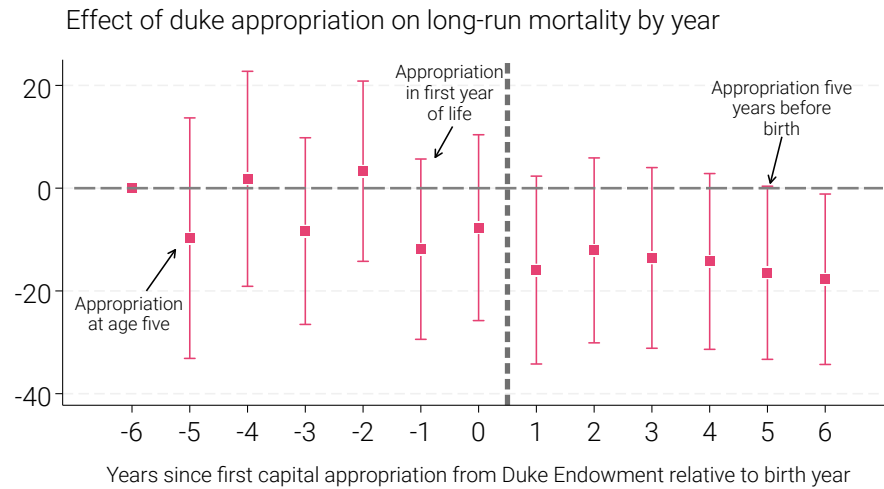
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The unit of observation is a birth county by birth year by follow-up age triplet. Birth cohorts are restricted to 1932 to 1941. Deaths are restricted to ages 56 to 64 and years 1988 to 2005. The dependent variable is the long-run mortality rate per 1,000 live births. Each coefficient represents the percent reduction in long-run mortality due to receiving a capital appropriation from The Duke Endowment around the time of birth. A coefficient of -10 would mean that long-run mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Baseline control variables in columns 3 to 6 include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Columns 4 to 6 also include flexible interactions of the baseline controls with age fixed effects. The weights are county-by-year birth cohort size. Panels B and C drop counties that include cohorts with zero births in any year between 1932 and 1941. Observations differ across the two samples as observations that are perfectly separated by either county-of-birth by follow-up age fixed effects or year-of-birth by follow-up age fixed effects are dropped. The bottom row presents a p-value from the interaction of race with our treatment variable from a model that fully interacts all variables with race. All specifications include county-of-birth fixed effects interacted with follow-up age fixed effects and year-of-birth fixed effects interacted with follow-up age fixed effects. Standard errors are estimated using the delta method and are clustered at the county level.

Figure F1. Life expectancy by race, coverage of Numident data, and the preferred cohorts



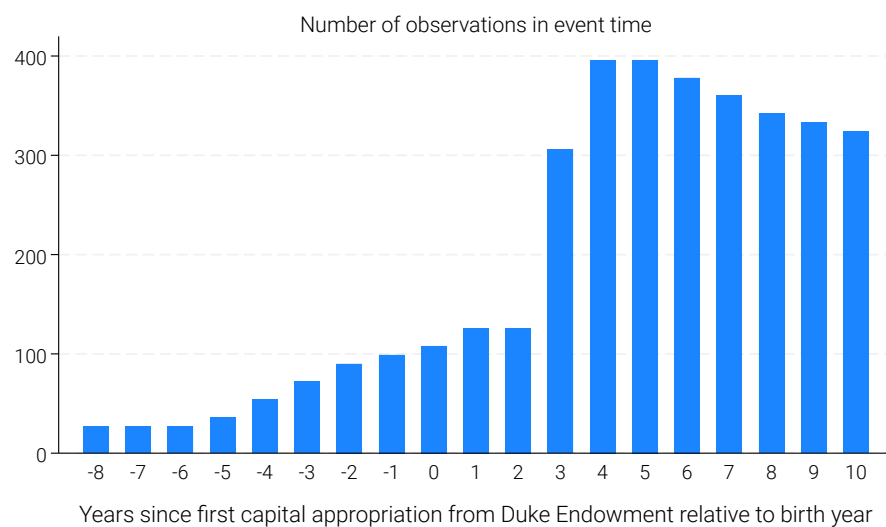
Notes: Life expectancy by birth cohort for White (solid line) and Black (dashed line) individuals from Haines (2008); Arias (2014); Arias et al. (2021). The parallelogram shaded in grey represents the set of birth cohorts and ages at death available in the Numident. The rectangle shaded in blue represents the set of birth cohorts used in our preferred specification for long-run mortality.

Figure F2. Long-run analysis: Event study estimates



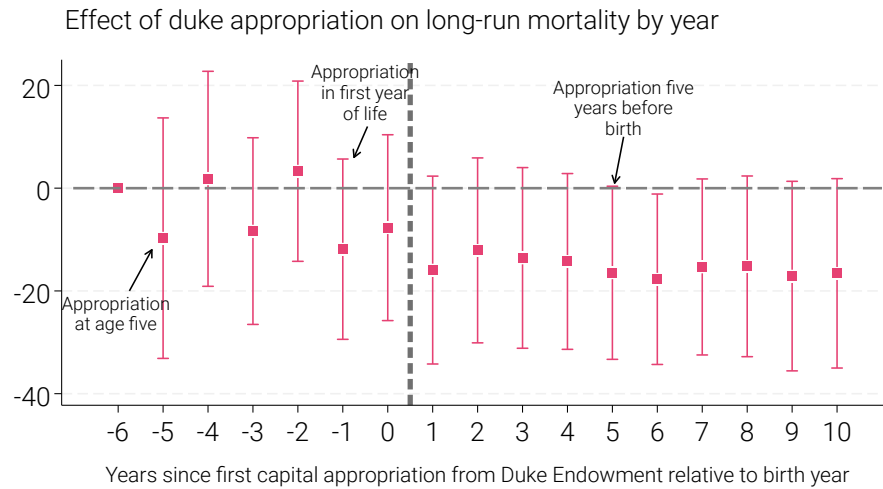
Notes: Plot contains long-run event-study estimates the effect of the *same* treatment as the short-run analysis – Duke support *around the time of birth*. In all event studies, event time is defined as the year of birth minus the year of first capital appropriation from The Duke Endowment. For example, for a county that received its first capital appropriation in 1935, the 1940 birth cohort would receive an event-time value of 5. Thus a positive event time indicates that treatment occurred in the year of birth *or before*. Likewise, -5 in event time value represents a cohort born in a county that received its first capital appropriation five years after the birth year (i.e., when the cohort was five years old). An observational unit is a county-of-birth, year-of-birth, follow-up age level. Coefficients reported are transformed in the following way: $100 \times (\exp(\beta) - 1)$. The dependent variable is deaths in for a given county-of-birth, year-of-birth, follow-up age group. Regressions are weighted by county-by-year of birth cohort size. All regressions include county-of-birth \times follow-up age and year-of-birth \times follow-up age fixed effects. Standard errors are estimated using the delta method and are clustered by county of birth. 95% confidence intervals are reported in brackets.

Figure F3. Unbalanced composition of treated observations across event times



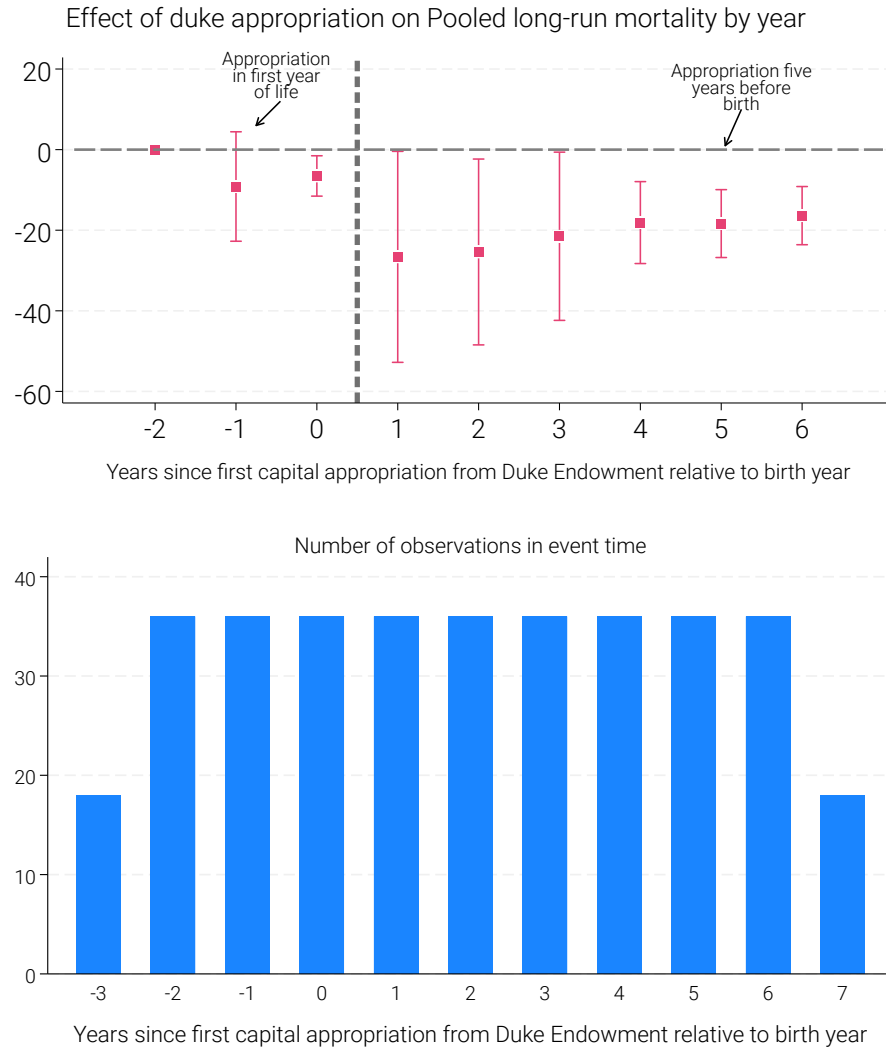
Notes: This Figure displays the treated number of observations for each event time that are included in the event studies reported in Figures F2 and F4.

Figure F4. Long-run analysis: Event study extended to ten years following treatment



Notes: This figure presents results from the same specification as in Figure F2, but extended to +10 in event-time. The long-run event-study estimates the effect of the *same* treatment as the short-run analysis – Duke support *around the time of birth*. In all event studies, event time is defined as the year of birth minus the year of first capital appropriation from The Duke Endowment. For example, for a county that received its first capital appropriation in 1935, the 1940 birth cohort would receive an event-time value of 5. Thus a positive event time indicates that treatment occurred in the year of birth *or before*. Likewise, -5 in event time value represents a cohort born in a county that received its first capital appropriation five years after the birth year (i.e., when the cohort was five years old). An observational unit is a county-of-birth, year-of-birth, follow-up age level. Coefficients reported are transformed in the following way: $100 \times (\exp(\beta) - 1)$. The dependent variable is deaths in for a given county-of-birth, year-of-birth, follow-up age group. Regressions are weighted by county-by-year of birth cohort size. All regressions include county-of-birth \times follow-up age and year-of-birth \times follow-up age fixed effects. Standard errors are estimated using the delta method and are clustered by county of birth. 95% confidence intervals are reported in brackets.

Figure F5. Long-run analysis: Event study using the same composition of treated units from -3 to 7 event time



Notes: This figure considers a restricted long-run specification that plots the event-time treatment path from a restricted set of treated counties, those that are observable from event times -2 to +6. The untreated cohorts are unchanged. The top panel presents the event-study coefficients and the bottom panel presents the treated number of observations for each event time that are included in the balanced event study estimates above.

G Other results

Stillborn infants: Here, we examine potential measurement error issues related to recording stillbirths. Our individual-level death certificate data for North Carolina include some reported stillbirths, but only until 1932. In our main specifications, we follow published infant mortality statistics in restricting attention to live births, thereby excluding stillbirths. Specifically, we use unnamed infants who died on the day of birth as a proxy for stillborn infants and exclude both unnamed infants and reported stillbirths from our measures of infant mortality. In Online Appendix Table G5, we show that our results are insensitive to including stillbirths in our infant mortality measure, irrespective of the exact definition of a stillborn infant. When constructing infant mortality rates, we adjust the numerator by the number of stillbirths. The results are unchanged if we include deaths of unnamed infants in our measure of infant mortality, regardless of whether we do so for the period when stillbirths were reported (1922 to 1932, columns 1 and 2) or for the entire sample period (column 3). Likewise, they are unaffected by including reported stillbirths as well (column 4). Thus, our infant mortality results are unlikely to be driven by measurement error in the reporting of stillbirths.

Using our proxy for stillbirths (unnamed infants who died on the day of birth) as the outcome, we find that Duke support reduced stillborn deaths per 1,000 (live and still) births by up to 20% (columns 5 and 6) and to a similar extent for both Black fetuses and White fetuses. Furthermore, this finding is independent of whether we use all sample years (column 5) or restrict the sample to the period 1922 to 1932 when reported stillbirths were explicitly included in our data (column 6). Interestingly, there is no effect on reported stillbirths per 1,000 (live and reported still) births (column 7), which suggests that our result based on the stillborn proxy is not driven by changes in reporting procedures during the time period, although we acknowledge that these estimates are fairly imprecise (as is clearly visible in Figure 5).

Shift-share DiD: In Equation 7, we estimate the effects of sulfa drugs on infant mortality in our main estimation sample using a within-North-Carolina shift-share design:³⁵

$$Y_{ct}^R = \exp(\delta_0 + \delta_1 \text{Pneumonia mortality}_c \times \text{Post sulfa}_t + \zeta_c + \eta_t + \Theta \mathbf{X}_{ct}) \epsilon_{ct} \quad (7)$$

where Y_{ct}^R , ζ_c , η_t , and \mathbf{X}_{ct} are defined as in Equation 3. $\text{Pneumonia mortality}_c$ is defined as the average county-level pneumonia mortality rate from 1922 to 1926 (our share factor) while Post sulfa_t takes the value of 1 for the years 1937 and onward, and zero for prior years (our shift factor).³⁶ Since $\text{Pneumonia mortality}_c$ is perfectly collinear with county fixed effects (ζ_c), while Post sulfa_t is perfectly collinear with birth cohort fixed effects (η_t), they are not separately identified. Nonetheless, the interaction term is identified and the coefficient δ_1 can be interpreted as the causal effect of innovation in sulfa drugs on infant mortality, provided that the standard difference-in-differences assumptions hold. In particular, since we only have one shock, we rely on the exogeneity of shares for identification (Goldsmith-Pinkham et al. 2020).³⁷ We cluster the standard errors (ϵ_{ct}) by county of birth to account for correlated errors within a county.

³⁵This approach is different from Jayachandran et al. (2010) and more closely resembles the identification strategies used by Bhalotra et al. (2017) and Lazuka (2020), but uses variation at a finer level of geography.

³⁶We use a 5-year average to define our shares for two reasons. First, the single-year pneumonia mortality rate can be volatile (especially in smaller counties) due to exogenous weather and health shocks. Second, it is ex-ante not clear which year we should chose as our baseline. Our results are robust to using shares from any specific year between 1922 and 1926, but as expected, the exact point estimates change somewhat.

³⁷We define pre-shock shares based on the years 1922 to 1926 rather than the years immediately prior to the discovery of sulfa because we need to ensure that they are unaffected by Duke support which started in 1927. This requirement is dictated by the exogeneity of shocks assumption which we need to interpret the interaction effect between the two interventions in a causal way.

Online Appendix Table G6 presents the results of estimating Equation 7 for the pooled infant mortality rate (panel A), the Black infant mortality rate (panel B) and the White infant mortality rate (panel C), both without (columns 1 and 2) and with (column 3) controls. The estimates are scaled by the interquartile range of baseline pneumonia mortality and can be interpreted as the percent reduction in infant mortality rate due to the availability of sulfa drugs when moving from the 25th to the 75th percentile of the baseline pneumonia mortality rate. Our findings in the preferred specification (column 3) suggest that infant mortality declined by 5.2%. To the best of our knowledge, prior papers on sulfa drugs have not analyzed effects on infant mortality. However, Jayachandran et al. (2010) using a different identification strategy found effects on maternal mortality in the range of 24% to 36%. Overall, we view our effect sizes as plausible, especially given how closely they align with the effects reported in Table 3.

Table G1. Balancing test: Effects of Duke support on control variables

	(1) Mean % difference	(2) p-value
<i>A. Pei et al. (2019) balancing test</i>		
% Illiterate	−4.30	0.18
% Black	−0.94	0.23
% Urban	−3.79	0.13
Retail sales per capita	0.46	0.88
County health department present (=1)	−0.13	0.99
<i>B. F-Test for joint significance of controls</i>		
p-value		0.192

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This analysis follows Pei et al. (2019) by re-estimating our preferred analysis using a Poisson regression and fixed effects, but where each control is included as the dependent variable. Here, a % difference of -10 would indicate that Duke support is associated with a change in outcome of 10% in treated versus control counties. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Standard errors are estimated using the delta method and are clustered at the county-level. Note we do not include % other race since within some counties this measure does not vary at all across our sample years, thus the county fixed effect for each of these counties perfectly predicts the outcome and these observations are dropped from the analysis. Since we prefer to present the test for a balanced sample we omit % other race.

Table G2. Extensive margin intent-to-treat effect of Duke support on maternal mortality

	Number of Deaths			Death rate per 1,000 women		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Maternal mortality</i>						
Percent effect from Duke (=1)	-18.26** (8.39)	-22.86*** (7.33)	-24.25*** (7.83)	-3.60 (17.43)	-16.84* (8.55)	-18.37** (8.43)
Observations	1,100	1,100	1,100	1,100	1,100	1,100
<i>B. Maternal mortality, resident</i>						
Percent effect from Duke (=1)	-14.15* (7.23)	-14.96*** (4.97)	-16.49*** (5.19)	-6.32 (17.30)	-13.68* (7.47)	-14.44** (7.04)
Observations	1,100	1,100	1,100	1,100	1,100	1,100
County of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. In columns 1 to 3, the dependent variable is the number of maternal deaths in a county and year, which we calculate by multiplying the published maternal mortality rate per 1,000 live births and the number of resident births. In columns 4 to 6, the dependent variable is the maternal mortality rate per 1,000 live births. In panel A, maternal mortality is reported by county of occurrence, while in panel B it is reported by county of residence. Each coefficient represents the percent reduction in maternal mortality due to receiving a capital appropriation from The Duke Endowment. A coefficient of -10 would mean that maternal mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. In columns 2 to 3 and columns 5 to 6, the weights are the number of births in a county and year. Control variables in columns 3 and 6 include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Standard errors are estimated using the delta method and are clustered at the county-level.

Table G3. Extensive margin intent-to-treat effect of Duke support on fertility

	Number of Births			Births per 1,000 women		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled births</i>						
Percent effect from Duke (=1)	2.05 (1.97)	0.52 (2.17)	-1.19 (2.20)	-0.43 (1.56)	-2.59* (1.51)	-2.22 (1.46)
Observations	2,100	2,100	2,100	2,100	2,100	2,100
<i>B. Black births</i>						
Percent effect from Duke (=1)	3.25 (2.98)	3.45 (3.35)	2.56 (2.78)	3.25 (2.98)	-0.89 (2.19)	-0.29 (1.87)
Observations	2,079	2,079	2,079	2,079	2,079	2,079
<i>C. White births</i>						
Percent effect from Duke (=1)	1.90 (2.25)	-1.29 (2.48)	-2.68 (2.59)	2.27 (1.87)	-2.11 (1.71)	-1.67 (1.75)
Observations	2,079	2,079	2,079	2,079	2,079	2,079
P-value for difference by race	0.56	0.13	0.13	0.12	0.55	0.55
County of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. In columns 1 to 3, the dependent variable is the number of births in a county and year, and in columns 4 to 6, it is the birth rate per 1,000 women in the population. Each coefficient represents the percent reduction in fertility due to receiving a capital appropriation from The Duke Endowment. A coefficient of -10 would mean that fertility declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. In columns 2 to 3 and columns 5 to 6, the weights are the number of women in a county and year. In panel A the number of women includes all women in the population, while in panels B and C it is the number of Black and White women, respectively. Panels B and C drop counties that ever have zero race-specific births between 1922 and 1942. Control variables in columns 3 and 6 include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Standard errors are estimated using the delta method and are clustered at the county-level.

Table G4. Effect of Duke support on infant mortality rate by timing of death

	Deaths on: day 0	Excluding those who died in the first:		
		day	week	month
	(1)	(2)	(3)	(4)
<i>A. Pooled infant mortality rate</i>				
Percent effect from Duke (=1)	-8.86* (4.63)	-6.07*** (2.12)	-8.25*** (2.40)	-9.58*** (2.79)
Observations	2,100	2,100	2,100	2,100
<i>B. Black infant mortality rate</i>				
Percent effect from Duke (=1)	-16.45*** (5.07)	-12.06*** (3.29)	-14.53*** (3.68)	-16.35*** (3.85)
Observations	1,974	1,995	1,995	1,995
<i>C. White infant mortality rate</i>				
Percent effect from Duke (=1)	-6.06 (5.94)	-3.14 (2.91)	-4.62 (3.33)	-5.30 (3.60)
Observations	1,974	1,995	1,995	1,995
P-value for difference by race	0.12	0.04	0.04	0.02
County of birth FE	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Extensive margin intent-to-treat estimates. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression based on preferred specification from column 6 of Table 3. In column 1 the dependent variable is the same day (day 0) infant mortality rate per 1,000 live births. In columns 2 to 4, they are the infant mortality rates per 1,000 live births excluding deaths on the same day, and within the first week and first month, respectively. Each coefficient represents the percent reduction in infant mortality due to receiving a capital appropriation from The Duke Endowment. A coefficient of -10 would mean that infant mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. All specifications are weighted by the number of births in a county and year. Panels B and C drop the counties that ever had zero births (columns 1 to 4) or zero deaths on day 0 (column 1) between 1922 and 1942. Standard errors are estimated using the delta method and are clustered at the county-level.

Table G5. Effects of Duke support: Robustness to including stillbirths and unnamed infants

	Adding to the main sample:				Including only:		
	Stillborn	Unnamed	Unnamed	Stillborn + Unnamed	Unnamed	Unnamed	Stillborn
	1922-32		All years		1922-32		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
<i>A. Pooled infant mortality rate</i>							
Percent effect from Duke (=1)	-6.41*** (2.13)	-8.06*** (2.36)	-7.24*** (2.12)	-6.41*** (1.98)	-18.76*** (6.45)	-15.52** (6.93)	-1.29 (8.55)
Observations	2,100	2,100	2,100	2,100	2,100	1,100	1,100
<i>B. Black infant mortality rate</i>							
Percent effect from Duke (=1)	-13.08*** (3.15)	-13.61*** (3.19)	-12.64*** (2.84)	-12.32*** (2.74)	-17.90** (8.70)	-13.56 (9.49)	-7.55 (9.55)
Observations	2,016	1,995	1,995	2,016	1,995	1,067	1,034
<i>C. White infant mortality rate</i>							
Percent effect from Duke (=1)	-4.79* (2.58)	-4.71 (2.93)	-3.72 (2.78)	-4.11 (2.48)	-20.20*** (7.05)	-19.10** (8.01)	5.95 (14.70)
Observations	2,016	1,995	1,995	2,016	2,016	1,078	1,034
P-value for difference by race	0.03	0.02	0.02	0.02	0.80	0.61	0.35
County of birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports extensive margin intent-to-treat estimates for the effects of Duke support on different measures of stillborn deaths and the robustness of the main results to including these deaths in the mortality outcomes. In columns 1 to 4 we modify our main infant mortality measure from column 6 of Table 3 by adding stillbirths to the numerator (death count). We use three measures of stillbirths: reported stillbirths (column 1), unnamed infants (columns 2 and 3), and the combination of the two measures (column 4). In columns 5 to 7, we explore effects of Duke support on stillbirths directly. In columns 1 to 2 and 6 to 7 we report results for the period 1922 to 1932 because reported stillbirths are only included in our data for these years. See Table 3 for a description of the specifications. Observations may differ across the samples by race as observations are dropped that are perfectly separated any fixed effects. Standard errors are estimated using the delta method and are clustered at the county-level.

Table G6. Effects of sulfa drugs on infant mortality: Shift-share difference-in-differences

	(1)	(2)	(3)
<i>A. Pooled infant mortality rate</i>			
Percent effect from $\text{Pneumonia}_{\text{IQR}} \times \text{Post sulfa}$	-3.70 (2.43)	-6.36*** (1.99)	-5.16** (2.49)
Observations	2,100	2,100	2,100
<i>B. Black infant mortality rate</i>			
Percent effect from $\text{Pneumonia}_{\text{IQR}} \times \text{Post sulfa}$	-19.83*** (7.29)	-5.16* (2.77)	-3.65 (3.52)
Observations	1,995	1,995	1,995
<i>C. White infant mortality rate</i>			
Percent effect from $\text{Pneumonia}_{\text{IQR}} \times \text{Post sulfa}$	-3.92 (2.86)	-7.08*** (2.49)	-5.85** (2.89)
Observations	1,995	1,995	1,995
County of birth FE	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes
Controls	No	No	Yes
Weights	No	Yes	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The dependent variable is the infant mortality rate per 1,000 live births. Years 1937 and onward are considered post sulfa (shift factor). The baseline shares are county-level pneumonia mortality rates per 100,000 population and are calculated as a simple average over the years 1922 to 1926. The displayed parameter of interest is an interaction between these two variables. See Equation 7 for details. Each coefficient represents the percent reduction in infant mortality due to the availability of sulfa drugs when moving from the 25th to 75th percentile of baseline pneumonia mortality. A coefficient of -10 would mean that infant mortality declines by 10%. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. In columns 2 and 3 the weights are the number of births in a county and year. Panels B and C drop counties that ever have zero births between 1922 and 1942. Standard errors are estimated using the delta method and are clustered at the county-level.

H Panel construction and sample size

In our main specification for short-run outcomes in panel A of Table 3, we estimate the effects of exposure to Duke support on the pooled infant mortality rate per 1,000 live births using a balanced panel with all 100 counties in North Carolina and 21 years of data (1922 to 1942) for a total of 2,100 observations. We take the following steps to deal with observations having zero births or zero deaths, as well as logical inconsistencies between birth and death counts:

1. There are 12 observations among 4 counties for which the number of deaths of Black infants exceeds the number of Black births. For these observations we replace the birth count with the mortality count.
2. There are 5 counties and 24 observations with zero Black births. In our main specifications for infant mortality rates by race (panels B and C of Table 3), we exclude these 5 counties in order to maintain a balanced panel. The observations with zero Black births will drop out since the weights are undefined for these observations.
3. After dropping all observations for the 5 counties that ever have zero Black births, there are 19 counties and 101 observations with zero Black deaths and non-zero Black births. There are 2 other counties and 4 observations with zero White deaths and non-zero White births. In specifications with the log of the infant mortality rate as the dependent variable we drop these 21 counties from the sample, in addition to the 5 counties already dropped (for a total of 26 counties). Our main specification is estimated using Poisson pseudo-maximum likelihood which can handle the presence of zero values for the dependent variable, unlike a log-level specification that would drop these observations.
4. We exclude the aforementioned 26 counties from estimation samples in specifications with the log of the infant mortality rate as the dependent variables in the following exhibits: Figure 5 where noted, Online Appendix Tables D2 and I1, and Online Appendix Figure I1.

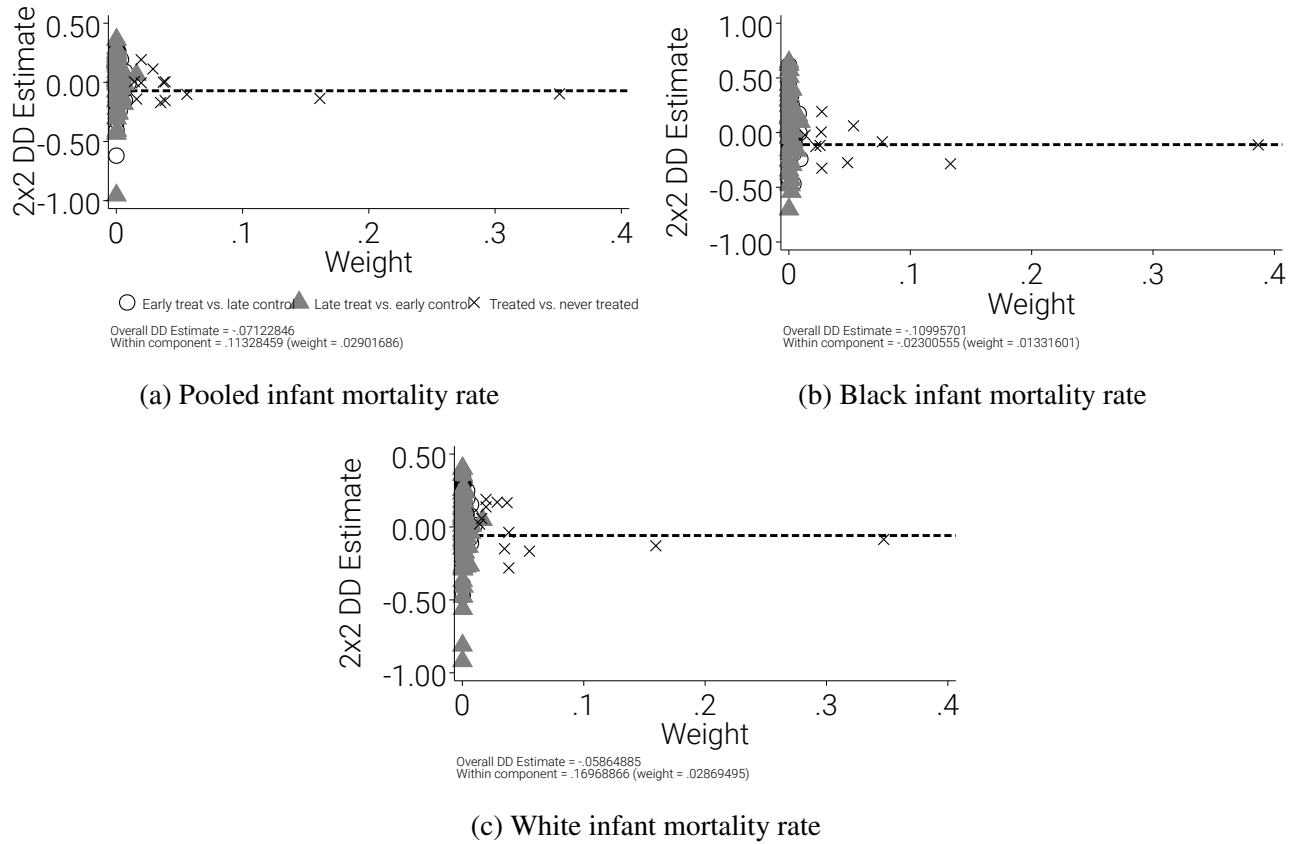
I Event study diagnostics

Table I1. Goodman-Bacon (2021a) decomposition diagnostic

Type of DD comparison	Average Estimate	Number of 2x2 Comparisons	Total Weight
<i>A. Pooled infant mortality rate</i>			
Earlier Treated vs. Later Treated Controls	−0.017	66	0.092
Later Treated vs. Earlier Treated Controls	−0.009	66	0.091
Treated vs. Untreated Controls	−0.084	12	0.817
Average DD estimate	−0.071	144	1.000
<i>B. Black infant mortality rate</i>			
Earlier Treated vs. Later Treated Controls	0.015	55	0.077
Later Treated vs. Earlier Treated Controls	−0.044	55	0.083
Treated vs. Untreated Controls	−0.128	11	0.840
Average DD estimate	−0.110	121	1.000
<i>C. White infant mortality rate</i>			
Earlier Treated vs. Later Treated Controls	−0.048	66	0.096
Later Treated vs. Earlier Treated Controls	0.043	66	0.096
Treated vs. Untreated Controls	−0.072	12	0.808
Average DD estimate	−0.059	144	1.000

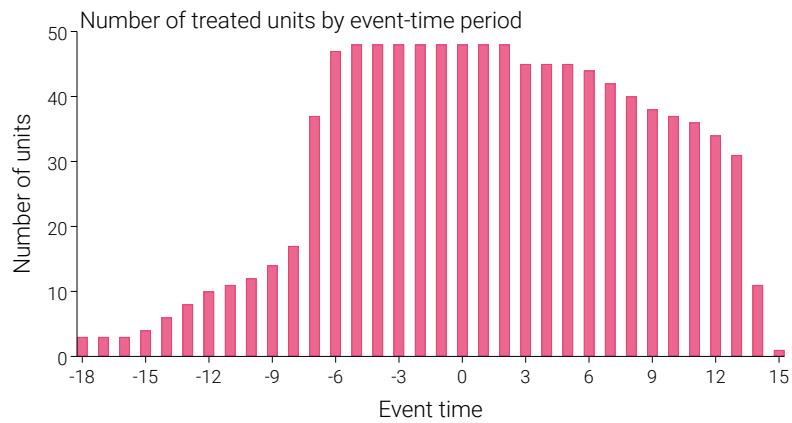
Notes: The table decomposes the static DiD two-way fixed effects estimate reported in column 1 of Online Appendix Table D2 into the average estimate and total weight contributed by earlier versus later treated comparisons, later versus earlier treated comparisons, and treated versus untreated comparisons, as well as the number of unique 2x2 comparisons found in each category. The specification does not include controls except for county and year fixed effects, is not weighted, and uses the log of the infant mortality rate as the dependent variable. The sample drops 26 counties that ever had zero Black or White births or zero Black or White deaths in a year during the sample period.

Figure I1. Goodman-Bacon (2021a) decomposition diagnostic



Notes: Figure I1 decomposes the static DiD two-way fixed effects estimate reported in column 1 of Online Appendix Table D2 into separate 2x2 DiD components. The specification does not include controls except for county and year fixed effects, is not weighted, and uses the log of the infant mortality rate as the dependent variable. The sample drops 26 counties that ever had zero Black or White births or zero Black or White deaths in a year during the sample period. The figure depicts the distribution of all unique treatment timing comparisons used to identify $\widehat{\delta}^{DD}$. For example, one symbol may represent a comparison between counties treated in 1935 and counties treated in 1937. The horizontal pink dotted line displays the overall DiD estimate.

Figure I2. Number of treated counties by event-time period



Notes: This figure plots the number of treated counties in each event-time period from $t = -18$ to $t = 15$, which corresponds to the full set of event-time indicators included in the event studies plotted in the bottom row of Figure 4.

J Instrumental variables, Alternate samples and non-Carolina control counties

Instrumental variables: We estimate an instrumental variables specification to provide additional evidence that our results are not driven by selection. The instrument interacts temporal and cross-sectional sources of variation. The first term in the interaction is the cumulative returns of The Duke Endowment's assets. We obtained original financial statements of The Duke Endowment containing these data from the Joseph and Matthew Payton Philanthropic Studies Library (The Duke Endowment 1925b). For exactness, we consider returns less operational overhead and less 20% (which is placed back into the principal, as outlined in the Indenture of Trust). However, our results are not sensitive to this decision and are virtually identical when we use total returns. The second term in the interaction takes the value of one if a county is in North Carolina and had a not-for-profit hospital in the year before The Duke Endowment began appropriating money for capital projects, and takes the value zero otherwise. We use two comparison groups, one consisting of counties outside of North Carolina with not-for-profit hospitals that were ineligible for Duke support, and another consisting of counties in North Carolina without existing not-for-profit hospitals.

Following our previous intensive margin specifications in Table 4, Table J1 presents results using both capital appropriations (left) and payments (right). Each set of specifications includes a Poisson regression, an OLS regression with the dependent variable equal to the natural log of the infant mortality rate, the first-stage regression of the potentially endogenous measure of Duke support on the instrument, the reduced form regression of log infant mortality rate on the instrument, and an instrumental variable specification. To simplify the exposition of this analysis, we report results of the instrumental variables analysis using the natural log of the pooled infant mortality rate as the dependent variable. Results using Black or White infant mortality measures are similarly larger than their accompanying non-IV estimates, but are not presented due to space constraints. We consider the effects for two samples. The first sample (panel A) ranges from 1922 to 1940 and includes only those North Carolina counties that had a not-for-profit hospital in the year before The Duke Endowment began capital appropriations, and all non-Carolina Southern counties that had a not-for-profit hospital. These data are from Fishback et al. (2007), whose data series stops in 1940. The second sample (panel B) includes all North Carolina counties from 1922-1940. We keep the years the same between Panels A and B for ease of comparison.

These instrumental variables results help to address various sources of potential selection. For example, if counties selected into treatment because they were better suited to take advantage of the modernization efforts of The Duke Endowment, it could be the case that our non-IV estimates overstate the effect of The Duke Endowment. Similarly, it could be the case that The Duke Endowment targeted projects which it believed would have the highest investment returns, although such selective behavior should be less pronounced in larger windfall years because resources are more plentiful. Finally, it could be the case that counties with non-profit hospitals were on different mortality trends than other counties. Given that we find larger, negative effect, these concerns are mitigated.

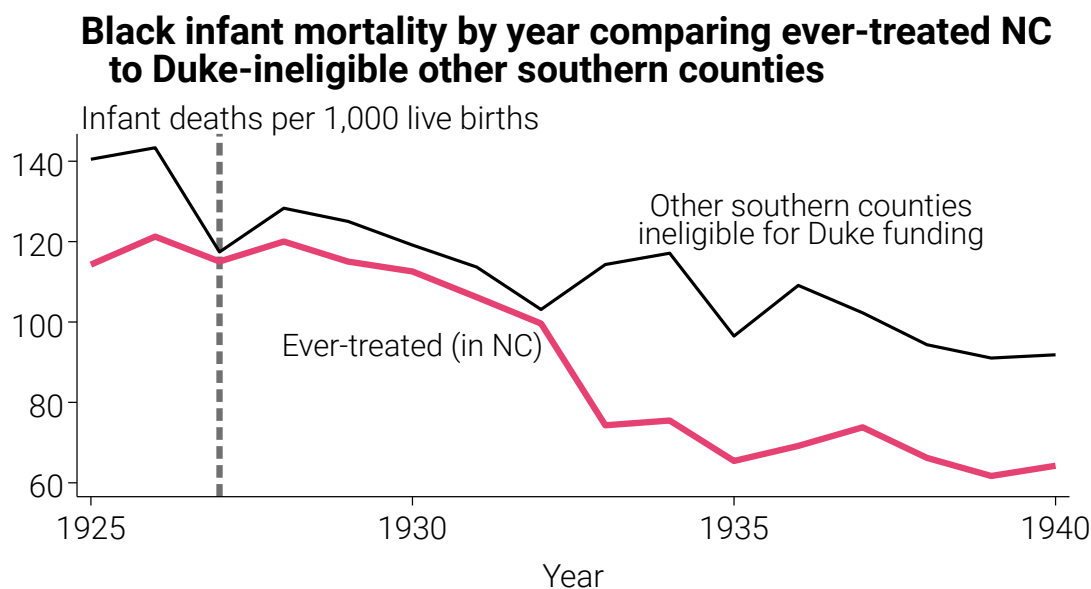
Table J1. Effect of Duke support on pooled infant mortality rate: Intensive margin instrumental variables estimates

Specification:	Appropriations					Payments				
	Poisson	OLS	First stage	Reduced form	IV	Poisson	OLS	First stage	Reduced form	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Y_{ct}^R :	IMR	ln(IMR)	Appropriations	ln(IMR)	ln(IMR)	IMR	ln(IMR)	Payments	ln(IMR)	ln(IMR)
<i>A. Southern counties with non-profit hospital (1922-1940)</i>										
Percent effect from \$1 million of Duke support	-7.84*** (1.12)	-7.28*** (1.36)			-14.58*** (4.00)	-9.73*** (2.05)	-8.65*** (2.04)			-16.17*** (4.36)
Anderson-Rubin 95% Confidence Set					[-25.88, -6.52]***					[-28.08, -7.34]***
tF 95% Confidence Interval					[-25.06, -2.65]**					[-27.42, -3.17]**
(Endowment returns, billions) X 1(Non-profit hospital before Duke)			0.21*** (0.06)	-2.95*** (0.89)				0.18*** (0.05)	-2.95*** (0.89)	
Observations	2,965	2,961	2,965	2,961	2,961	2,965	2,961	2,965	2,961	2,961
<i>B. All NC counties (1922-1940)</i>										
Percent effect from \$1 million of Duke support	-6.92*** (1.19)	-6.77*** (1.31)			-17.42*** (6.53)	-8.17*** (1.70)	-7.75*** (1.74)			-17.37*** (6.11)
Anderson-Rubin 95% Confidence Set					[-36.91, -5.08]**					[-33.88, -5.32]**
tF 95% Confidence Interval					[-34.56, 4.20]					[-32.54, 1.21]*
(Endowment returns, billions) X 1(Non-profit hospital before Duke)			0.18*** (0.05)	-3.06*** (1.11)				0.18*** (0.05)	-3.06*** (1.11)	
Observations	1,900	1,900	1,900	1,900	1,900	1,900	1,900	1,900	1,900	1,900

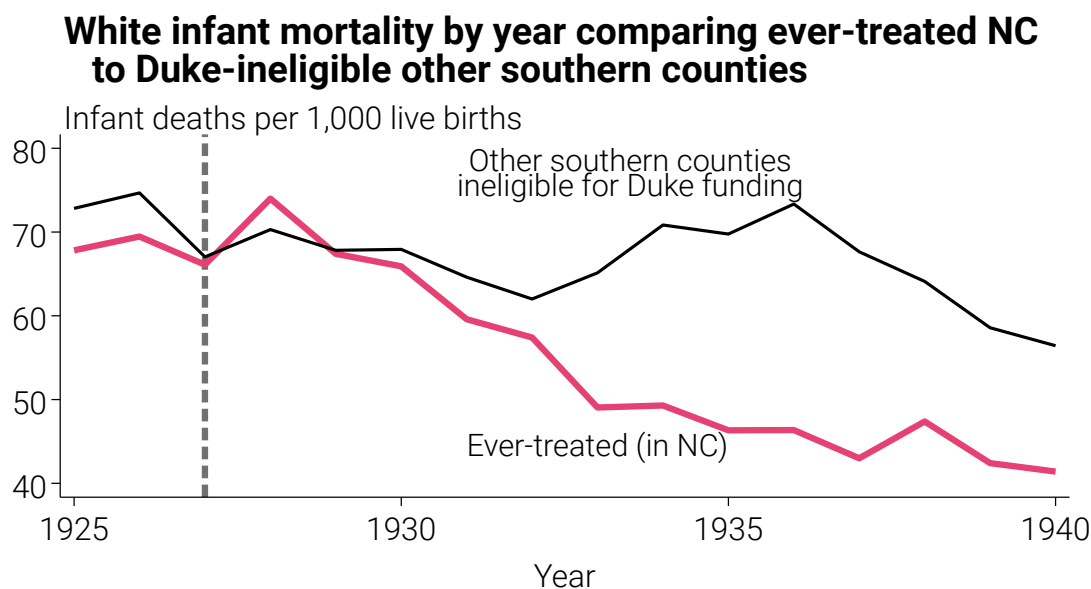
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression. Columns 1 through 5 focus on cumulative appropriations from The Duke Endowment in 2017 \$. Columns 6 through 10 focus on cumulative payments from The Duke Endowment in 2017 \$. All specifications in the top panel A, use the same sample that includes all North Carolina counties that had a not-for-profit hospital in the year before The Duke Endowment began capital appropriations and all non-Carolina Southern counties from Bailey et al. (2015) that have a not-for-profit hospital. The Bailey et al. (2015) data are unbalanced before 1930 and after 1940. Thus, to ensure balance across years, the data in this panel are restricted to birth years 1930 to 1940. All specifications in the bottom panel B, use the same sample that is used in our preferred analysis from Table 3, which includes all North Carolina counties from 1922 to 1942. Columns 1 and 6 show the effect of \$1 million of Duke appropriation, or payments, on infant mortality using a Poisson estimator. These results are analogous to those presented in Table 4, except the years are from 1922-1940 for comparison with top panel. Columns 2 and 7 conduct the same exercise, but use OLS and a natural log transform of the infant mortality rate. Columns 3 and 8 show the first-stage relationship between the instrumental variable and the potentially endogenous variable, cumulative appropriations (or payments). In both cases the instrumental variable is the same interaction. The first term in the interaction is the cumulative returns of The Duke Endowment's assets less operational overhead and 20% which is placed back into the principle, following instructions in the Indenture of Trust. The second term in the interaction takes the value of one if a county in North Carolina had a not-for-profit hospital in the year before The Duke Endowment began appropriating money for capital projects. This instrument exploits the fact that non-North Carolina counties were ineligible for Duke-support, that most Duke-support went to improve existing not-for-profit hospitals, and that as more money was earned by the Endowment there was greater ability to appropriate funds. Columns 4 and 9 show the reduced form relationship between the instrumental variable and the natural log of the infant mortality rate. Columns 5 and 10 show results from an instrumental variables specification with corrected standard errors. Below the standard errors the 95% Anderson and Rubin (1949) confidence set and the 95% tF confidence interval following Lee et al. (2022) are reported. Each regression includes year and county fixed effects and control variables. Control variables include % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. Observations are weighted by the number of births in a county and year. Each coefficient where infant mortality rate is the dependent variable represents the percent reduction in infant mortality rates due to receiving a capital appropriation or payment from The Duke Endowment. A coefficient here of -10 would mean that mortality declines by 10%. For the first stage regressions, the coefficient is the relationship between \$ 1 billion in cumulative returns and either cumulative appropriations or payments. Coefficients reported in the table are transformed in the following way: $100 \times (\exp(\beta) - 1)$. Standard errors are estimated using the delta method and are clustered at the county-level.

Figure J1. Infant mortality by year: Ever-treated NC vs. other Duke-ineligible Southern counties

(a) Black infant mortality

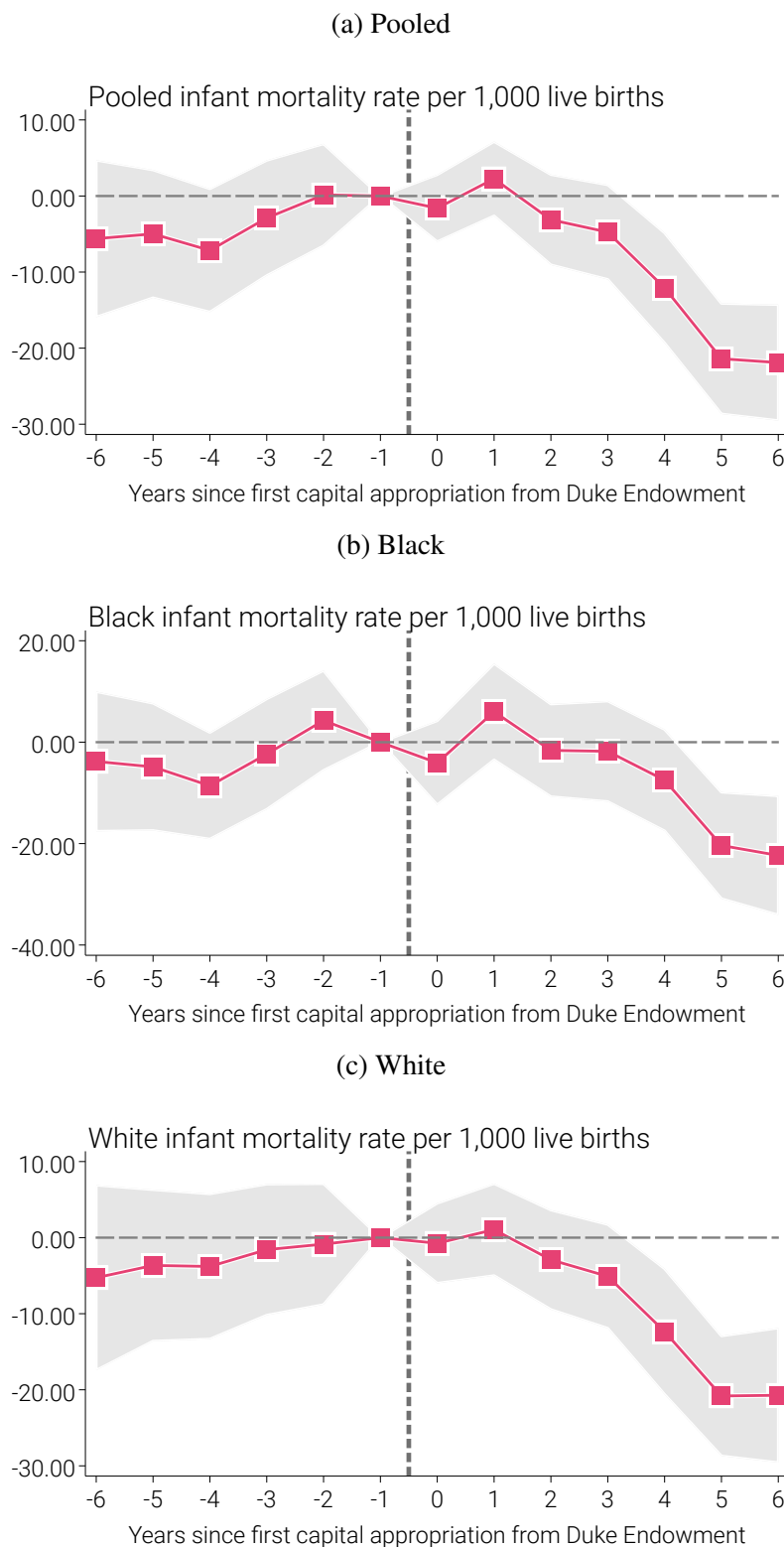


(b) White infant mortality



Notes: This figure compares county-year infant mortality by race for counties in North Carolina that are “Ever treated” during our sample (pink solid line) to infant mortality rates from from other Southern counties ineligible for Duke funding (thick black line). These data come from Fishback et al. (2007). Non-Carolina counties are mechanically ineligible as Duke funding was only available for communities in North and South Carolina.

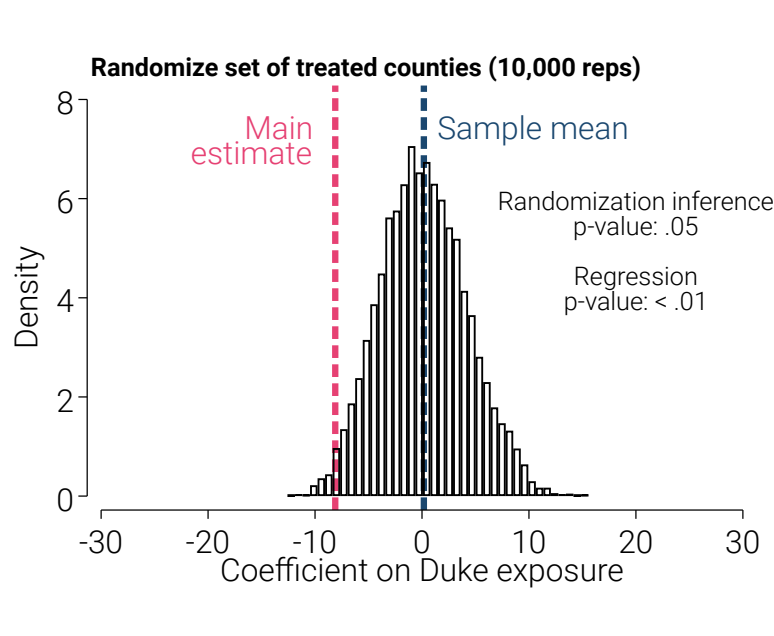
Figure J2. Event studies: Replacing untreated North Carolina counties with other ineligible Southern counties



Notes: Each panel is an event study that corresponds to the regression from column 2 of Table F3. Panel A pools Black and White infants together, panel B examines the Black infant mortality rate, and panel C examines the White infant mortality rate. Each regression drops all North Carolina counties that were not treated by Duke (i.e., did not receive a capital appropriation) before 1942. Untreated control counties are Southern counties outside the Carolinas that had cities between 1930 to 1940. Data come from Bailey et al. (2015) and are unbalanced before 1930 and after 1940. Thus, to ensure balance across years, the data in this figure are restricted to the 1930 to 1940 birth years. The weights are the number of births in a county and year. Standard errors are estimated using the delta method and are clustered at the county-level.

K Randomization of the treatment

Figure K1. Randomization of Duke support for infant mortality rate



Notes: This figure presents a histogram of coefficient estimates from 10,000 iterations of modifying the regression specification in column 5 of Table 3. The dependent variable is the pooled infant mortality rate. In each iteration, we randomly select 48 counties out of the 100 counties in North Carolina and consider them to be treated by Duke funding. In each case, we preserve the true treatment path, i.e., the years when treatment turns on. The number of counties in each treatment timing group also does not change. The dashed blue line indicates the sample mean of the 10,000 estimates. The dashed pink line indicates the estimate from column 5 of Table 3. Each coefficient represents the percent reduction in infant mortality due to receiving a capital appropriation from The Duke Endowment. A coefficient of -10 would mean that infant mortality declines by 10%. Coefficients reported in the figure are transformed in the following way: $100 \times (\exp(\beta) - 1)$. All regressions include county and year-of-birth fixed effects but no other controls. The weights are the number of births in a county and year.

L Adding additional years of data to the end of the sample

Table L1. First-stage hospital analysis extended to 1950

	$Y_{ct}^R = \text{Beds or Hospitals}$			$Y_{ct}^R = \text{Beds or Hospitals per 1,000 births}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Beds</i>						
Total	37.40*** (9.35)	44.76*** (13.55)	28.88* (15.56)	28.04*** (5.55)	25.33*** (5.96)	24.51*** (5.61)
Non-profit/church/public	41.64*** (9.68)	48.81*** (14.47)	32.76** (16.09)	33.55*** (5.72)	28.97*** (6.47)	28.06*** (6.14)
Proprietary	-4.66** (2.16)	-5.38* (2.89)	-5.52* (2.84)	-5.26* (2.82)	-4.34* (2.35)	-4.44* (2.30)
<i>B. Hospitals</i>						
Total	0.21** (0.09)	0.12 (0.11)	0.14 (0.10)	0.30 (0.19)	0.13 (0.12)	0.22* (0.12)
Non-profit/church/public	0.35*** (0.09)	0.25* (0.13)	0.28** (0.12)	0.50*** (0.16)	0.25** (0.12)	0.35*** (0.12)
Proprietary	-0.13** (0.06)	-0.14* (0.07)	-0.16** (0.07)	-0.17 (0.11)	-0.11 (0.07)	-0.13* (0.07)
Observations	2,900	2,900	2,900	2,900	2,900	2,900
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	No	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	No	No	Yes

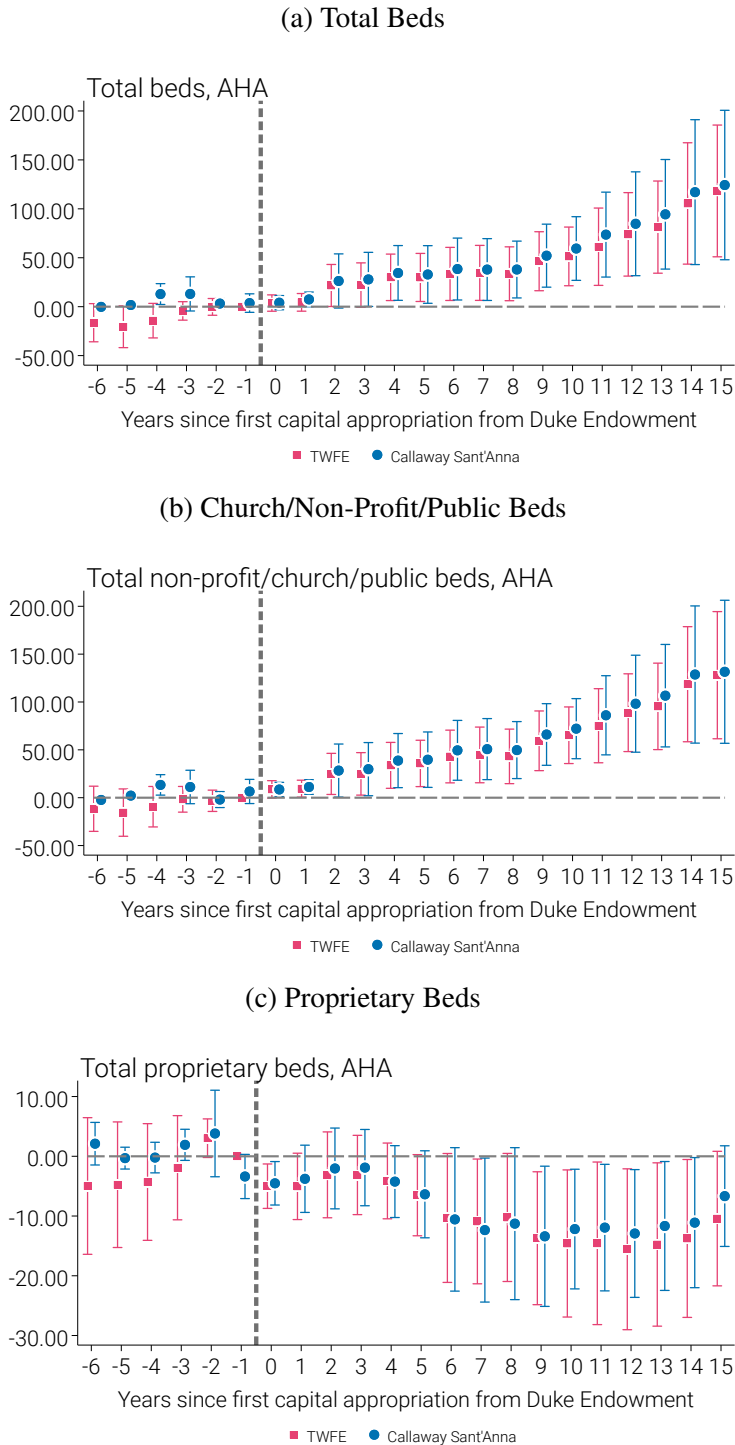
Notes: See the notes to Table 1 for a description of the specifications reported in the table. This table differs only in the extension of the sample period from 1922-1942 to 1922-1950, which in turn implies that some counties that are never treated in the main sample are considered treated if they received Duke funding between 1943 and 1950.

Table L2. Effects on infant mortality: Non-Carolina controls extended to 1962

	Fishback data		Bailey data			
	1922-1940		1922-1962			
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
Percent effect from Duke (=1)	-9.87*** (2.10)	-8.59*** (2.20)	-10.54*** (2.08)	-10.64*** (2.10)	-17.95*** (2.49)	-18.38*** (2.50)
Observations	3,801	2,813	4,001	3,564	9,398	8,456
County of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Exclude untreated NC	No	Yes	No	Yes	No	Yes

Notes: The specifications reported in this table extend the analysis with non-Carolina control counties by including data up to 1962. All specifications use our main infant mortality measure for North Carolina counties for our main sample years (1922 to 1942). For the non-Carolina controls, columns 1 and 2 use infant mortality data from Fishback et al. (2007) which are only available up to 1940. Column 2 drops counties in North Carolina that are untreated up to and including 1940. Columns 3 and 4 are equivalent to columns 1 and 2 but use data from Bailey et al. (2015) for the non-Carolina controls, which cover a similar of counties. Columns 5 and 6 extend the sample to 1962 – the full extend of data from Bailey et al. (2015)

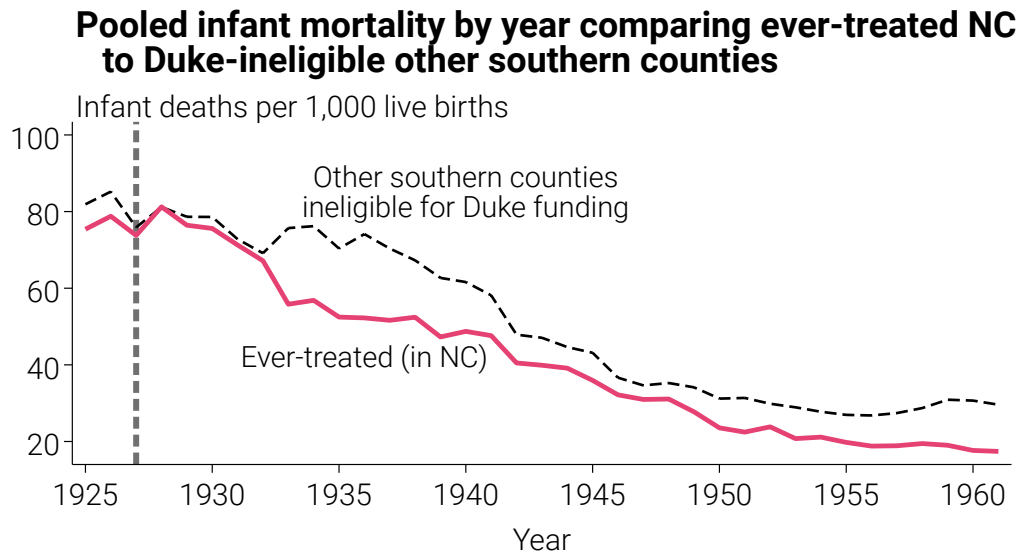
Figure L1. First-stage analysis of hospital beds extended to 1950: Event studies



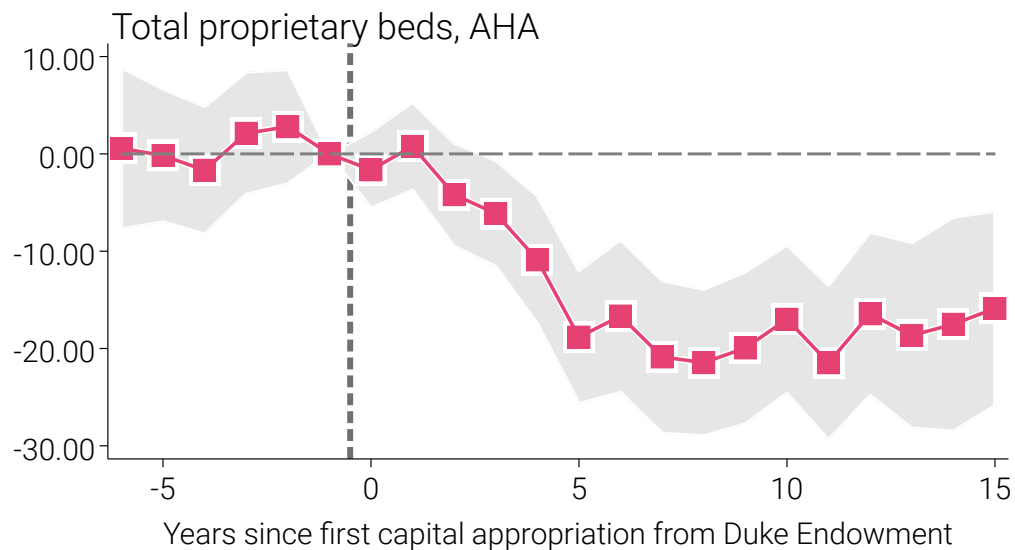
Notes: The event studies correspond to the specifications reported in column 2 of panel A in Online Appendix Table L1.

Figure L2. Analysis with non-Carolina controls extended to 1962

(a) Infant mortality rate across time



(b) Event study



Notes: Panel A uses our main infant mortality measure constructed from North Carolina death certificates for North Carolina counties between 1925 and 1942. We use infant mortality data from Bailey et al. (2015) to extend the series until 1962 and add other Southern non-Carolina counties ineligible for Duke funding as controls. Panel B presents an event study that corresponds to column 5 of Online Appendix Table L2.

M Propensity score matching: Alternative control group and falsification test

In this section we conduct two exercises. First, we perform a falsification test using only untreated non-Carolina counties. Second, we re-estimate our treatment effects using a selected set of non-Carolina control counties that look the most like the treated North Carolina counties. Both exercises are built on the same propensity score match. Collectively, these tests help dissuade concerns that underlying trends in places that appear to be similar to treated North Carolina counties (e.g., places with hospitals) were not simply on differential trends with respect to infant mortality than other places.

We use a probit regression at the county level to identify the non-Carolina counties that look most like the ever-treated North Carolina counties. The outcome variable takes the value 1 if a county is ever treated during our sample (i.e., is a county in North Carolina that received a capital appropriation from the Duke Endowment between 1927 and 1942). The explanatory variables used for matching are defined at the county level and are held fixed to values from the time period before the Duke Endowment began appropriating funds. Included variables are % illiterate, % black, % other race, % urban, retail sales per capita, and total population from the 1920 census; county health department presence in 1925; the number of proprietary hospital beds, non-proprietary hospital beds, and the number of hospitals in 1927; and proxies for the 1920 infant mortality rate and 1920-1924 childhood (1-5 year old) mortality rate by race.

The infant and childhood mortality rates used in our probit regression are constructed following Feigenbaum et al. (2023) and represent consistent measures of infant and childhood mortality in all Southern counties regardless of whether or not a state collected and reported mortality data. These measures proxy for mortality by finding the number of “missing” infants and children from one census wave to the next. We create these proxies using the publicly-available complete-count US Census data from IPUMS in 1920 and 1930 which are linked across time at the person and household levels (Ruggles et al. 2023). We limit the sample to households that were surveyed in both census years, who resided in the South, and that were enumerated as a “married-couple family household.”

For infant mortality, we then create an indicator that takes the value one if a child under the age of one is present in the surveyed household in the 1920 census, but is not present in the same surveyed household in the 1930 census. We do an analogous exercise for children aged one to five in the 1920 census. We

collapse these data to the county-level and create the share of missing children by age group and by race. Feigenbaum et al. (2023) point out that for the counties with data on infant and childhood mortality, the 1-5 year old proxies have greater correlation with actual data than the infant mortality proxies. Thus, while we think these values help match counties on underlying health status, we do not think it is best to use these constructed proxies instead of actual infant mortality data in our main analyses.

We obtain predicted ever-treated probabilities using estimates from the cross-sectional probit regression and retain only Southern non-Carolina counties. We then construct three different comparison sets using these predicted probabilities. The first set keeps only the top 10% and bottom 10% of counties based on predicted probabilities. The second and third sets do the same, but with break points at 25% and 50%.

Figure M1 uses data from Fishback et al. (2007) to show for each set how the average infant mortality rate in non-Carolina counties that look the most like treated NC counties compares to the average infant mortality rate in counties that look the least like treated NC counties. The first column compares the top 10% to the bottom 10%, the second and third columns do the same using the top and bottom 25% and 50%. The first row presents the comparison for pooled infant mortality while the second and third rows present the comparison for Black and White infant mortality rates, respectively. Broadly, there are no apparent visual differences in the average infant mortality rate between Southern non-Carolina counties that look the most like treated North Carolina counties and non-Carolina counties that look the least like treated North Carolina counties.

We formalize this placebo analysis using regressions in Table M1. We consider two types of pseudo-treatment for Southern counties that look the most like treated North Carolina counties. The first pseudo-treatment begins treatment for each “top X%” county in a random year. Once treatment begins, the county remains treated. These pseudo-treatment effect estimates are reported in the odd columns. The second pseudo-treatment randomly allocates a treatment year that corresponds to the actual roll-out of treatment years in North Carolina. We do not find any differences in infant mortality in either analysis.

The second exercise we show in this section uses the propensity scores to restrict the set of control counties to be those non-Carolina counties that look the most like the treated North Carolina counties. These results are reported in Tables M2, M3, and M4. In Table M2, the only non-Carolina control counties that

are included are those whose propensity scores are in the top ten percent. That is, those counties that look the most like treated North Carolina counties. Tables M3 and M4 show analogous results for those counties whose propensity scores are in the top 25% and 50%, respectively. Results are essentially unchanged from our other analyses that use non-Carolina control variables.

Table M1. Compare non-Carolina counties that look the most like treated NC counties to those that look the least like like treated NC counties

	Top 10% vs bottom 10%		Top 25% vs bottom 25%		Top 50% vs bottom 50%	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
Percent effect from Duke (=1)	1.88 (2.08)	2.12 (2.64)	2.16 (2.05)	-0.03 (2.08)	2.92 (2.33)	2.15 (1.91)
Observations	1,098	1,098	1,572	1,572	1,922	1,922
<i>B. Black</i>						
Percent effect from Duke (=1)	6.82* (3.68)	3.92 (4.35)	-3.22 (3.40)	1.59 (2.42)	1.97 (3.78)	-1.32 (2.55)
Observations	959	959	1,352	1,352	1,682	1,682
<i>C. White</i>						
Percent effect from Duke (=1)	0.56 (2.04)	3.52 (3.35)	2.42 (2.61)	-3.25 (2.76)	1.92 (2.43)	3.36 (2.35)
Observations	973	973	1,368	1,368	1,703	1,703
County of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Exclude NC	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo-treatment begins in random year	Yes	No	Yes	No	Yes	No
Pseudo-treatment begins in years that match actual roll-out	No	Yes	No	Yes	No	Yes

Figure M1. Compare infant mortality rates in non-Carolina counties that look the most like treated NC counties to infant mortality rates in non-Carolina counties that look the least like treated NC counties

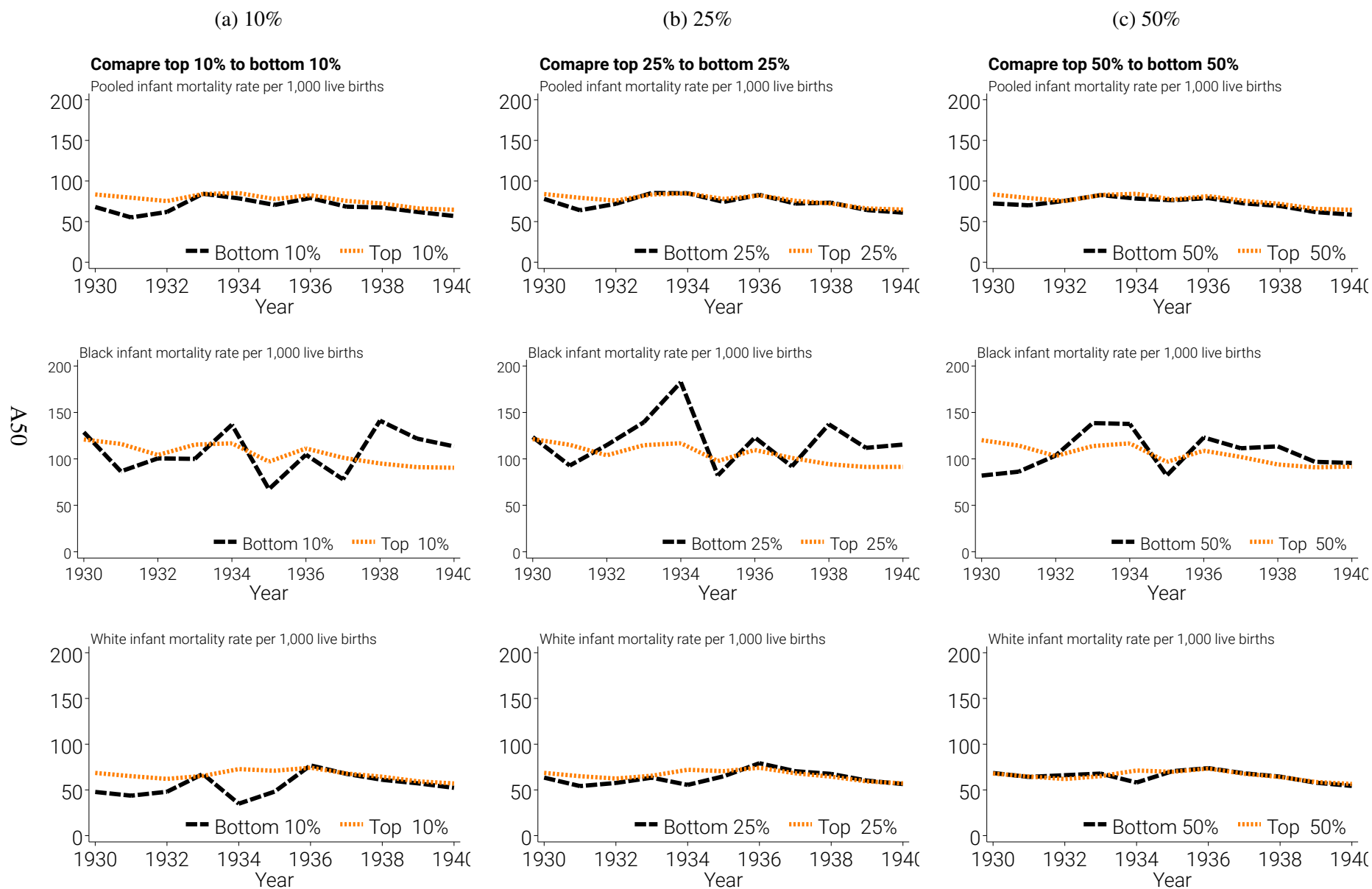


Table M2. Propensity score match - top 10%

	$Y_{ct}^R = \text{Infant mortality rate}$			$Y_{ct}^R = \text{Long-run deaths}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
Percent effect from Duke (=1)	-10.57*** (2.16)	-10.02*** (2.42)	-10.05*** (2.41)	-12.45*** (3.09)	-13.16*** (3.29)	-13.16*** (3.29)
Observations	2,936	1,948	1,940	12,861	8,181	8,181
<i>B. Black</i>						
Percent effect from Duke (=1)	-12.33*** (2.70)	-9.71*** (2.69)	-9.72*** (2.69)	-10.57** (4.57)	-14.42*** (4.75)	-14.42*** (4.75)
Observations	2,709	1,778	1,772	12,269	7,959	7,952
<i>C. White</i>						
Percent effect from Duke (=1)	-10.46*** (2.63)	-11.07*** (2.88)	-11.14*** (2.88)	-11.84*** (3.44)	-12.03*** (3.49)	-12.01*** (3.49)
Observations	2,714	1,783	1,776	12,923	8,523	8,460
P-value for difference by race	0.55	0.67	0.66	0.82	0.65	0.65
County of birth FE	Yes	Yes	Yes	No	No	No
County of birth X Age FE	No	No	No	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	No	No	No
Year of birth X Age FE	No	No	No	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Exclude untreated NC	No	Yes	Yes	No	Yes	Yes
Exclude without non-profit hosp.	No	No	Yes	No	No	Yes
PSM percentile cutoff for including other southern counties	10%	10%	10%	10%	10%	10%

Table M3. Propensity score match - top 25%

	$Y_{ct}^R = \text{Infant mortality rate}$			$Y_{ct}^R = \text{Long-run deaths}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
Percent effect from Duke (=1)	-10.90*** (2.11)	-10.64*** (2.26)	-10.66*** (2.26)	-12.26*** (3.07)	-12.87*** (3.26)	-12.87*** (3.26)
Observations	3,384	2,396	2,388	14,796	10,116	10,116
<i>B. Black</i>						
Percent effect from Duke (=1)	-11.96*** (2.68)	-9.28*** (2.62)	-9.29*** (2.62)	-10.32** (4.49)	-13.62*** (4.79)	-13.62*** (4.79)
Observations	3,077	2,146	2,140	14,168	9,858	9,851
<i>C. White</i>						
Percent effect from Duke (=1)	-11.10*** (2.58)	-11.95*** (2.77)	-12.00*** (2.77)	-11.73*** (3.39)	-11.82*** (3.43)	-11.81*** (3.44)
Observations	3,084	2,153	2,146	14,903	10,503	10,440
P-value for difference by race	0.78	0.39	0.39	0.79	0.74	0.73
County of birth FE	Yes	Yes	Yes	No	No	No
County of birth X Age FE	No	No	No	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	No	No	No
Year of birth X Age FE	No	No	No	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Exclude untreated NC	No	Yes	Yes	No	Yes	Yes
Exclude without non-profit hosp.	No	No	Yes	No	No	Yes
PSM percentile cutoff for including other southern counties	25%	25%	25%	25%	25%	25%

Table M4. Propensity score match - top 50%

	$Y_{ct}^R = \text{Infant mortality rate}$			$Y_{ct}^R = \text{Long-run deaths}$		
	(1)	(2)	(3)	(4)	(5)	(6)
<i>A. Pooled</i>						
Percent effect from Duke (=1)	-10.27*** (2.07)	-9.61*** (2.20)	-9.62*** (2.20)	-12.42*** (3.00)	-13.07*** (3.13)	-13.07*** (3.13)
Observations	3,646	2,658	2,650	16,065	11,385	11,385
<i>B. Black</i>						
Percent effect from Duke (=1)	-11.46*** (2.78)	-8.35*** (2.69)	-8.38*** (2.69)	-9.89** (4.43)	-12.73*** (4.86)	-12.73*** (4.86)
Observations	3,327	2,396	2,390	15,463	11,153	11,146
<i>C. White</i>						
Percent effect from Duke (=1)	-10.60*** (2.55)	-11.10*** (2.70)	-11.14*** (2.71)	-11.97*** (3.37)	-12.16*** (3.38)	-12.15*** (3.39)
Observations	3,337	2,406	2,399	16,262	11,862	11,799
P-value for difference by race	0.78	0.38	0.38	0.70	0.92	0.92
County of birth FE	Yes	Yes	Yes	No	No	No
County of birth X Age FE	No	No	No	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	No	No	No
Year of birth X Age FE	No	No	No	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Exclude untreated NC	No	Yes	Yes	No	Yes	Yes
Exclude without non-profit hosp.	No	No	Yes	No	No	Yes
PSM percentile cutoff for including other southern counties	50%	50%	50%	50%	50%	50%

N Summary statistics

Table N1. Summary statistics: Short-run and long-run mortality

	Mean	S.D.	Min.	Max.	N
<i>Short-run treatment and mortality</i>					
County-year treatment status (=1)	0.28	0.45	0.00	1.00	2100
Appropriations, millions	0.03	0.18	0.00	4.73	2100
Payments, millions	0.02	0.16	0.00	4.52	2100
<i>Pooled</i>					
Infant deaths	49.97	47.08	1.00	331.00	2100
Infant deaths per 1,000 births	60.01	24.56	6.99	218.39	2100
Births	801.53	575.67	72.00	3843.00	2100
<i>Black</i>					
Infant deaths	20.05	22.83	0.00	150.00	2100
Infant deaths per 1,000 births	100.45	127.10	0.00	1000.00	2082
Births	246.61	233.18	0.00	1123.00	2100
<i>White</i>					
Infant deaths	29.91	29.56	0.00	243.00	2100
Infant deaths per 1,000 births	53.06	22.90	0.00	289.47	2100
Births	554.95	441.36	30.00	2901.00	2100
<i>Long-run treatment and mortality</i>					
County-birth-year treatment status (=1)	0.26	0.44	0.00	1.00	9000
<i>Pooled</i>					
Deaths in follow-up year	6.19	5.55	0.00	46.00	9000
<i>Black</i>					
Deaths in follow-up year	2.07	2.53	0.00	20.00	9000
<i>White</i>					
Deaths in follow-up year	4.13	4.09	0.00	35.00	9000
<i>Sulfa-specification</i>					
Average pooled pneumonia mortality rate, 1922 to 1926	93.76	22.79	38.77	157.84	100
<i>Controls</i>					
% illiterate	7.44	3.00	0.00	18.54	2100
% population Black	27.67	18.09	0.02	65.28	2100
% population other race	0.46	2.23	0.00	22.54	2100
% population urban	14.97	18.53	0.00	80.66	2100
Retail sales per capita	159.26	83.06	21.53	584.46	2100
County health department present (=1)	0.36	0.48	0.00	1.00	2100

Table N2. Summary statistics: First stage outcomes for hospitals and doctors

	Mean	S.D.	Min.	Max.	N
<i>Short-run hospital data</i>					
<i>Total</i>					
Beds	56.22	108.31	0.00	1012.89	2100
Beds per 1,000 births	49.40	64.12	0.00	478.20	2100
Hospitals	1.11	1.34	0.00	7.00	2100
Hospitals per 1,000 births	1.29	1.49	0.00	9.80	2100
<i>Non-profit/Public/Church</i>					
Beds	48.42	104.90	0.00	1012.89	2100
Beds per 1,000 births	40.18	60.96	0.00	478.20	2100
Hospitals	0.85	1.18	0.00	7.00	2100
Hospitals per 1,000 births	0.91	1.24	0.00	9.80	2100
<i>Proprietary</i>					
Beds	6.97	17.67	0.00	140.00	2100
Beds per 1,000 births	7.97	23.22	0.00	351.76	2100
Hospitals	0.24	0.49	0.00	2.00	2100
Hospitals per 1,000 births	0.33	0.89	0.00	7.19	2100
<i>Short-run doctor data</i>					
<i>Pooled</i>					
Doctors	23.70	31.36	1.00	232.00	1100
Doctors per 1,000 births	25.59	13.42	2.93	114.64	1100
High-quality doctors	6.74	15.04	0.00	190.00	1100
High-quality doctors per 1,000 births	6.45	8.64	0.00	92.57	1100
Low-quality doctors	16.54	20.56	0.00	144.00	1100
Low-quality doctors per 1,000 births	18.51	9.37	0.00	80.54	1100
<i>Black</i>					
Doctors	1.44	2.75	0.00	16.00	1100
Doctors per 1,000 births	3.92	7.18	0.00	142.86	1095
High-quality doctors	0.39	1.21	0.00	10.00	1100
High-quality doctors per 1,000 births	0.94	2.52	0.00	16.46	1095
Low-quality doctors	1.05	1.95	0.00	13.00	1100
Low-quality doctors per 1,000 births	2.98	6.26	0.00	142.86	1095
<i>White</i>					
Doctors	22.25	29.06	1.00	219.00	1100
Doctors per 1,000 births	38.33	22.72	3.20	216.61	1100
High-quality doctors	6.36	14.02	0.00	180.00	1100
High-quality doctors per 1,000 births	9.98	14.61	0.00	158.84	1100
Low-quality doctors	15.49	19.08	0.00	139.00	1100
Low-quality doctors per 1,000 births	27.60	15.22	0.00	92.59	1100

Appendix References

- Anderson, T. W. and Herman Rubin (1949) “Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations,” *The Annals of Mathematical Statistics*, Vol. 20, No. 1, pp. 46 – 63.
- Arias, E. (2014) “United States life tables, 2010,” *National vital statistics reports*, Vol. 63, No. 7.
- Arias, E., B. Tejada-Vera, F. Ahmad, and K.D. Kochanek (2021) “Provisional life expectancy estimates for 2020,” *Vital Statistics Rapid Release*, No. 15, pp. 50–61.
- Bailey, Martha, Karen Clay, Price Fishback, Michael Haines, Shawn Kantor, Edson Severnini, and Anna Wentz (2015) *U.S. County-Level Natality and Mortality Data, 1915-2007*: Inter-university Consortium for Political and Social Research [distributor].
- Bhalotra, Sonia, Martin Karlsson, and Therese Nilsson (2017) “Infant Health and Longevity: Evidence from a Historical Intervention in Sweden,” *Journal of the European Economic Association*, Vol. 15, No. 5, pp. 1101–1157.
- Callaway, Brantly and Pedro H.C. Sant’Anna (2021) “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, Vol. 225, No. 2, pp. 200–230.
- Feigenbaum, James, Lauren Hoehn-Velasco, and Sophie Li (2023) “Germ Theory at Home: The Role of Private Action in Reducing Child Mortality during the Epidemiological Transition,” Working Paper.
- Fishback, Price V., Michael R. Haines, and Shawn Kantor (2007) “Births, Deaths, and New Deal Relief during the Great Depression,” *The Review of Economics and Statistics*, Vol. 89, No. 1, pp. 1–14.
- Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020) “Bartik Instruments: What, When, Why, and How?” *American Economic Review*, Vol. 110, No. 8, pp. 2586–2624.
- Goodman-Bacon, Andrew (2021) “The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes,” *American Economic Review*, Vol. 111, No. 8, pp. 2550–2593.
- Haines, Michael (2008) “Fertility and Mortality in the United States,” in Robert Whaples ed. *EH.Net Encyclopedia*: Economic History Association.
- Jayachandran, Seema, Adriana Lleras-Muney, and Kimberly Smith (2010) “Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs,” *American Economic Journal: Applied Economics*, Vol. 2, No. 2, pp. 118–146.
- Lazuka, Volha (2020) “Infant Health and Later-life Labor Market Outcomes: Evidence from the Introduction of Sulfa Antibiotics in Sweden,” *Journal of Human Resources*, Vol. 55, No. 2, pp. 660–698.
- Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack Porter (2022) “Valid t-Ratio Inference for IV,” *American Economic Review*, Vol. 112, No. 10, pp. 3260–90.
- Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt (2019) “Poorly Measured Confounders Are More Useful on the Left than on the Right,” *Journal of Business & Economic Statistics*, Vol. 37, No. 2, pp. 205–216.
- Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Anne Chen, Grace Cooper, Stephanie Richards, and Megan Schouweiler (2023) *IPUMS USA: Version 14.0 [dataset]*, Minneapolis, MN: IPUMS.
- Wooldridge, Jeffrey M. (2021) “Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators,” Working Paper.