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

We 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, Miami University, The Ohio State University Departments of Economics and Human Sciences, and the University of Copenhagen for helpful comments and feedback. We are grateful to Patrick Carlin 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.

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
JEL No. I14,J13,N32

ABSTRACT

The past century witnessed a dramatic improvement in public health, the rise of modern medicine, and the transformation of the hospital from a fringe institution to one essential to the practice of medicine. In this paper, we explore how access to the hospital and modern medicine affects mortality. We do so by leveraging a combination of novel data and a unique quasi-experiment: a large-scale hospital modernization program introduced by The Duke Endowment 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 access to a Duke-supported hospital reduced infant mortality by 10%, saving one life for every $20,000 (2017 dollars) spent. Effects were larger for Black infants (16%) than for White infants (7%), implying a reduction in the Black-White infant mortality gap by one-third. We show that the effect of Duke support persisted into later life with a 9% reduction in mortality between the ages of 56 and 65. We further provide evidence on the mechanisms that enabled these effects, finding that Endowment-supported hospitals attracted higher-quality physicians and were better able to take advantage of new medical innovations.

Alex Hollingsworth
O'Neill School of Public and
Environmental Affairs
Indiana University
1315 E. Tenth St.
Bloomington, IN 47405
and NBER
hollinal@indiana.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; Sutch 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 of the modern hospital, transforming the institution "from places of dreaded impurity and exiled human wreckage into awesome citadels of science and bureaucratic order" (Starr 2017). At the same time, health has improved dramatically, 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; Acemoglu and Johnson 2007). Despite this viewpoint, little is actually 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.

In this paper we ask the following questions relevant to understanding both the role of medical care in the health transition and contemporary healthcare decisions: Did access to modern medical knowledge and technology through the transformation of the hospital contribute to declining mortality? Given the many racial inequities in society and in health, were these effects similar or different across race? What were the mechanisms that enabled modern medicine and hospitals to improve health? In particular, did better infrastructure attract higher-quality physicians and allow for more productive use of new medical technologies? And finally, were these effects limited to gains in infancy or did they persist into later life?

We answer these questions by leveraging 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 1926, The Duke Endowment expanded access to modern medicine and hospitals across the Carolinas, 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 younger and higher-quality physicians to the Carolinas (The Duke Endowment 1934; Durdin 1998, p. 44). By the end of 1942, the Endowment had appropriated over $126 million (2017 dollars) for hospital expansion and modernization, an investment that significantly improved access to medical facilities and the quality of the medical infrastructure in North Carolina, thereby propelling the practice of medicine in supported communities towards the modern era.[1]

Our main empirical approach exploits quasi-random variation in Duke support, comparing outcomes for cohorts born in counties that had received a funding commitment from the Endowment around the time of their birth to outcomes for those born in counties without such a commitment. 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. Before an appropriation turned into a payment, the Endowment typically required hospitals to undertake additional steps, such as changing hospital accounting or bookkeeping practices; ensuring that surgeries were conducted by licensed physicians; allowing Endowment employees to visit the hospital for a de facto inspection; 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 for the possibility that hospitals with failed appropriations (i.e., those that never received a payment) could still have made changes that improved health.

We find that the efforts by Duke to modernize hospitals improved the health of the surrounding community and reduced infant mortality by 10%. Treatment effect estimates for Black infants (16%) are twice as large as for White infants (7%), implying that the Endowment narrowed the infant mortality gap by one-third. Those exposed to Duke support around the time of birth also

[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.

enjoyed lasting health effects, as long-run mortality rates (between the ages 56 and 65) declined by 9%. We also find similar evidence when considering the intensive margin (i.e., how much money was appropriated or paid). We explore various channels through which these effects could have emerged, and show that Duke support increased the stock of high-quality physicians in a county by 11% and reduced the share of poorly trained physicians by 9%. This change shifted the average graduation year of doctors in Endowment-supported counties by 12 years, advancing the average state of medical knowledge in these locations. 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 not sensitive to the use of a variety of recent techniques designed to address concerns about treatment effect heterogeneity and staggered treatment adoption in a difference-in-differences setting. We also 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. 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. We do so by considering balance tests, specifications with an alternate set of non-Carolina Southern control counties that were ineligible for Duke funding, and an instrumental variables framework that exploits exogenous variation in Duke eligibility and returns of Endowment assets. Moreover, statistical significance is not affected by using a randomization inference approach that randomizes the set of treated counties, holding fixed the number of treated counties and the treatment path across time. These tests mitigate concerns that selection or spurious trends are driving our treatment effect estimates.

The program was cost effective. Considering only our infant mortality results, one life was saved for every $20,000 (2017 dollars) paid to hospitals, implying that the Endowment's inter-

vention passes any cost-benefit analysis using a reasonable estimate for a Value of Statistical Life (VSL). The intervention was also particularly effective in reducing Black infant mortality. Duke support caused a decrease in the mortality rate for Black infants that was roughly three times larger than the decrease for White infants (a reduction of 12 vs. 4 deaths per 1,000 births), narrowing the Black-White infant mortality gap by one-third. This new evidence that access to high-quality medical care and infrastructure disproportionately reduced Black mortality rates adds to previous work that has examined the Black-White infant mortality gap, some of which finds no relationship between increased healthcare access and the gap (Miller 2001; Collins and Thomasson 2004; Almond et al. 2006; Elder et al. 2011, 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., Cutler and Miller 2005; Ferrie and Troesken 2008; Clay et al. 2014, 2020; Anderson et al. 2019, 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; Acemoglu and Johnson 2007). In contrast, we provide one of the first glimpses into the role that modern medical care played in the mortality transition, finding large short-run effects and lasting mortality effects in later life.

Second, we build on the broader literature examining the short- and long-run consequences of prenatal and infant health.[3] We also complement studies examining the determinants of mortal-

[2]Extending work by Almond et al. (2006), Anderson et al. (2021a) show that the 1966 federal campaign to desegregate Southern hospitals had no observable effect on the Black-White infant mortality gap. A hypothesized reason for the null result is that desegregation was "cosmetic" and did not actually extend access to high-quality facilities for Black residents. While North Carolina is not included in their analysis, these results are not inconsistent with our findings. The 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, p.185).

[3]Prior work has examined the effects of a number of interventions including the provision or expansion of health or social insurance (Currie and Gruber 1996a,b; Cohodes et al. 2016; Goodman-Bacon 2018, 2021b; Lührmann and Wilson 2018; Bauernschuster et al. 2019; Miller and Wherry 2019; Brown et al. 2020); access to maternity wards

ity and longevity, in which a common theme is that early life circumstances matter for later-life health (Almond and Currie 2011; 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. A unique feature of our paper is that we are not simply studying the effects of a well-known intervention applied to a different vulnerable population (e.g., how insurance expansion affects a newly enrolled group). We are instead examining an ambitious experiment that sought to modernize medical care at a critical time in the evolution of the field.

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, with the first director of the Endowment's Hospital Section, Dr. Watson S. Rankin, stating 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 "the main remedy for the shortage of doctors" (Durdin 1998, p. 31, 54). Our findings support this view. Related studies have explored how forms of occupational licensing in this time period increased access to higher-quality medical personnel and how this in turn affected health (Law and Kim 2005; Anderson et al. 2020; Moehling et al. 2020).[4] Our setting differs in that the Endowment did not affect legislation or directly contract with physicians; instead, it was an inducement, making counties with modern, Duke-supported facilities more attractive to physicians.

An additional complementarity that we examine is the relationship between technological in-

(Fischer et al. 2021; Lazuka 2021); county health departments (Hoehn-Velasco 2021); home visits from nurses (Olds et al. 1999; Hjort et al. 2017; Bütikofer et al. 2019); and reductions in air pollution (Chay and Greenstone 2003; Greenstone and Hanna 2014; Clay et al. 2016; Knittel et al. 2016).

[4]Anderson et al. (2020) show how midwifery licensing reforms reduced maternal and infant mortality, finding a smaller average effect for infant relative to maternal mortality. When split by race, they find a statistically insignificant relationship with White infant mortality and a relatively larger, significant relationship with non-White infant mortality. Law and Kim (2005) document that physician licensing changed physician quality, but they find an insignificant relationship between physician quality and mortality from physician-treatable diseases. Moehling et al. (2020) find that the Sheppard-Towner Act increased one-on-one contacts with licensed physicians and nurses – leading to reductions in infant mortality, with particularly large effects for Black infants.

novation and improvements in physical and human capital. For example, Soares (2007) argues that the diffusion of knowledge is necessary for the effective spread and use of medical technologies. In related work, Alsan and Goldin (2019) examine how the provision of clean water and innovations in sewage disposal affected child mortality in late-nineteenth century Massachusetts. However, most other studies in this area only consider interactions in the context of human capital, and are inspired by the idea of dynamic complementarities (Cunha and Heckman 2007).[5] Our focus differs in that we are interested in complementarities between hospital infrastructure and medical innovation. Our investigation into this interaction is enabled by the timing of the discovery and use of sulfa drugs, the first chemical substances systematically used to prevent and treat bacterial infections. Prior work has documented that the 1937 introduction of sulfa drugs led to large declines in infectious disease mortality (Thomasson and Treber 2008; Jayachandran et al. 2010; Bhalotra and Venkataramani 2015; Lazuka 2020). We contribute to this literature by showing that the higher-quality facilities and workforce promoted by Duke enabled communities to take greater advantage of sulfa drugs. Specifically, we find that for communities with higher baseline infectious disease mortality, the reductions in infant mortality caused by residing in a Duke-supported community roughly *doubled* after sulfa drugs were developed.

Fourth, we contribute to the literature on management style, standardization, and firm performance (Bloom and Van Reenen 2007; Tsai et al. 2015; McConnell et al. 2016; Bloom et al. 2020; Figinski and Troland 2020; Scur et al. 2021; Gray-Lobe et al. 2022). A distinct aspect of the intervention we study is that treatment is not limited to a monetary component. Duke support was a multifaceted bundle of treatments coupled with ample oversight from the Endowment. In addition to providing funding for facilities and equipment, core management practices were altered and hospital accounting procedures were standardized.[6] As a result, the Endowment "exerted considerable

[5]Examples of recent studies that investigate interactions between multiple inputs and estimate the human capital production function include Gunnsteinsson et al. (2019) and Bütikofer et al. (2021) who find that negative prenatal health shocks can be offset in the early postnatal period; Rossin-Slater and Wüst (2020) who find that programs involving home visits by nurses have a greater effect in the presence of high-quality preschools; or Adhvaryu et al. (2021) and Duque et al. (2021) who examine the interactions between financial shocks and poverty-alleviation programs.

[6]Summarizing a letter written by Graham Davis, the assistant director of The Duke Endowment in the time period we study, 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

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 and it is possible that the Duke intervention was so successful *because* these many changes served to reinforce one another. Our findings support the idea that bundling healthcare access with treatments such as changing management style and increasing oversight can have complementary effects (Bloom et al. 2020; Figinski and Troland 2020).

Finally, our work connects with the literature examining the effects of charitable contributions on public health and welfare. For instance, Bleakley (2007) examines the long-term effects of the hookworm eradication campaign sponsored by the Rockefeller Sanitary Commission across the American South in the early-twentieth century and finds increases in school enrollment, attendance, and literacy.[7] We add to this literature by showing large, lasting, and cost-effective direct health effects from private charitable contributions.[8] For many years, The Duke Endowment was the third largest charity in the US and it remains one of the largest charitable foundations in the world (Durdin 1998, p. 1). Despite its size and success, to our knowledge, no economists and virtually no other academics (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 program was so pronounced that the Endowment's efforts to support hospitals "served as both inspiration and model" for the Hill-Burton Act of 1946 (Durdin 1998, p. 33), a large program designed to stimulate hospital construction and access across the US (Chung et al. 2017). We provide some of the first evidence demonstrating how private charitable funding can represent

expertise, and assistance in converting to uniform record keeping, of whom Duke-supported facilities "were the largest group in the nation employing."

[7] A related set of papers evaluate the Rosenwald Rural Schools Initiative, a charitable effort that constructed nearly 5,000 schools for Black children living in the rural South. Prior work shows that this initiative had large, lasting effects as measured by changes in education, fertility, labor market outcomes, and incarceration probability (Aaronson and Mazumder 2011; Aaronson et al. 2014; Eriksson 2020; Mohammed and Mohnen 2022).

[8] Those establishing The Duke Endowment were careful to take into consideration issues common to other charitable organizations, such as crowding out (Andreoni and Payne 2013; Rask and Rask 2000), being careful to "exercis[e] their discretionary powers in assisting hospitals… to distinguish carefully and sharply between stimulating local interests and self help and substituting [for] local interests and self help." (Durdin 1998).

a blueprint for future public expenditures.[9] Second, the impact of the Endowment highlights the potential costs of closing rural hospitals today and the importance of attracting physicians to practice in rural and underserved communities.[10] If our results are interpreted as the effect of access to the first high-quality medical care in a community, then removing the last high-quality healthcare option may have a symmetric negative impact. 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 (McCoy et al. 2009; Yamey et al. 2019). A critical feature of the Endowment's success was its multifaceted treatment. Thus, we expect to see larger and longer-lasting health improvements in developing countries when donations targeted for health infrastructure are accompanied by reinforcing policies and oversight.

1 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 to improve the social and economic welfare of communities in North and South Carolina. In the Indenture of Trust for the Endowment, Duke recognized that hospitals were "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 hospitals (The Duke Endowment 1925, p. 10-11). Duke died unexpectedly in October 1925 and his will left the Endowment with an additional $69 million; 90% of the annual income from this

[9] Related work by Berkes and Nencka (2022) finds that libraries constructed using private charitable funds from Andrew Carnegie led to an increase in publicly funded libraries, which in turn increased long-run patenting.

[10] Germack et al. (2019) document that closures of rural hospitals lead to reductions in the supply of physicians in these communities which is consistent with and the exact reverse of our result that improved infrastructure attracted higher-quality physicians to treated counties. Similarly, Alexander and Richards (2021) find that rural hospital closures reduce the size of the healthcare workforce in the surrounding community. Efforts to study effects of rural hospital closures on health have produced mixed results. Fischer et al. (2022) find no statistically significant effect on infant mortality but their 95% confidence interval includes positive values as high as 7%, which is similar to our estimated gains in White infant mortality. Kozhimannil et al. (2018) find negative effects on birth outcomes and Gujral and Basu (2020) find large adverse mortality effects of closures that are not specific to infants.

contribution was specifically directed to hospitals in the Carolinas (Durdin 1998, p. 26).[11]

Following his 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 accepting applications for funding in 1926, and the first approvals were made 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 appropriation were made prior to 1927, but by 1929, almost 30% of counties had been promised financial support. In subsequent years the geographic spread slowed down as the share increased by only an additional 20 percentage points between 1929 and 1942. We present the spatial distribution of first appropriations by county and year in panel B of Figure 1.

Over the first two decades following the founding of The Duke Endowment, the number of hospitals in North Carolina increased significantly. At the time of the Endowment's baseline survey in 1925, there were 122 hospitals with 5,334 beds in 51 counties, with an average size of 44 beds per hospital. Nearly half of the counties (49%) did not have any hospital facility. By 1942, the end year of our preferred sample, only 7% of North Carolina's counties lacked a hospital. By 1949, the number of hospitals had increased 41% to 172, and the number of beds increased by 410% to 27,178, raising the average number of beds per hospital to 158.

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, directing 86% of appropriations and 79% of payments in North Carolina between 1927 and 1942 to hospitals that accepted patients of both races. White-only hospitals received 10% of capital appropriations (13% of payments), while Black-only hospitals received 4% (8% of payments).[12]

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

[12] While 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

In the same spirit, Duke support was only available for not-for-profit hospitals, an intentional measure designed to reduce financial barriers to access (Durdin 1998). These investments helped in narrowing the long-standing racial disparities in access to health care. Another of Duke's stated objectives in improving hospital facilities was to attract more highly-qualified doctors. In its first 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). Indeed, in North Carolina, physician density increased between 1923 and 1940 from 26 to 56 per 1,000 population.

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 hospital-level data on appropriations and payments for capital projects reported in the *Annual Reports of the Hospital Section* of The Duke Endowment. Online Appendix Figure A1 shows 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. 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 institutions. These costs were not usually incurred in the same fiscal year. Figure B1 in the Online Appendix shows the distribution of years since appropriation when an initial payment was received. Most payments (73%) were delivered by the start of the third year following an appropriation, while some appropriations (15%) never resulted in a payment. Our

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."

preferred estimates are intent-to-treat, which consider the first treated year to be the first year a non-zero appropriation from the Endowment occurs in a given county. Using this measure, panels A and B of Figure 1 show the rollout of Duke-support at the county-level across time and space.

From 1927 to 1942, 271 projects (in 48 counties) received an appropriation, out of which 164 projects (in 41 counties) received a non-zero payment. Most entries report a broad "purpose" category for the project (25% of approved projects had no category stated). Online Appendix Table B1 presents summary statistics for appropriations and payments by purpose. 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. Additions to existing not-for-profit hospitals represent the majority of approved projects at 48%, purchases of equipment and of existing facilities (often changing a for-profit facility into a not-for-profit) represent 13% and 10% of approved projects, respectively, and 6% of approvals were for building new facilities.

2.2 Births and mortality

Infant mortality: We construct a new measure of infant mortality by county of birth from the universe of North Carolina death certificates between 1917 and 1963 (Cook et al. 2014, 2016). Online Appendix Figure A1 shows an example of a death certificate. Importantly, the data include place and year of birth, which we use to aggregate the individual deaths to the county-by-year-of-birth level. Existing aggregate data on infant deaths for North Carolina during our study period are reported by place of *occurrence* only, which is potentially endogenous to hospitalization since more deaths are likely to occur in places where more hospitals are built, and are thus ill-suited to our objectives as they may introduce bias.[13] We define infant death as occurring between day 0 (i.e., the day of birth) and day 365 of life, limiting our sample to deaths of individuals born in North Carolina, because similar infant mortality data are not available for South Carolina.

[13] Data 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." 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.

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*. 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.[14] 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.

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, there is a precipitous decline in infant mortality for both races. 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. Nonetheless, we also collected data on Duke Endowment funding for the period up to 1962 and, in Section 5 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.[15] Third, as we discuss further in Section 5, two-way fixed effects estimators place greater weight

[14] The 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).

[15] After 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.

on groups treated in the middle of the panel, so extending the panel would shift the initial year of treatment to the early panel years for most counties, potentially biasing our results. Finally, this choice facilitates consistency between the short-run and long-run analyses.

Later-life mortality: We face two primary challenges in estimating the effect of Duke Endowment funding on later-life 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. 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 year of death. We link these data to exposure to Duke Endowment support at the birth-county by year-of-birth level. For administrative reasons, the public-use Numident is incomplete prior to 1988 (Goldstein et al. 2021), which affects the earliest year we are able to include in our long-run analysis.

Construction of the estimation sample for later-life 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. Here, a follow up year refers to the year in which we observe later-life 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. Second, we restrict the sample to the 1932 to 1942 birth cohorts due to the left- and right-censoring of the public-use Numident and the fact that we intend

to measure later-life mortality up to age 65, when individuals can retire and become eligible for Medicare. Deaths 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. Therefore, we start our follow up at age 56 and limit the sample to cohorts born in 1932 or later. The right-censoring of the public-use Numident in 2007 implies that we can observe mortality up to age 65 for cohorts born as late as 1942, and thus we do not extend the sample beyond the 1942 birth cohort.[16] To avoid biasing our estimates due to differential survival probabilities by age, we use a balanced panel of follow up years for each county-by-year birth cohort, which implies that all cohorts are observed exactly 10 times at ages 56 to 65 during follow up years (e.g., the 1932 cohort is observed from 1988 to 1997). We discuss the econometric implications of these sample restrictions in Section 7.

2.3 Other data

Physicians: One of our hypothesized mechanisms for the improvement in mortality is that the number and the quality of physicians who moved to counties that received Duke support both increased. We explore this possibility using data on physicians working in North Carolina from the volumes of the *American Medical Directory* (AMD) for 1914, 1918, 1923, 1927, 1931, 1936 and 1940. Online Appendix Figure A1 shows an extract of a page from a volume of the AMD. 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. We aggregate these data to the county-by-year level and scale the numerator by 100,000 population using data from NHGIS. We proxy for physician quality with the combination of years since graduation and the admission requirements of the medical school for their graduating cohort.[17] We define a doctor as high quality if they were licensed more than two years after the medical school they attended introduced a two-year degree requirement as an admission pre-requisite.

[16] A death at age 66 from the 1942 birth cohort and a death at age 65 from the 1943 would have occurred in 2008 and would only be found in the restricted-use data.

[17] See Moehling et al. (2020) for a discussion of how medical education admission requirements changed in the early twentieth century and were associated with improved physician quality.

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 NHGIS.

3 Methods

Our preferred analysis uses a difference-in-differences research design with staggered treatment adoption to estimate the effects of hospital modernization around the time of birth on infant and age 56 to 65 mortality rates. We assign treatment to individuals based on their county and year of birth, and consider a birth cohort in a county to be treated starting in the first year that the county received an appropriation for capital expenditures from The Duke Endowment and lasting through all subsequent years. We tackle two challenges for estimating the causal effect of Duke funding that arise in our empirical setting. First, we give careful consideration to the functional form of the dependent variable, which includes zeroes, in order to avoid model misspecification. Second, we implement recent innovations that address the concern 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).

Under our preferred estimation method of Poisson pseudo-maximum likelihood (Correia et al. 2020), coefficients of interest can be interpreted as a semi-elasticity – the percent change in the outcome due to Duke funding exposure around the time of birth. We prefer Poisson estimation to OLS for three reasons.[18] First, the Poisson model represent the natural log of the conditional mean of the dependent variable, rather than the conditional mean of the natural log of the dependent variable that is estimated by the common log-level specification. While the latter has been shown to suffer from bias due to Jensen's inequality when heteroskedasticity is present, the Poisson specification does not have this issue (Santos Silva and Tenreyro 2006). Second, all of our outcomes

[18]Throughout the paper, "Poisson" refers to the Poisson pseudo-maximum likelihood estimation method, rather than the assumption that the dependent variable has a Poisson distribution.

contain zeroes, sometimes in large numbers, which are explicitly modeled and handled well by the Poisson (Santos Silva and Tenreyro 2011). Thus, the Poisson specification obviates the need to drop zero-valued observations, as would happen with a dependent variable specified in logs, or manipulate the data to avoid dropping observations by adding one to each value of the dependent variable before taking logs. These are both common approaches that have been shown to introduce bias (Santos Silva and Tenreyro 2006; Cohn et al. 2021). Finally, in the later-life analysis, we are examining mortality rates that vary mechanically with age. The Poisson allows for a proportional analysis that ensures the comparability of estimates across specifications, and for the joint estimation of the long-run mortality effects while accounting for the changing baseline mortality rate. As robustness checks for the main specification of our outcomes and for comparability with other studies, we also estimate a linear model, a log-level model that drops any zero-value observations, and a log-level specification that adds one to the dependent variable for all observations.

Short-run mortality: For our short-run outcomes, we estimate the following regression:

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

where c indexes county of birth and t indexes birth year. Our primary outcome of interest Y^{R}_{ct} is the county-by-birth year count of deaths or infant mortality rate defined in Section 2.2. We also consider variations of this variable (e.g., first-day infant mortality rate) and other short-run outcomes (e.g., the number of high-quality physicians working in a county-year). We superscript this variable with R to indicate separate outcomes for Black and White persons or outcomes that pool together both racial groups.[19]

In Online Appendix Table C1 we perform a balance test following Pei et al. (2019), and we do not find a statistically significant relationship between changes in covariates and treatment status, easing concerns that selection or differential trends are driving our effects. Most notably, Endowment support does not have an effect on our best measure of economic activity, retail sales per

[19]Hispanic 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.

capita. This indicates that our findings are driven by changes in the healthcare provided rather than by general economic improvements due to the hospital modernization. Nonetheless, to reduce residual variation and decrease standard errors, in our preferred specification we add $\mathbf{X}_{ct}$, a vector of control variables including the percent of the population that is illiterate, Black, of other race, and residing in urban areas. It also includes 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 (Hoehn-Velasco 2018, 2021). Furthermore, 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. In our preferred Poisson specification, we weight observations by the number of births that occur in each county-year.[20]

Treatment effect: The treatment variable is $1(t \geqslant \text{First appropriation year})_{ct}$, 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 $exp(\alpha_1) - 1$, 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; Ciani and Fisher 2019; Lee and Lee 2021). This is an intent to treat effect, as capital appropriations from The Duke Endowment represented a commitment to provide financial resources that does not account for either the utilization (i.e., actual spending) or the allocation of funding (i.e., what the money was spent on).

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

[20]This method produces identical results to a model that uses the count of deaths as the outcome variable and includes an exposure variable equal to the number of births. We prefer our specification because it provides a clearer comparison to the alternative specifications we run as robustness checks.

$$Y_{ct}^{R} = exp(\alpha_0 + \sum_{j=-7}^{-2} \alpha_j 1(t - \text{First appropriation year}_{ct} = j) + \sum_{j=0}^{7} \alpha_j 1(t - \text{First appropriation year}_{ct} = j) + \zeta_c + \eta_t)\epsilon_{ct} \quad (2)$$

where j indicates the time period relative to the first appropriation year. The coefficients α_{-7} to α_{-2} capture trends prior to the start of the funding while coefficients α_0 to α_7 capture 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 and right-most coefficients, which are binned. Thus, the coefficient on α_{-7} indicates relative time periods between -18 and -7, while the coefficient on α_7 represents relative time periods between 7 and 15. The choice of cutoffs is informed by Online Appendix Figure F2, which shows that the number of treated counties by event time drops off beyond ± 6. The results are virtually unchanged if we do not bin the coefficients. We weight observations by county- and year-of-birth cohort size and do not include additional controls for ease of interpretation and to be in line with recent recommendations in the literature. Standard errors (ε_{ct}) are clustered by county of birth to account for correlated errors within a county.

Alternative estimators: While the recent literature has developed new DiD estimators that ensure that no previously treated counties are included in the control group, these methods typically involve linear models and thus they are not well-suited for our preferred specification. To circumvent this limitation and to ensure the internal validity of our estimates, we also implement a stacked regression model following Cengiz et al. (2019), which can be estimated by Poisson pseudo-maximum likelihood. We do so by creating treatment-timing group specific data sets that include counties first exposed to Duke Endowment funding in a particular year and "clean control" counties that have not yet been treated in various estimation windows. We stack these data sets and estimate DiD regressions and event study specifications that modify Equations 1 and 2, respectively, by saturating the county- and year-of-birth fixed effects with indicators for each of the stacked data sets. Our preferred estimation window runs from $t = -6$ to $t = 6$, but we consider smaller windows as alternatives in Online Appendix E to demonstrate robustness.

In Section 5, we also present results based on a variety of the new difference-in-differences estimators, with the caveat that these methods involve linear models. Thus, we execute our robustness specifications with $\log(Y_{ct}^{R})$ as the dependent variable and estimate the two-way fixed effects models using OLS rather than the preferred Poisson pseudo-maximum likelihood with Y_{ct}^{R}. For this $\log(Y_{ct}^{R})$ outcome, we conduct event studies using estimators from de Chaisemartin and D'Haultfœuille (2020), Callaway and Sant'Anna (2021), Gardner (2021), and Sun and Abraham (2021), as well as using two-way fixed-effects for comparison.

4 Results

We begin by presenting raw infant morality data by calendar year and treatment status. The first row of Figure 2 depicts infant mortality rates per 1,000 live births in North Carolina between 1922 and 1942 separately for counties that 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 and White infants (columns B and 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 should hold. Second, with the exception of the White sample, 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 under served communities. Third, after 1927, the infant mortality rate in treated counties dipped below 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 2 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 time period to serve as a simple comparison 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.

Having discussed the descriptive patterns in the infant mortality rates in North Carolina, we now turn to our main difference-in-differences results from estimating Equation 1, which are presented in Table 1. 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; in both cases it is measured at the county-by-year level. 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. Our preferred specification in column (6) uses the infant mortality rate as the dependent variable and includes both weights and the aforementioned set of control variables. Panel A presents pooled estimates while panels B and C estimate effects separately for Black and White infants, respectively.[21] 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 8.4% to 10.9% (panel A). The effects are larger for Black infants compared with White infants. In our preferred specification in column (6), the coefficients imply that Duke support caused a 10.0% reduction in pooled infant mortality, a 16.4% reduction in Black infant mortality, and a 7.1% reduction in White infant mortality. The difference between the race-specific estimates is statistically significant at the 5% level. Generally, we cannot reject the equality of coefficients across the six specifications, and thus we view them as being quantitatively and qualitatively similar. For this reason, unless otherwise noted, in the reminder of the paper we mostly focus on our preferred specification from column (6).

[21] There 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 steps we take to deal with cases of zero births or zero deaths, and other adjustments to birth or death counts, in Online Appendix D.

The average treatment effects documented in Table 1 are consistent with the graphical evidence and further validated by event studies based on estimating Equation 2. These results are presented in the bottom row of Figure 2. In each case, prior to the first dollar being appropriated to a county, there are no significant differences in infant mortality between treatment and control groups. This finding supports our identifying assumptions of parallel trends and no anticipation. By contrast, in the post-treatment periods we do see marked declines in mortality starting from the second year after the initial appropriation.[22] The statistically insignificant estimates in the first two post-period years are consistent with a lag between when the funds were appropriated by The Duke Endowment and when improvements to hospitals could be effectively implemented (Online Appendix Figure B1).

Magnitudes: To put our estimates into context, it is useful to 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 above imply that 27.8% 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 43.2% of the reduction in Black infant mortality but only 20.3% of the reduction for White infants.

We can also compare our effect sizes to other estimates in the literature, especially those ex-

[22] In Online Appendix C we document effects on fertility measured by the number of births per 1,000 women (Table C3) and on maternal mortality (Table C2). Point estimates for fertility neither imply major changes in childbearing due to improved health care infrastructure, nor suggest a cohort-composition concern that one might have when interpreting our main results. On the other hand, we find point estimates indicating reduced maternal mortality that are comparable in magnitude to our infant mortality estimates, but they are imprecisely estimated and in most cases are not statistically different from zero. This is not inconsistent with general trends in maternal mortality. From 1932 to 1942, maternal mortality in North Carolina declined by 51% from 7.2 to 3.5 deaths per 1,000 births (this was similar to the decline in infant mortality over the same time period of 41% from 66.7 to 39.5 deaths per 1,000 live births). Despite these large reductions, one reason for the lack of precision in our treatment effect estimate may be because the data are only reported beginning in 1932, and thus the sample includes roughly half the number of observations as in the main specification. Additionally, since maternal mortality is a rarer outcome than infant mortality, being roughly ten times less likely to occur for each live birth, the variation may be insufficient to precisely estimate the effects.

amining 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 (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). Finally, the introduction of water filtration in US cities between 1900 and 1940 led to 11 to 12% declines in infant mortality, but other public health and sanitation measures did not (Anderson et al. 2022).

Our estimated effects are also in line with results of more recent healthcare interventions. For example, Goodman-Bacon (2018) focuses on the introduction of Medicaid and finds that a 1 percentage point change in the initial Aid to Families with Dependent Children (AFDC) rate in a state lead to 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 1996a).

Finally, when comparing our treatment effect magnitudes to health interventions outside of the US, we again see close parallels. For example, the introduction of Bismark's compulsory social health insurance in the German Empire during the late-nineteenth century reduced infant mortality by 1.6%, explaining 22% of the decline in German infant mortality up to the beginning of the twentieth century (Bauernschuster et al. 2019). The introduction of Norwegian health care centers reduced infant mortality by up to 18% (Bütikofer et al. 2019) and a similar intervention in Sweden increased infant survival by about 24% (Bhalotra et al. 2017). Overall, the Scandinavian findings appear somewhat larger than what we find in North Carolina, perhaps because both interventions specifically targeted mothers and infants whereas The Duke Endowment supported general healthcare investments in the community. Given that we focus on individuals treated around the time of their birth, due to data limitations, and since hospital modernization likely benefited other groups as well, our estimates are likely conservative with respect to broader welfare improvements.

Intensive margin: While we have focused thus far on the extensive margin of appropriations from The Duke Endowment, we also consider intensive margin effects and conduct a cost-benefit analysis. For each county, we know the exact amount of money appropriated and paid out by the foundation. Thus, we re-estimate Equation 1 using cumulative appropriations or cumulative payments instead of a dummy treatment variable. The results of this analysis are presented in Table 2. Columns (1) and (2) present 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. It is clear from this analysis that only about half of the promised appropriations were paid out, highlighting the focus on frugality and the careful oversight from the foundation in how the money was spent. For example, Endowment payments could have been contingent upon certain bureaucratic standards being met by the hospital. On some occasions money appropriated, but not paid out, would carryover to the next year as another appropriation.

Due to the relatively low conversion rate, per dollar treatment effect estimates appear much larger for payments compared with appropriations. When we consider the actual money spent by the foundation, we find reductions in infant mortality in the range of 6.6% (White infants) to 9.1% (Black infants) per $1 million. Based on the amount appropriated and spent, one life was saved for every $39,000 appropriated or $20,000 spent (in 2017 dollars), implying that the Endowment's intervention passes any cost-benefit analysis using a reasonable estimate for a VSL.[23] Most other papers in the literature lack the detailed cost data needed to provide similar estimates, with the notable exceptions of Currie and Gruber (1996a) and Goodman-Bacon (2018), who find that an infant death could be averted at a cost of about $1.9 million and $175,000 (in 2017 dollars), respectively, which suggests that investment in hospital infrastructure is orders of magnitude cheaper than other previously studied interventions.

Heterogeneity: In Online Appendix Figure B2 we examine heterogeneity in the effects on infant mortality by four project categories targeted for funding: new hospitals, additions to hospitals,

[23]For 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.

equipment, and purchases of existing facilities (typically for conversion to a non-profit). Each estimate comes from a separate regression with the category of interest as the treatment, while excluding commitments for other project types. Regardless of whether we consider the extensive margin intent-to-treat specification (panel A), or the intensive margin of appropriations (panel B) or payments (panel C), we find that the effects of each project type, along with the effect of residual commitments, are comparable to the main estimates. The lone exception is for estimates of the effects of purchases of existing facilities, for which the impact of appropriations is noisy and the effect of payments is zero. Given that we still find an extensive margin effect of converting a private facility to a non-profit hospital, the weaker intensive margin estimates could imply that extending access to an existing facility does not convey additional benefits when it is simply more expensive. Overall, these findings support our interpretation of Duke support as a multifaceted and bundled treatment.

5 Robustness Checks

Point estimates and statistical significance for our main infant mortality results in Table 1 are robust to the choice of estimator, specification, transformation of the dependent variable, choice of the control group, or other sample restrictions. We provide a visual overview of many of these checks in Figure 3. Our event study findings in Figure 2 are likewise insensitive to the use of estimators designed to address concerns about treatment effect heterogeneity and bias in a difference-in-differences setting with staggered treatment adoption (Figure 4). This array of tests demonstrates that our results are not driven by the two primary econometric threats to identification: differential pre-trends and selection into treatment.

Stacked regression (Cengiz et al. 2019): While our preferred event study model is estimated by Poisson pseudo-maximum likelihood, most of the recently proposed estimators that address treatment effect heterogeneity are only implemented for linear models. As an alternative, we consider a stacked regression design following Cengiz et al. (2019) which can be estimated both by OLS and Poisson. In Online Appendix Figure E1 we report estimates from stacked Poisson event studies

and find treatment effects that are similar to the main Poisson event study estimates, as well as support for parallel pre-trends. We also document that our results are invariant to adjusting the width of the event-time window from ± 3 to ± 6 and insensitive to the inclusion of control variables. In Online Appendix Table E1 we further present Poisson and OLS average treatment effects for infant mortality rate.

Alternatives to two-way fixed effects: As it is common in the literature to estimate linear models with the natural log of the infant mortality rate per 1,000 live births as the dependent variable, we present results based on this specification in Figure 4. In addition to the standard two-way fixed effects OLS specification (blue X marks), we compute the estimators proposed by de Chaisemartin and D'Haultfœuille (2020) (red circles), Callaway and Sant'Anna (2021) (grey diamonds), Sun and Abraham (2021) (green triangles), and Gardner (2021) (orange squares). The results are similar across all of the estimators, ruling out bias due to treatment effect heterogeneity. As before, we see an absence of pre-trends.

Goodman-Bacon (2021a) decomposition: The validity of our event study results is further supported by the Goodman-Bacon (2021a) decomposition reported in Online Appendix Table F2 and Figure F1. The 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, as expected given the rapid roll out of the program and the reasonable fraction of untreated counties in our sample. Reassuringly, these comparisons give estimates close to the overall DiD estimate (Online Appendix Table F2). With one exception, the early versus later and later versus early treated comparisons are negative and smaller in magnitude compared to the overall DiD estimate. In all cases, these comparisons receive much lower weight (below 10%) than the treated versus untreated ones. These findings ease our concerns about bias in the two-way fixed effects DiD results stemming from comparisons of early versus later treated units.

Functional form and aggregation: Alongside the event study results discussed above, we show that our results from Table 1 are similarly robust. Online Appendix Table G1 reports estimates of linear models with the dependent variable Y_{ct}^{R} specified in levels, as $\log(Y_{ct}^{R})$, and as $\log(Y_{ct}^{R}+1)$, which indicate that the results are unaffected by the functional form of the dependent variable. For a subset of the alternate DiD estimators, we also aggregate the event study estimates to obtain ATT estimates that are comparable to those in Table 1. These estimates are reported in Online Appendix Table F1 and in Figure 3, and use the $\log(Y_{ct}^{R}+1)$ transformation as the outcome. The average treatment effects are comparable to or larger than those estimated using our preferred specification, ranging from negative 9.2% to negative 20.4%, which suggests that our preferred estimates my be especially conservative.

Non-Carolina controls: Next, in Online Appendix H, we replace North Carolina counties in the control group with southern counties that were never eligible for any form of assistance from the Endowment (i.e., counties in southern states excluding the Carolinas). This exercise addresses concerns about selection into treatment and spillover effects. 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 non-profit hospitals. 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 general equilibrium 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.

Our DiD regression results in Table H1 are almost identical to our preferred point estimates regardless of whether we add other southern counties to the control group (column 1) or if we also drop the North Carolina control counties (column 2). We also consider the possibility that non-profit hospitals in the treatment group could be more motivated or better equipped to attract funding – which could correlate with the quality of care they provide and, ultimately, with improved health

outcomes. Thus, we limit the control group to non-Carolina southern counties that had a non-profit hospital at some point during the sample period (column 3). Our results and conclusions remain unchanged. Event study results – based on the sample from column 2 of Table H1 – presented in Figure H1 likewise support the robustness of our results.

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 exploits 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. We interact the temporal and cross-sectional sources of variation to create our instrument. The first term in the interaction is the cumulative returns of The Duke Endowment's assets.[24] 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 2, Table H2 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 for Duke-support on the instrument, the reduced form regression of log infant mortality rate on the instru-

[24] We obtained original financial statements of The Duke Endowment containing these data from the Joseph and Matthew Payton Philanthropic Studies Library. 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.

ment, and an instrumental variable specification.[25] We consider the effects for two samples. The first sample (panel A) ranges from 1930 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.[26] The second sample (panel B) includes all North Carolina counties from 1922-1942.

Compared to the Poisson and OLS based intensive margin results, the instrumental variables framework yields larger point estimates that are still statistically significant at conventional levels. 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 million dollars of Duke appropriations in a county, the infant mortality rate declined between 9 and 11%. For payments, the effect sizes are between 15 and 17% per one million dollars. These point estimates are larger than their non-IV counterparts. 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.[27] The finding of a larger, negative effect mitigates these concerns and provides additional evidence in favor of our overall argument.

Randomization inference: We next consider 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 I1, we present results from a randomization inference exercise (based on the specification from column 5 of Table 1) where we randomly assign 48 of the North Carolina counties to the treatment path of the true treated 48 counties. As

[25] 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.

[26] These data are from Bailey et al. (2015) and are unbalanced before 1930 and after 1940.

[27] 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.

expected under the no-selection hypothesis, the average placebo estimate is symmetric and centered around zero. We also use this exercise to construct two-sided randomization inference based p-values for our preferred point estimates (Young 2019). We 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. Using this procedure, our results remain statistically significant at conventional levels, and p-values are largely unaffected, although they are slightly larger at 2% for the pooled specification, 3% for the Black specification, and 6% for the White specification. This procedure indicates that our results are not driven by spurious time trends in infant mortality across all counties or by spillovers across treated versus untreated counties in North Carolina.

Sample restrictions: When 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 bias our estimates due to dynamic treatment effects that vary across treatment timing groups (Goodman-Bacon 2021a; Baker et al. 2022). Thus, in Figure 3, we document that our results are not sensitive to changing the endpoints of the panel. Additionally, the results are not affected by including the World War II period in the sample or by changing the start of the panel.

A related concern is a potential bias due to the inclusion of small counties in the sample. Although we weight observations by county-by-year-of-birth cohort size in our preferred specification, doing so may not eliminate such bias entirely. We confirm this bias is not an issue by sequentially dropping the smallest counties (in increasing increments of five) until only half of the original sample remains. Our results are largely unaffected by these changes to the sample composition.

We also explore whether our results are driven by mortality immediately after birth, a period

with the largest scope for measurement error in recording deaths.[28] In Online Appendix Table C4 we estimate effects on day-one mortality and test the sensitivity of the main results to excluding deaths on the first day, in the first week, or in the first month. While we find somewhat more negative coefficients for day-one mortality, our results are unaffected by dropping deaths reported in the first month of a child's life. Therefore, we are not concerned that measurement error in deaths during early infancy is driving our results.

Finally, it may be possible that Duke support changed the likelihood of birth occurring in a supported county, which is important given that we use the number of births in some specifications to create both the infant mortality rate and to weight observations. To ensure that this potential bias does not affect our results, we consider specifications that do not include the number of births and focus solely on deaths. Our overall findings are largely unaffected by this restriction. For example, column (1) of Table 1 uses the count of infant deaths by county and year of birth and includes no weights. Here, we find a 9% reduction in pooled infant mortality (vs. 10% for our preferred specification that includes births), a 13% (vs. 16%) reduction for Black infants, and a 6% (vs. 7%) reduction for White infants.

6 Mechanisms and complementarities

6.1 Effects on physician labor market

Next, we consider whether the effects of support from The Duke Endowment extended beyond the patients and into the physician labor market. Even though Duke support was not directly tied to hiring physicians, those who ran the foundation believed that improved funding and working conditions in North Carolina's hospitals should attract higher-quality doctors. Table 3 presents the physician labor supply results in which we focus on the extensive margin of the treatment, as in our preferred set of results for infant mortality in Table 1.

We document the effects on the number of doctors (columns 1-3) and doctors per 100,000 population (columns 4-6) measured at the county-by-year level. Column (1) represents the simplest

[28] In our main specifications, we follow published infant mortality statistics in restricting attention to live births, thereby excluding stillbirths. In Online Appendix C, we examine potential measurement error in recording stillbirths.

specification, where weights and controls are not included and the dependent variable is the count of physicians.[29] Column (6) is analogous to our preferred specification from the infant mortality results. Outcomes include all physicians (panel A) as well as high quality (panel B) and low quality (panel C) doctors. Physician quality is based on whether the medical school a physician graduated from required a 2-year undergraduate degree at the time of their admission (Moehling et al. 2020).

In comparison to our infant mortality findings, these results are somewhat more sensitive to the econometric specification, but they broadly point towards two conclusions. First, in three out of six specifications there is statistically significant positive effect on the total labor supply of physicians in a county. Second, this effect is driven by two offsetting forces: an increase in higher-quality physicians and a decrease in lower quality physicians. Indeed, in the preferred specification in column (6), we find a 10.7% increase and a 9.0% decrease in these two groups, respectively. This substitution effect led to the average graduation year of physician shifting forward in time by 12 years, thereby realizing the Endowment's aim of attracting better-qualified physicians to the hospitals it supported.[30]

Importantly, our results are not driven by a reshuffling of physicians within North Carolina. Online Appendix Figure C1 shows that the number of high and low quality physicians per 100,000 population 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, but that decrease was greater in treated counties.

For brevity, we do not perform all of the robustness checks for the physician outcomes. How-

[29]This specification is not our preferred one, since it weights counties with small and large populations equally. Nonetheless, we find the results informative because population is not included as either a weight or in the denominator of the dependent variable (i.e., in constructing a rate). Thus, the results sidestep potential concerns that population may be responding to Duke support, thereby obscuring the interpretation of our treatment effect estimates.

[30]Physician age is positively correlated with experience and could be viewed as a proxy for quality, but in our setting, only the younger doctors were educated in higher-quality institutions in which modern medical practices and techniques were taught. Many of the older physicians obtained their education from schools that dissolved after the Flexner report (Moehling et al. 2020).

ever, in Online Appendix Table G2, we demonstrate that our findings are robust to estimators other than Poisson and are qualitatively similar to the preferred results when using OLS, a log-level specification that drops zero-valued observations, and a log-level specification that adds one to all observations.

6.2 Interactions with medical innovation

Given the meaningful improvements in the average human capital of physicians practicing in Duke-supported counties that we documented above, we now ask if medical innovation in the form of sulfa drugs – the first modern treatment that could effectively control and fight bacterial infections – had differential effects in these locations. In other words, we examine whether the effects of Duke support are magnified (complements) or diminished (substitutes) by medical innovation. For example, it is plausible that while doctors were undergoing training in selective medical schools, they were exposed to new medical interventions and thus they may have been more likely to administer them, use them more effectively, and inform their colleagues about them.[31] To the extent that sulfa drugs have positive effects (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. Finally, the inputs might be independent as some prior "lightning strikes twice" studies have found no interaction effects (Duque et al. 2021).

In order to answer this question, we follow a three-steps procedure. First, we have already estimated the effects of Duke support using Equation 1, as discussed in Section 3 and presented in Table 1. Second, in Online Appendix C, we estimate the effects of sulfa drugs on infant mortality using a within North Carolina shift-share design. The findings, based on our preferred specification from column 3 of Online Appendix Table C6, suggest that sulfa drugs reduced infant mortality

[31] For example, Escarce (1996) shows that access to information about laparoscopic cholecystectomy influenced surgeons' adoption behavior, while Centola (2010) shows that individuals are more likely to register for a health forum if they received social reinforcement from their neighbors in the social network. This evidence is consistent with a broader literature on peer effects and social networks in technology adoption that mostly uses experimental evidence from agriculture in developing countries (Conley and Udry 2010; Beaman et al. 2021).

from 6.5% to 10.5%. Third, given the results presented in Tables 1 and C6, we ask if the two inputs – Duke support and sulfa drugs – are complements, substitutes, or independent treatments, by interacting our roll-out (Duke support) and shift-share (sulfa drugs) difference-in-differences designs and estimating the following equation:

$$\begin{aligned} Y^{R}_{ct} &= \exp(\gamma_0 + \gamma_1 1(t \geqslant \text{First appropriation year})_{ct} \\ &+ \gamma_2 1(t \geqslant \text{First appropriation year})_{ct} \times \text{Post sulfa}_t \\ &+ \gamma_3 1(t \geqslant \text{First appropriation year})_{ct} \times \text{Pneumonia mortality}_c \\ &+ \gamma_4 \text{Pneumonia mortality}_c \times \text{Post sulfa}_t \\ &+ \gamma_5 1(t \geqslant \text{First appropriation year})_{ct} \times \text{Pneumonia mortality}_c \times \text{Post sulfa}_t \\ &+ \zeta_c + \eta_t + \mathbf{\Theta X_{ct}}) \epsilon_{ct} \end{aligned} \tag{3}$$

where Y^{R}_{ct}, ζ_c, η_t, $\mathbf{X_{ct}}$, and $1(t \geqslant \text{First appropriation year})_{ct}$ are defined as in Equation 1. The variable Pneumonia mortality$_c$ is defined as the average county-level pneumonia mortality for the years 1922 to 1926 (our share factor) while the variable Post sulfa$_t$ takes the value of 1 for the years 1937 and later, and zero for prior years (our shift factor). 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, across counties with different levels of baseline pneumonia mortality. The auxiliary parameters, γ_1 to γ_4, further enable 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 4 presents linear combinations of these coefficients to ease the interpretation. Panel A presents an estimate for the interaction of Duke funding with the introduction of sulfa drugs when moving from a county in the 25th percentile of the baseline pneumonia mortality rate to one in the 75th percentile. It suggests that gains from Duke support grew when medical innovation became available, and did so to a greater extent in counties (with high baseline pneumonia mortality) poised to benefit from the new treatment, 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, but are in contrast with some studies that focus on human capital formation which find shocks to be substitutes rather than complements (e.g., Rossin-Slater and Wüst 2020; Bütikofer et al. 2021).

In panels B and C, we unpack this interaction coefficient into its components by computing the effects of Duke support separately in locations with and without high baseline pneumonia, as well as before and after the introduction of sulfa drugs. An inspection of these coefficients facilitates the interpretation of the interaction term. First, there are no effects of Duke funding in low baseline pneumonia mortality counties either before or after the introduction of sulfa drugs (rows 2 and 4 in Panel B). Furthermore, the pre-vs-post sulfa difference in the effect of Duke in these counties is small and insignificant (row 2 of Panel C). Second, in high baseline pneumonia mortality counties, coefficients on Duke support are statistically significant at conventional levels in both the pre- and post-sulfa periods (rows 1 and 3 of Panel B), but they are larger after sulfa drugs become available. Put differently, the reduction in infant mortality due to Duke funding is increasing with baseline pneumonia mortality, but to a greater extent in the post-sulfa period (rows 3 and 4 of Panel C). We conclude from these findings that reductions in infant mortality due to hospital modernization occurred even absent medical innovation. Moreover, the result that the effect of Duke support is not only larger in the post-sulfa period, but also increasing in baseline pneumonia mortality indicates that the hospital modernization was complementary with medical innovation.

7 Long-run mortality

Having documented that Duke funding reduced infant mortality, disproportionately so for Black compared to White infants, and discussed two plausible mechanisms behind these findings, we next ask if these health gains translated into reductions in later-life mortality for those exposed to Duke funding around the time of their birth. All our long-run analyses are conditional on survival to at least age 56 for reasons described in Section 2. 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, Duke funding might have induced the most vulnerable individuals to survive past the first

year of life. If these individuals likewise survived past age 55, then we could expect to find null effects or even see increases in later-life mortality steaming from the fact that those adults saved as children may still be more unhealthy and more likely to die than peers who were not reliant on Duke support for surviving infancy.

We begin by presenting descriptive plots of later-life mortality at ages 56 to 65, separately for cohorts exposed (pink solid line) and not exposed (black dashed line) to Duke funding around the time of birth during our sample period (Online Appendix Figure C2). Regardless of whether we examine effects on overall later-life mortality, or mortality separately by race, we find that individuals who were not treated around the time of birth had persistently higher mortality at all ages from 56 to 65. Furthermore, in all cases, the two lines diverge with age, implying a steeper age-mortality gradient among cohorts that were not affected by Duke support. Equipped with this descriptive evidence, we formalize the analysis and estimate the following equation:

$$Y^{R}_{ctf} = exp(\beta_0 + \beta_1 1(t \geqslant \text{First appropriation year})_{ct} + \zeta_c + \eta_t + \mathbf{\Theta X_{ctf}})\epsilon_{ctf} \tag{4}$$

which is very similar to Equation 1 with two exceptions: (i) we stack birth county-by-year observations over follow up years f and (ii) due to the incompleteness of the Numident prior to 1988 (as described in Section 2) we rely on the 1932 to 1942 birth cohorts, which can all be observed for 10 follow up years between ages 56 and 65. Using Poisson maximum likelihood estimation is particularly relevant in this context since the probability of death increases with age (Online Appendix Figure C2). In Equation 4, Y^{R}_{ctf} is the number of deaths or the death rate per 1,000 population for those born in county c and in birth cohort t and observed in follow-up year f.[32] 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.[33]

[32] Since we observe each birth county-by-year cohort at ages 56 to 65 during follow up, we allow baseline controls to flexibly vary with follow up years. We also verified that our results are invariant to including follow-up year fixed effects as well as birth-cohort-by-follow-up-year fixed effects. These additional control variables are included in the vector $\mathbf{X_{ctf}}$.

[33] We do not present event studies for the long-run outcomes as these are problematic to interpret for three reasons. First and foremost, since we are interested in effects on long-run outcomes that operate through exposure around the time of birth, and any subsequent exposures or changes could be endogenous to the original treatment, the relevant parallel trends assumption applies to the infant mortality rate. Second, cohorts represented by the pre-treatment co-

We present the long-run results in Table 5 which is structured similarly to Table 1 with the number of later-life deaths as the dependent variable in columns 1 to 3 and the later-life mortality rate in columns 4 to 6. We present pooled estimates in panel A and the effects separately for Black and White adults in panels B and C, respectively. In each case, the reported coefficients correspond to effects of treatment around the time of birth on mortality between ages 56 and 65, conditional on survival to age 56. We find consistently negative and statistically significant coefficients, with the exception of column 4 which does not include weights and thus estimates effects for an average county rather than an average individual. In our preferred specification (column 6), the average treatment effect is a 8.8% reduction in long-run mortality. This effect size is slightly smaller than our preferred estimate from Table 1, but it is close enough for us to conclude that much of the gains that accrued in infancy seem to have persisted to adulthood.

When it comes to the racial differences, the effects on Black and White later-life mortality are very similar across specifications, a striking difference from our infant mortality results in which the effects were twice as large for Blacks. This discrepancy could stem 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 the known differences in life expectancy at birth for Black and White babies.[34] Alternatively, it could simply be that the marginal effects of Duke support were not differential by race beyond age 56. This result is different from Goodman-Bacon (2021b) who finds that early age Medicaid eligibility reduced non-Aids adult mortality before age 37 by 14.5% and 8.7% for White and non-White individuals, respectively.[35] Quantitatively our estimates appear to be of similar magnitude as the estimates from a relatively limited literature which can trace mortality from

efficient α_{-2} in Equation 2 should not have had their infant mortality affected, but could have been treated at age 2 and thereafter, thus plausibly affecting their later-life mortality. Finally, due to the incompleteness of the Numident, we use the 1932 to 1942 birth cohorts rather than the 1922 to 1942 cohorts in the long-run analysis, resulting in a relatively shorter panel.

[34]For 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.

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

childhood interventions, such as providing health insurance (Goodman-Bacon 2021b) or access to in-hospital deliveries (Fischer et al. 2021), to adulthood. On the other hand, they are much larger than the effect sizes of up to 0.7% for survival until age 64 from a nurse home visiting program (Hjort et al. 2017). Given that the average effect sizes for the pooled population persisted almost unchanged from infancy to pre-retirement adulthood, we conclude that Duke support improved the long-run welfare of those treated around the time of birth.

Due to data limitations and for brevity we do not preform all of the robustness checks for long-run mortality that we do for infant mortality. However, in Online Appendix Table G3, we demonstrate that our findings are robust to estimators other than Poisson and are similar when using OLS, and a log-level specification that drops zero-valued observations. The only imprecisely estimated specification is when the value one has been added to the outcome for every observation, which is not unexpected as there are a large number of zero-valued observations in this analysis, so adding one represents a substantial alteration.

In a manner analogous to the infant mortality results, we also conduct analyses with non-Carolina control counties that were ineligible for Duke funding. These results are presented in columns (4) through (6) of Online Appendix Table H1: column (4) adds the non-Carolina data as additional untreated observations, column (5) then excludes untreated observations from North Carolina, and column (6) restricts the non-North Carolina counties to be those with at least one not-for-profit hospital. Across all of these specifications, the treatment effect estimates are qualitatively similar to those from our preferred specification. However, the inclusion of non-Carolina data does increase the difference between the point estimates for the effect of treatment on White and on Black mortality rates.

8 Conclusions

The barriers that limit access to effective health care in the US today – a lack of providers and facilities in rural and impoverished areas, financial difficulties, differential access by race and socioeconomic status, insufficient insurance, and structural impediments such as a lack of transportation – are not unique to the modern era. These were also challenges almost 100 years ago in

1924, when James Buchanan Duke established The Duke Endowment to relieve such burdens in the Carolinas. In this paper, we use 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 for infant and later-life mortality, with particular focus on racial gaps.

Our results indicate that Duke support reduced average infant mortality by about 10% at a cost of $20,000 (2017 dollars) per life saved, which can be viewed as extremely cost-effective compared with any reasonable VSL estimate or prior literature on the introduction of Medicaid or its expansion. These gains were larger for Black infants compared with White infants at 16.4% and 7.1%, respectively, leading to the closing of the pre-intervention racial gap in infant mortality gap by one-third. Furthermore, they persisted until at least modern retirement age for those exposed around the time of birth, reducing the average mortality between the ages 56 and 65 by 9%.

We propose and provide empirical evidence for two mechanisms which could explain such sizeable gains in health. First, we document that Duke support had a positive impact on the labor market for physicians in affected counties, which attracted better educated and more recently trained physicians. Importantly, improvements in human capital in treated counties are unlikely to have resulted in parallel declines in control counties. Second, counties supported by the foundation disproportionately benefited from the invention of sulfa drugs. This finding provides some of the first causal evidence that medical infrastructure and medical innovation are complements rather than substitutes. Better-educated physicians may also have been more likely to adopt such innovation, use it more effectively, and transfer this knowledge though physicians' networks in the county, thus making innovation more productive in Duke-supported counties.

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 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. Given our findings, and the fact that the program was modelled after Duke support, it is 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. 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 (Davis et al. 2017).

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. It is likely that the effects we document for North Carolina were mirrored in South Carolina, but we lack data to evaluate this hypothesis. Since we find large and positive health effects that persist in the long-run, 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, and because it induced permanent changes in access to modern hospitals in supported communities, Duke-support could have affected many more people in different age and sex groups than are considered in this paper, rendering our estimates a lower bound on the total positive welfare effects. We believe that part of the long-run reductions in mortality we observe could be explained by the

fact that those treated around the time of birth were also likely to be treated in childhood and perhaps even in adulthood to the extent that they did not migrate elsewhere. 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.

Acemoglu, Daron and Simon Johnson (2007) "Disease and Development: The Effect of Life Expectancy on Economic Growth," *Journal of Political Economy*, Vol. 115, No. 6, pp. 925–985.

Adhvaryu, Achyuta, Anant Nyshadham, Teresa Molina, and Jorge Tamayo (2021) "Helping Children Catch Up: Early Life Shocks and the Progresa Experiment," *Economic Journal*, Vol. forthcoming.

Alexander, Diane E and Michael R Richards (2021) "Economic Consequences of Hospital Closures," NBER Working Paper No. 29110.

Almond, Douglas and Janet Currie (2011) "Killing Me Softly: The Fetal Origins Hypothesis," *Journal of Economic Perspectives*, Vol. 25, No. 3, pp. 153–172.

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.

Almond, Douglas Vincent, Kenneth Y. Chay, and Michael Greenstone (2006) "Civil Rights, the War on Poverty, and Black-White Convergence in Infant Mortality in the Rural South and Mississippi," Working Paper.

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.

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, Claudio Las Heras Olivares, and Daniel I. Rees (2019) "Was the First Public Health Campaign Successful?" *American Economic Journal: Applied Economics*, Vol. 11, No. 2, pp. 143–175.

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.

Andreoni, James and A. Abigail Payne (2013) "Charitable Giving," in *Handbook of Public Economics*, Vol. 5: Elsevier, pp. 1–50.

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.

Bauernschuster, Stefan, Anastasia Driva, and Erik Hornung (2019) "Bismarck's Health Insurance and the Mortality Decline," *Journal of the European Economic Association*, Vol. 17, No. 6, pp. 1–47.

Beaman, Lori, Ariel Ben Yishay, Jeremy Magruder, and Ahmed Mushfiq Mobarak (2021) "Can Network Theory-Based Targetting Increase Technology Adoption?" *American Economic Review*, Vol. 111, No. 6, pp. 1918–1943.

Berkes, Enrico and Peter Nencka (2022) "Knowledge Access: Effects of Carnegie Libraries on Innovation," Working Paper.

Bhalotra, Sonia, Martin Karlsson, and Therese Nilsson (2017) "Infant Health and Longevity: Evidence from a Historical Intervenetion in Sweden," *Journal of the European Economic Association*, Vol. 15, No. 5, pp. 1101–1157.

Bhalotra, Sonia and Atheendar Venkataramani (2015) "Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions," Working Paper.

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 Hookwarm 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.

Bloom, Nicholas, Renata Lemos, Raffaella Sadun, and John Van Reenen (2020) "Healthy Business? Managerial Education and Management in Health Care," *The Review of Economics and Statistics*, Vol. 102, No. 3, pp. 506–517.

Brown, David, Amanda Kowalski, and Ithai Lurie (2020) "Long-term Impacts of Childhood Medicaid Expansions on Outcomes in Adulthood," *Review of Economic Studies*, Vol. 87, No. 2, pp. 792–821.

Bütikofer, Aline, Krzysztof Karbownik, and Ingrid Semb (2021) "Getting a Second Chance? Prenatal Health Shocks and Infant Health Care," Working Paper.

Bütikofer, Aline, Katrine Løken, and Kjell Salvanes (2019) "Infant Health Care and Long-term Outcomes," *Review of Economics and Statistics*, Vol. 101, No. 2, pp. 341–354.

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.

Centola, Damon (2010) "The Spread of Behavior in an Online Social Network Experiment," *Science*, Vol. 329, No. 5996, pp. 1194–1197.

de Chaisemartin, Clément and Xavier D'Haultfœuille (2020) "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, Vol. 110, No. 9, pp. 2964–96.

Chay, K. Y. and M. Greenstone (2003) "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *The Quarterly Journal of Economics*, Vol. 118, No. 3, pp. 1121–1167.

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.

Ciani, Emanuele and Paul Fisher (2019) "Dif-in-Dif Estimators of Multiplicative Treatment Effects," *Journal of Econometric Methods*, Vol. 8, No. 1, p. 20160011.

Clay, Karen, Peter Juul Egedesø, Casper Worm Hansen, Peter Sandholt Jensen, and Avery Calkins (2020) "Controlling Tuberculosis? Evidence from the First Community-Wide Health Experiment," *Journal of Development Economics*, Vol. 146, p. 102510.

Clay, Karen, Joshua Lewis, and Edson Severnini (2016) "Canary in a Coal Mine: Infant Mortality, Property Values, and Tradeoffs Associated with Mid-20th Century Air Pollution," *Review of Economics and Statistics*, Vol. forthcoming.

Clay, Karen, Warner Troesken, and Michael Haines (2014) "Lead and Mortality," *Review of Economics and Statistics*, Vol. 96, No. 3, pp. 458–470.

CMS (2022) "NHE Fact Sheet | CMS," U.S. Centers for Medicare & Medicaid Services.

Cohn, Jonathan, Zach Liu, and Malcolm Wardlaw (2021) "Count Data in Finance," Working Paper.

Cohodes, Sarah, Daniel Grossman, Samuel Kleiner, and Michael Lovenheim (2016) "The Effect of Child Health Insurance Access on Schooling. Evidence from Public Insurance Expansions," *Journal of Human Resources*, Vol. 51, No. 3, pp. 727–759.

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.

Conley, Timothy and Christopher Udry (2010) "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, Vol. 100, No. 1, pp. 35–69.

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.

Cunha, Flavio and James Heckman (2007) "The Technology of Skill Formation," *American Economic Review*, Vol. 97, No. 2, pp. 31–47.

Currie, J. and J. Gruber (1996b) "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *The Quarterly Journal of Economics*, Vol. 111, No. 2, pp. 431–466.

Currie, Janet and Jonathan Gruber (1996a) "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.

Cutler, David and Grant Miller (2005) "The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States," *Demography*, Vol. 41, No. 1, pp. 1–22.

Davis, Jonathan, Jonathan Guryan, Kelly Hallberg, and Jens Ludwig (2017) "The Economics of Scale-up," Working Paper.

Duque, Valentina, Maria Rosales-Rueda, and Fabio Sanchez (2021) "How Do Early-Life Shocks Interact with Subsequent Human Capital Investments? Evidence from Administrative Data," Working Paper.

Durdin, Robert F. (1998) *Lasting Legacy to the Carolinas*, Durham, NC: Duke University Press.

Elder, Todd E, John H Goddeeris, and Steven J Haider (2011) "A Deadly Disparity: A Unified Assessment of the Black-White Infant Mortality Gap," *The B.E. Journal of Economic Analysis & Policy*, Vol. 11, No. 1.

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.

Eriksson, Katherine (2020) "Education and Incarceration in the Jim Crow South: Evidence from Rosenwald Schools," *Journal of Human Resources*, Vol. 55, No. 1, pp. 43–75.

Escarce, Joes (1996) "Externalities in Hospitals and Physician Adoption of a New Surgical Technology: An Exploratory Analysis," *Journal of Health Economics*, Vol. 15, No. 6, pp. 715–734.

Ewbank, Douglas (1987) "History of Black Mortality and Health before 1940," *Milbank Quarterly*, Vol. 65, No. 1, pp. 100–128.

Ferrie, Joseph and Warner Troesken (2008) "Water and Chicago's Mortality Transition, 1850-1925," *Explorations in Economic History*, Vol. 45, No. 1, pp. 1–16.

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 Prodromidis (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. Horrace, 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.

Gardner, John (2021) "Two-stage Difference-in-Differences," Working Paper.

Germack, Hayley Drew, Ryan Kandrack, and Grant Martsolf (2019) "When Rural Hospitals Close, the Physician Worlforce Goes," *Health Affairs*, Vol. 38, No. 12, pp. 2086–2094.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift (2020) "Bartik Instrments: What, When, Why, and How?" *American Economic Review*, Vol. 110, No. 8, pp. 2586–2624.

Goldstein, Joshua R., Monica Alexander, Casey Breen, Andrea Miranda Gonzàlez, 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.

Gray-Lobe, Guthrie, Anthony Keats, Michael Kremer, Isaac Mbiti, and Owen W. Ozier (2022) "Can Education Be Standardized? Evidence from Kenya," Working Paper.

Greenstone, Michael and Rema Hanna (2014) "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India," *American Economic Review*, Vol. 104, No. 10, pp. 3038–3072.

Gujral, Kritee and Anirban Basu (2020) "Impact of Rural and Urban Hospitals Closures on Inpatient Mortality," NBER Working Paper No. 26182.

Gunnsteinsson, Snaebjorn, Achyuta Advaryu, Paul Christian, Alain Labrique, Jonathan Sugimoto, Aub Ahmed Shamim, and Keith West Jr (2019) "Protecting Infants from Natural Disasters: The Case of Vitamin A Supplementation and a Tornado in Bangladesh," Working Paper.

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.

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.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders (2016) "Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health," *Review of Economics and Statistics*, Vol. 98, No. 2, pp. 350–366.

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.

Law, Marc T. and Sukkoo Kim (2005) "Specialization and Regulation: The Rise of Professionals and the Emergence of Occupational Licensing Regulation," *The Journal of Economic History*, Vol. 65, No. 3, pp. 723–756.

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.

——— (2021) "It's a Long Walk: Lasting Effects of Maternity Ward Openings on Labour Market Performance," *Review of Economics and Statistics*, Vol. forthcoming.

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.

Lee, Myoung-jae and Sanghyeok Lee (2021) "Difference in Differences and Ratio in Ratios for Limited Dependent Variables," arXiv:2111.12948.

Lührmann, Melanie and Tanya Wilson (2018) "Long-run Health and Mortality Effects of Exposure to Universal Health Care at Birth," Working Paper.

McConnell, K. John, Richard C. Lindrooth, Douglas R. Wholey, Thomas M. Maddox, and Nick Bloom (2016) "Modern Management Practices and Hospital Admissions," *Health Economics*, Vol. 25, No. 4, pp. 470–485.

McCoy, David, Sudeep Chand, and Devi Sridhar (2009) "Global Health Funding: How Much, Where it Comes From and Where it Goes," *Health Policy and Planning*, Vol. 24, pp. 407–417.

Miller, Doug (2001) "What Underlies the Black-White Infant Mortality Gap? The Importance of Birth Weight, Behavior, Environment, and Health Care," Working Paper.

Miller, Sarah and Laura Wherry (2019) "The Long-term Effects of Early Life Medicaid Coverage," *Journal of Human Resources*, Vol. 54, No. 3, pp. 785–824.

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.

Mohammed, A R Shariq and Paul Mohnen (2022) "Black Economic Progress in the Jim Crow South: Evidence from Rosenwald Schools," Working Paper.

National Center for Health Statistics (2017) "Health, United States, 2017: Table 015," Hyattsville, MD.

Olds, David L., Charles R. Henderson, Harriet J. Kitzman, John J. Eckenrode, Robert E. Cole, and Robert C. Tatelbaum (1999) "Prenatal and Infancy Home Visitation by Nurses: Recent Findings," *The Future of Children*, Vol. 9, No. 1, p. 44.

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.

Rask, Kevin N and Kimberly J Rask (2000) "Public Insurance Substituting for Private Insurance: New Evidence Regarding Public Hospitals, Uncompensated Care Funds, and Medicaid," *Journal of Health Economics*, Vol. 19, No. 1, pp. 1–31.

Rossin-Slater, Maya and Miriam Wüst (2020) "What is the Added Value of Preschool for Poor Children? Long-term and Intergenerational Impacts and Interactions with an Infant Health Intervention," *American Economic Journal: Applied Economics*, Vol. 12, No. 3.

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 Tenreyro (2006) "The Log of Gravity," *The Review of Economics and Statistics*, Vol. 88, No. 4, pp. 641–658.

——— (2011) "Further Simulation Evidence on the Performance of the Poisson Pseudo-maximum Likelihood Estimator," *Economics Letters*, Vol. 112, No. 2, pp. 220–222.

Scur, Daniela, Raffaella Sadun, John Van Reenen, Renata Lemos, and Nicholas Bloom (2021) "The World Management Survey at 18: Lessons and the Way Forward," *Oxford Review of Economic Policy*, Vol. 37, No. 2, pp. 231–258.

Soares, Rodrigo R. (2007) "On the Determinants of Mortality Reductions in the Developing World," *Population and Development Review*, Vol. 33, No. 2, pp. 247–287.

Starr, Paul (2017) *The Social Transformation of American Medicine*, New York: Basic Books, updated edition edition.

Sun, Liyang and Sarah Abraham (2021) "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, Vol. 225, No. 2, pp. 175–199.

Sutch, Richard (2006) "National Income and Product," 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.

The Duke Endowment (1925) *Annual Report of the Hospital Section of the Duke Endowment for the year 1925*: 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 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.

Tsai, Thomas C., Ashish K. Jha, Atul A. Gawande, Robert S. Huckman, Nicholas Bloom, and Raffaella Sadun (2015) "Hospital Board And Management Practices Are Strongly Related To Hospital Performance On Clinical Quality Metrics," *Health Affairs*, Vol. 34, No. 8, pp. 1304–1311.

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.

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.

Young, Alwyn (2019) "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," *Quarterly Journal of Economics*, Vol. 134, No. 2, pp. 557–598.

9 Tables

Table 1. Effect of Duke support on infant deaths and infant mortality rate: Extensive margin intent-to-treat estimates

	Y_{ct}^{R} = Infant deaths			Y_{ct}^{R} = Infant mortality rate		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Pooled infant deaths or infant mortality rate						
Percent effect from Duke (=1)	–8.75***	–8.35***	–8.37***	–10.84***	–10.88***	–9.97***
	(2.50)	(2.65)	(2.33)	(2.88)	(2.67)	(2.59)
Observations	2,100	2,100	2,100	2,100	2,100	2,100
B. Black infant deaths or infant mortality rate						
Percent effect from Duke (=1)	–12.84***	–11.40**	–13.69***	–11.88*	–15.47***	–16.37***
	(4.34)	(4.56)	(3.83)	(7.06)	(4.28)	(3.72)
Observations	1,995	1,995	1,995	1,995	1,995	1,995
C. White infant deaths or infant mortality rate						
Percent effect from Duke (=1)	–6.42**	–6.79*	–6.55**	–10.66***	–8.48***	–7.06**
	(2.88)	(3.94)	(2.96)	(3.32)	(2.97)	(2.98)
Observations	1,995	1,995	1,995	1,995	1,995	1,995
P-value for difference by race	0.21	0.45	0.13	0.88	0.14	0.03
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 corrected for bias 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 the five counties that ever have zero race-specific births between 1922 and 1942. P-values reported in the bottom row come from tests for the difference in coefficients presented in panels B and C. They are estimated using 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 2. Effect of Duke support on infant mortality rate: Intensive margin estimates

	Y_{ct}^{R} = Payments	Y_{ct}^{R} = Infant mortality rate	
	Appropriations (1)	Appropriations (2)	Payments (3)
A. Pooled infant mortality rate			
Percent effect from $1 million of Duke support		–4.03***	–6.98***
		(0.49)	(1.32)
Appropriations	0.46***		
	(0.09)		
Observations	2,100	2,100	2,100
B. Black infant mortality rate			
Percent effect from $1 million of Duke support		–5.90***	–9.13***
		(0.69)	(1.92)
Appropriations	0.48***		
	(0.08)		
Observations	1,995	1,995	1,995
C. White infant mortality rate			
Percent effect from $1 million of Duke support		–3.40***	–6.60***
		(0.67)	(1.50)
Appropriations	0.45***		
	(0.09)		
Observations	1,995	1,995	1,995
P-value for difference by race	0.18	0.01	0.21

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Estimates from econometric specifications based on column 6 of Table 1. 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. P-values reported in the bottom row come from tests for the difference in coefficients presented in panels B and C. They are estimated using 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 3. Effect of Duke support on number of doctors or doctors per 100,000 population

	Number of Doctors			Doctors per 100,000 population		
	(1)	(2)	(3)	(4)	(5)	(6)
A. All doctors						
Percent effect from Duke (=1)	20.82***	16.15***	7.32	8.62	8.56*	6.60
	(7.02)	(5.60)	(4.65)	(5.53)	(5.06)	(4.29)
Observations	700	700	700	700	700	700
B. High quality doctors						
Percent effect from Duke (=1)	25.02***	19.93***	11.08**	12.99	14.20**	10.74*
	(7.93)	(6.13)	(5.04)	(7.84)	(6.17)	(5.67)
Observations	700	700	700	700	700	700
C. Low quality doctors						
Percent effect from Duke (=1)	1.84	4.43	–3.37	–11.67**	–8.64**	–9.03**
	(5.18)	(5.24)	(3.84)	(5.02)	(3.60)	(3.79)
Observations	665	665	665	665	665	665
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. The dependent variables are specified for all doctors (panel A), high quality doctors (panel B), and low quality doctors (panel C). In columns 1 to 3, these variables are expressed as counts, while in columns 4 to 6, they are modified to represent the number of doctors per 100,000 population. The sample includes doctors listed in the *American Medical Directory* for the years 1914, 1918, 1923, 1927, 1931, 1936, and 1940. Doctors are considered high quality if they were licensed more than two years after the medical school they attended introduced a two-year degree requirement as an admission pre-requisite. All other doctors are considered low quality. Control variables in columns 3 and 6 are listed in Table 1. Regressions with weights are weighted by total population in county c and year t. Standard errors are estimated using the delta method and are clustered at the county level.

Table 4. Interaction of Duke-support rollout and sulfa-discovery shift-share DiD designs

	Pooled	Black	White
	(1)	(2)	(3)
A. Interaction of Duke rollout × sulfa shift-share DiD			
Post-pre sulfa Pneumonia$_{.75}$ – Post-pre sulfa Pneumonia$_{.25}$ (γ_5)	–14.34*** (5.28)	–12.98* (7.39)	–15.07** (5.84)
B. Duke vs. no Duke			
Post-Sulfa, Pneumonia$_{.75}$ $\gamma_1+\gamma_2+\eta_{.75}\times(\gamma_3+\gamma_5)$	–18.46*** (3.81)	–22.90*** (5.37)	–16.52*** (4.17)
Post-Sulfa, Pneumonia$_{.25}$ $\gamma_1+\gamma_2+\eta_{.25}\times(\gamma_3+\gamma_5)$	3.13 (7.48)	0.64 (11.70)	3.40 (7.69)
Pre-Sulfa, Pneumonia$_{.75}$ $\gamma_1+\eta_{.75}\times\gamma_3$	–7.98*** (2.56)	–12.82*** (3.52)	–5.76* (3.26)
Pre-Sulfa, Pneumonia$_{.25}$ $\gamma_1+\eta_{.25}\times\gamma_3$	–0.29 (3.73)	–0.97 (6.80)	–0.87 (4.15)
C. (Duke vs. no Duke) × (Pre vs. post sulfa) or (Duke vs. no Duke) × (Pneumonia IQR)			
Pneumonia$_{.75}$, Post-pre Sulfa $\gamma_2+\eta_{.75}\times\gamma_5$	–11.40** (4.41)	–11.56** (5.75)	–11.41** (4.91)
Pneumonia$_{.25}$, Post-pre Sulfa $\gamma_2+\eta_{.25}\times\gamma_5$	3.43 (6.75)	1.63 (7.55)	4.31 (8.36)
Post-Sulfa, Pneumonia$_{IQR}$ $(\eta_{.75}-\eta_{.25})\times(\gamma_3+\gamma_5)$	–20.94*** (5.65)	–23.39*** (8.20)	–19.26*** (5.84)
Pre-Sulfa, Pneumonia$_{IQR}$ $(\eta_{.75}-\eta_{.25})\times\gamma_3$	–7.71*** (2.59)	–11.96*** (4.08)	–4.94* (2.93)
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 3. Because we report exponentiated linear combinations of the underlying coefficients, linear combinations of this nonlinear transform may not be equal to the reported coefficient due to Jensen's inequality (i.e., $\exp(a+b) \leqslant \exp(a)+\exp(b)$). 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.

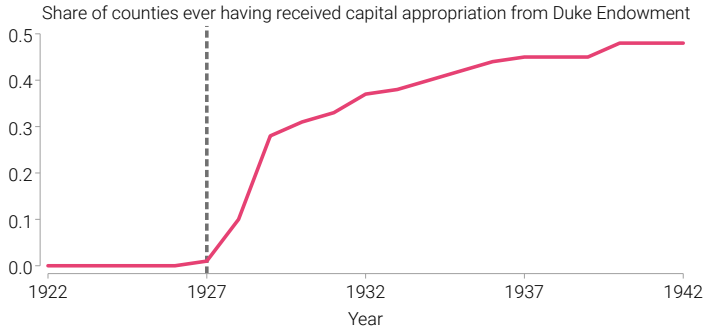
Table 5. Effect of Duke support around time of birth on mortality at ages 56 to 65

	Y_{ct}^{R} = Later-life deaths			Y_{ct}^{R} = Later-life mortality rate		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Pooled later-life deaths or later-life mortality rate						
Percent effect from Duke (=1)	–7.85***	–10.76***	–8.60***	–5.53	–8.31**	–8.81***
	(2.37)	(2.18)	(2.39)	(3.65)	(3.32)	(3.15)
Observations	11,000	11,000	11,000	11,000	11,000	11,000
B. Black later-life deaths or later-life mortality rate						
Percent effect from Duke (=1)	–8.33***	–10.94***	–10.83***	–12.03*	–9.90***	–8.50***
	(2.77)	(3.01)	(2.44)	(6.48)	(2.69)	(2.61)
Observations	9,570	9,570	9,570	9,570	9,570	9,570
C. White later-life deaths or later-life mortality rate						
Percent effect from Duke (=1)	–7.30**	–10.74***	–7.74***	–1.20	–7.57*	–8.65**
	(3.12)	(2.57)	(2.51)	(5.07)	(4.50)	(3.53)
Observations	9,570	9,570	9,570	9,570	9,570	9,570
P-value for difference by race	0.78	0.96	0.30	0.22	0.60	0.96
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

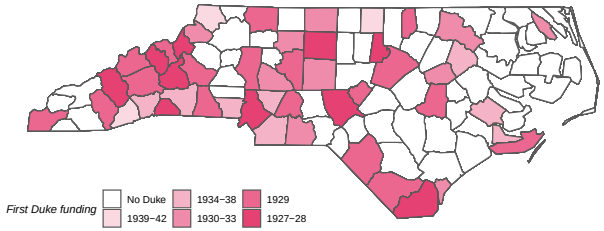
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The unit of observation is a birth county by birth year by follow-up year triplet. Birth cohorts are restricted to 1932 to 1942. Deaths are restricted to ages 56 to 65 and years 1988 to 2007. 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. 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 corrected for bias in the following way: $100 \times (exp(\beta) - 1)$. Control variables include flexible interactions of follow-up year FEs 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 the 13 counties that include cohorts with zero births in any year between 1932 and 1942 or zero surviving population at any age between 56 and 65 from 1988 to 2007. P-values reported in the bottom row come from tests for the difference in coefficients presented in panels B and C. They are estimated using a model that fully interacts all variables with race. Standard errors are estimated using the delta method and are clustered at the county level.

10 Figures

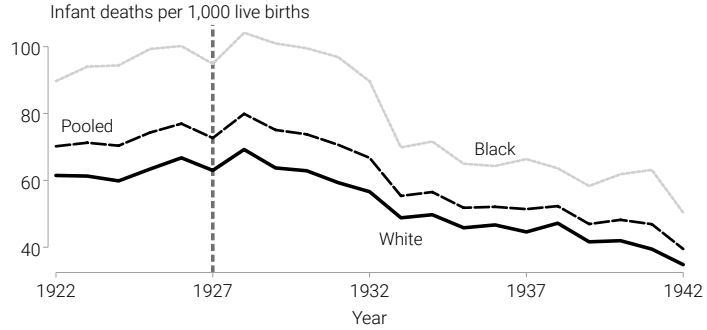
Figure 1. Rollout of Duke support and changes in infant mortality rate, 1922-1942



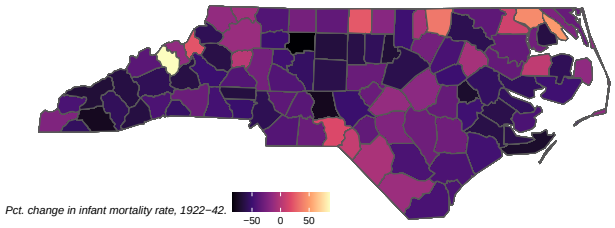
(a) Share of counties treated by a given year



(b) Duration of exposure to Duke support by county



(c) Infant mortality rate by year and race



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

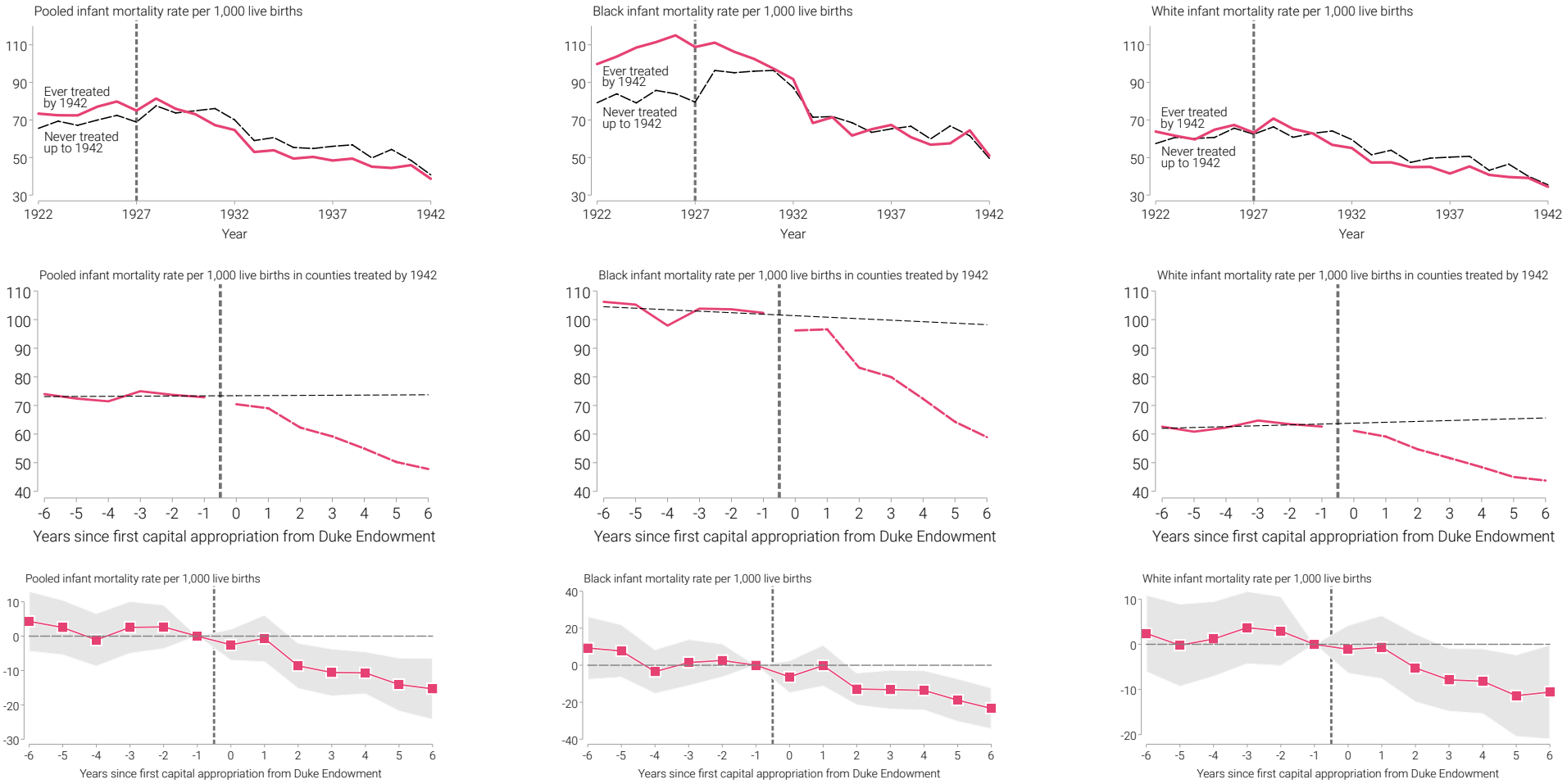
Notes: Figure 1a plots the fraction of ever-treated counties in North Carolina by calendar year. Figure 1b shows a map of county boundaries for the 100 counties in the state of North Carolina. The color gradient provides a visualization of the time variation in the rollout of funding from The Duke Endowment. 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 1c 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. Infant mortality rates by treatment status and event time. Event study estimates.

(a) Pooled infant mortality rate

(b) Black infant mortality rate

(c) White infant mortality rate

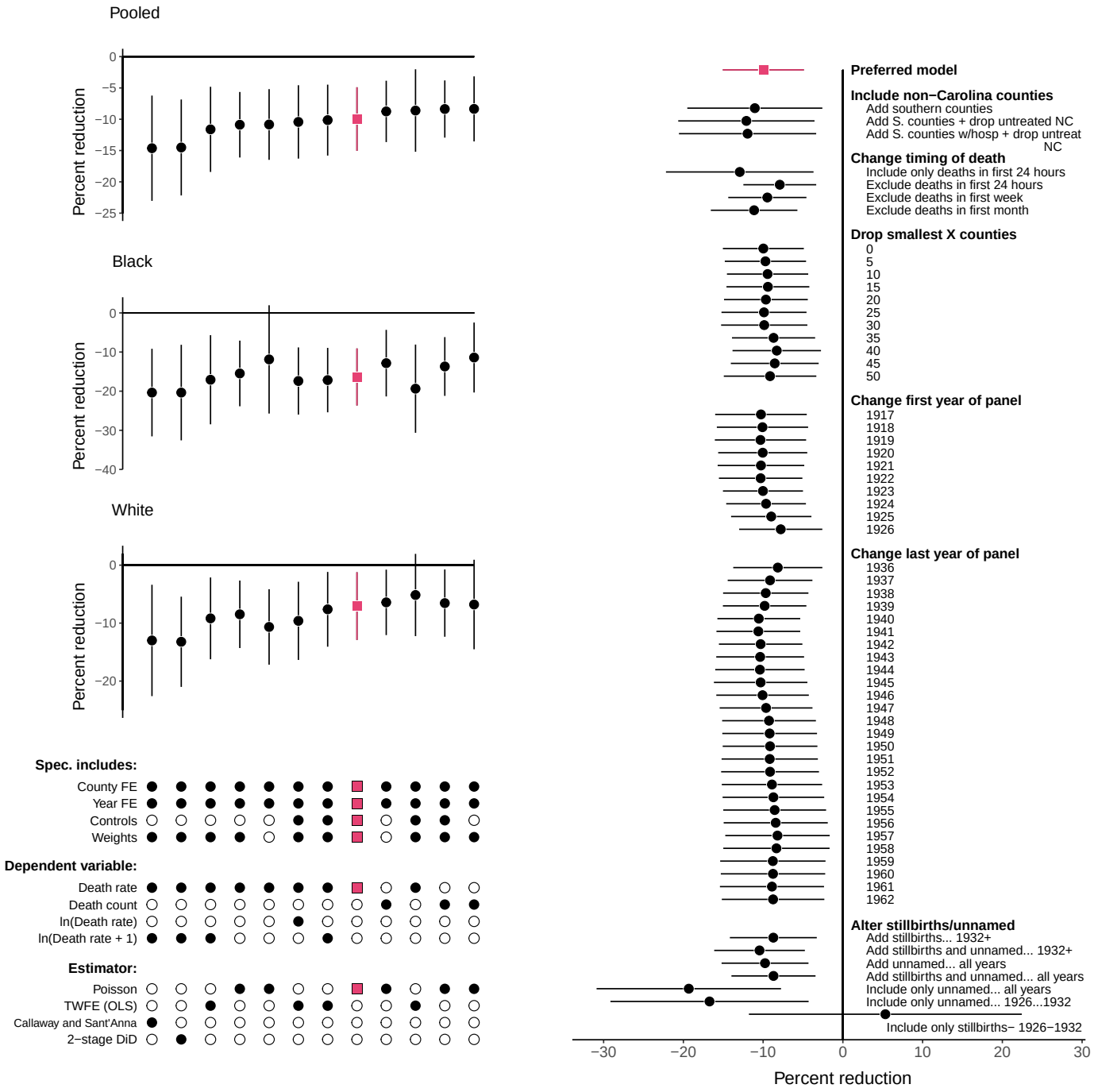


Notes: The top row of plots shows the annual infant mortality rate per 1,000 live births by Duke treatment status for Blacks and Whites pooled together, Blacks only, and Whites only, respectively. Counties "Ever treated by 1942" first received Duke funding during the sample period, between 1927 and 1942, 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 and right-most coefficients, which are binned. Thus, the coefficient on $t = -7$ indicates relative time periods between -18 and -7, while the coefficient on $t = 7$ represents 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 are corrected for bias 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. The bottom row of panels B and C drop five counties that ever had years with zero Black births during the sample period.

Figure 3. A visual representation of key robustness checks

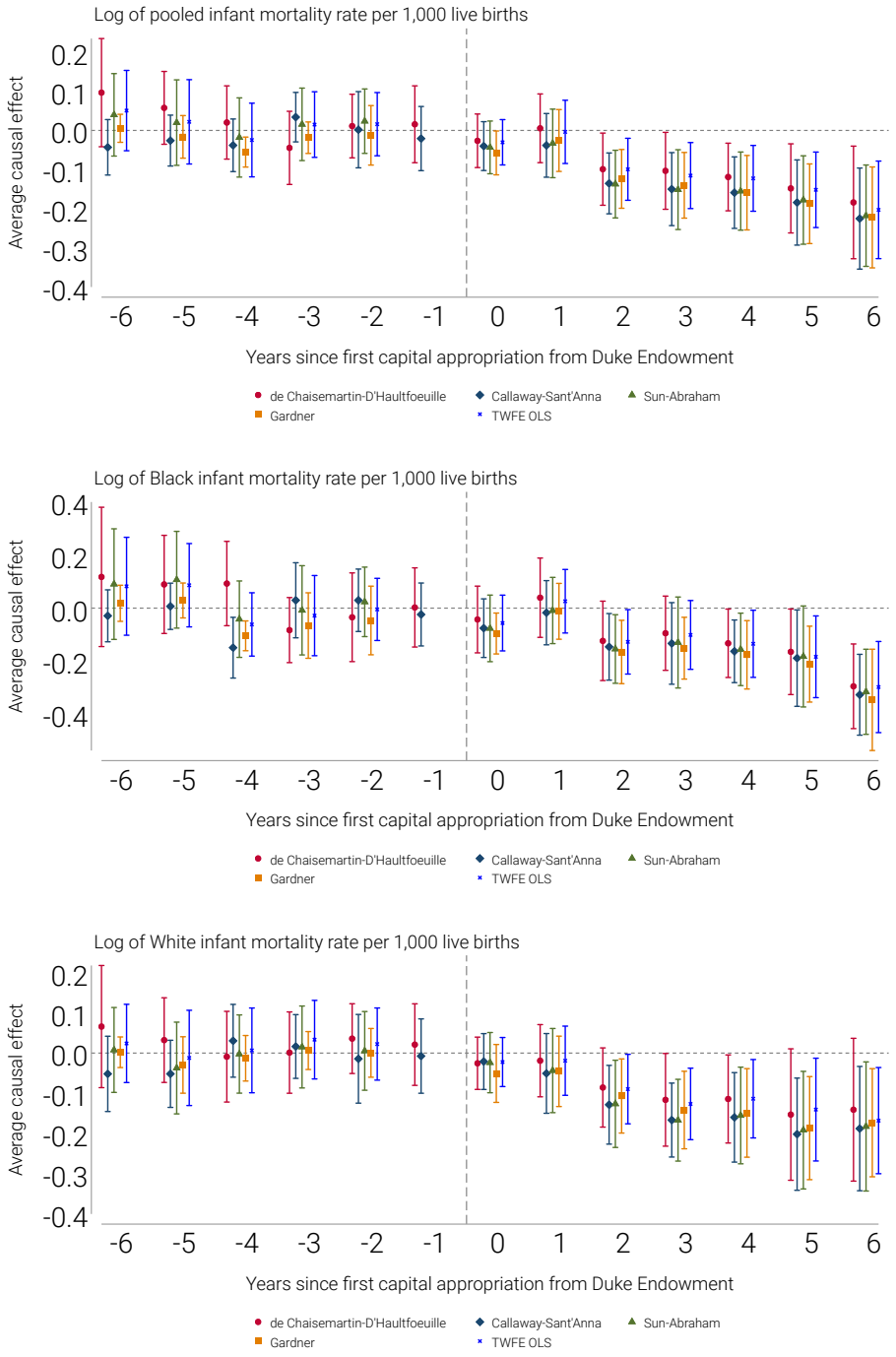
(a) Varying specification, dependent variable, and estimator

(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 1 and Online Appendix Tables F1 and G1 for a description of the specifications plotted in panel A. See Online Appendix Tables H1, C4, and C5 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 1. Standard errors are estimated using the delta method and are clustered at the county level.

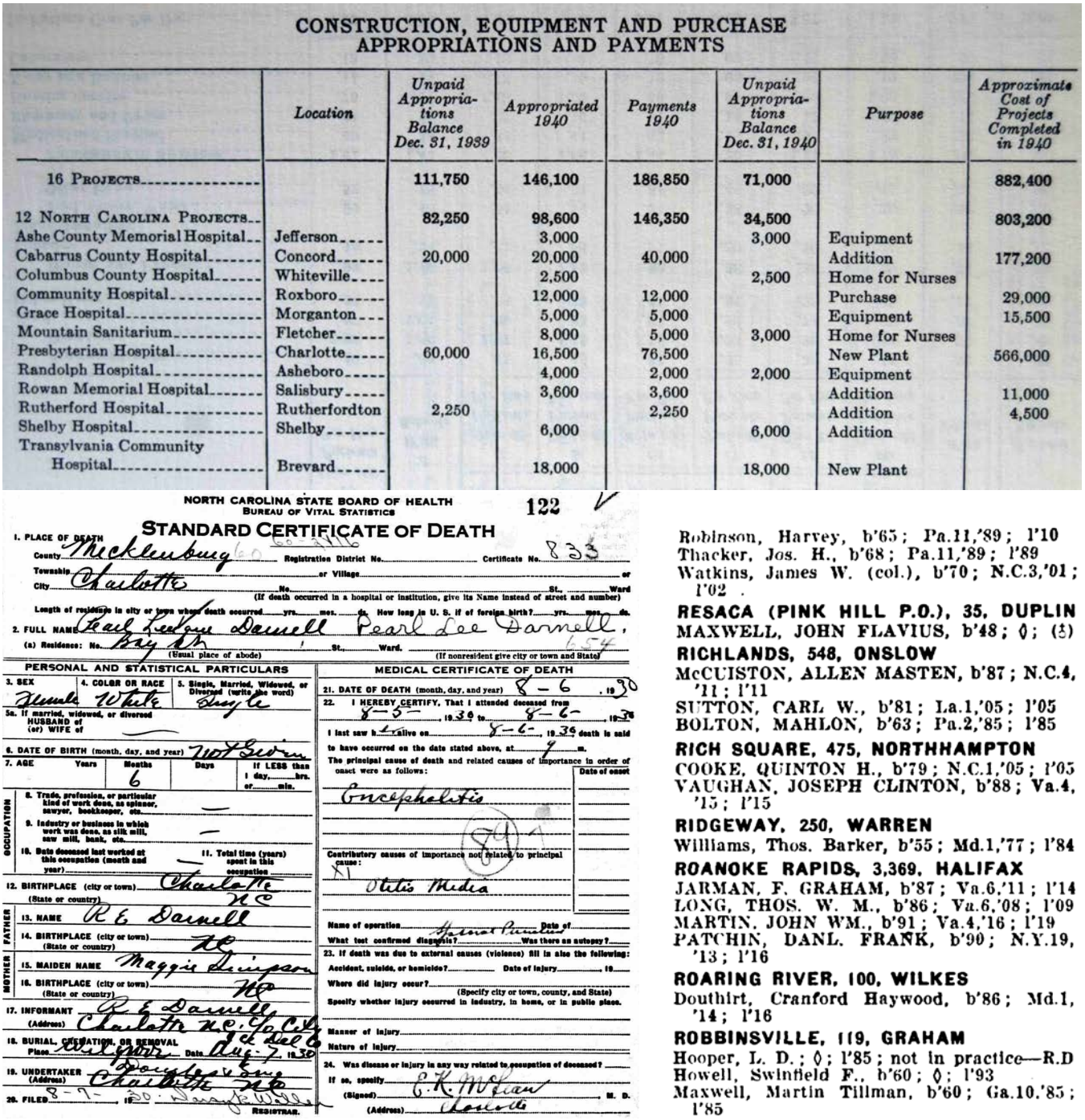
Figure 4. Event studies: Robustness to recent innovations in difference-in-differences literature



Notes: This figure plots event study estimates for the effects of Duke support separately by time period relative to the first year that a county received a capital appropriation from The Duke Endowment. The dependent variable is the natural log of the infant mortality rate per 1,000 live births for Black and White infants pooled together (top), Black infants only (middle), and White infants only (bottom). The sample drops 26 counties from the sample that ever had zero Black or White births or zero Black or White deaths in a year during the sample period. We plot results using the OLS two-way fixed effects estimator (blue X-marks) as well as new estimators developed in de Chaisemartin and D'Haultfœuille (2020) (red circles), Callaway and Sant'Anna (2021) (gray diamonds), Gardner (2021) (orange squares), and Sun and Abraham (2021) (green triangles). The reported coefficients multiplied by 100 approximately represent the percent effect of the Endowment funding. Where appropriate, regressions include county and year-of-birth fixed effects. Observations are weighted by the number of births in each county-by-year-of-birth cohort. Standard errors are clustered by county of birth.

A Sample images of data sources

Figure A1. Sample images



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	

NORTH CAROLINA STATE BOARD OF HEALTH
BUREAU OF VITAL STATISTICS

122

STANDARD CERTIFICATE OF DEATH

1. PLACE OF DEATH — County: Mecklenburg — Registration District No. — Certificate No. 833
Township — or Village — Charlotte — City — No. — St. — Ward
(If death occurred in a hospital or institution, give its Name instead of street and number)
Length of residence in city or town where death occurred — yrs. — mos. — ds. How long in U. S. if of foreign birth? — yrs. — mos. — ds.
2. FULL NAME: Pearl Lee Darnell — Pearl Lee Darnell
(a) Residence: No. Bay St. — St. — Ward. — (If nonresident give city or town and State)
(Usual place of abode)

PERSONAL AND STATISTICAL PARTICULARS

3. SEX	4. COLOR OR RACE	5. Single, Married, Widowed, or Divorced (write the word)
Female	White	Single

5a. If married, widowed, or divorced — HUSBAND of (or) WIFE of
6. DATE OF BIRTH (month, day, and year): Not Given
7. AGE — Years — Months: 6 — Days — If LESS than 1 day, hrs. or min.
OCCUPATION
8. Trade, profession, or particular kind of work done, as spinner, sawyer, bookkeeper, etc.
9. Industry or business in which work was done, as silk mill, saw mill, bank, etc.
10. Date deceased last worked at this occupation (month and year) — 11. Total time (years) spent in this occupation
12. BIRTHPLACE (city or town): Charlotte — (State or country): N.C.
FATHER
13. NAME: R E Darnell
14. BIRTHPLACE (city or town) — (State or country): N.C.
MOTHER
15. MAIDEN NAME: Maggie Simpson
16. BIRTHPLACE (city or town) — (State or country): N.C.
17. INFORMANT: R E Darnell — (Address): Charlotte N.C. Father
18. BURIAL, CREMATION, OR REMOVAL — Place: Elmwood — Date: Aug 7 1930
19. UNDERTAKER: Dougherty & Long — (Address): Charlotte N.C.
20. FILED: 8-9- 1930 — Registrar

MEDICAL CERTIFICATE OF DEATH

21. DATE OF DEATH (month, day, and year): 8-6 1930
22. I HEREBY CERTIFY, That I attended deceased from 8-5- 1930 to 8-6- 1930
I last saw him alive on 8-6- 1930 death is said to have occurred on the date stated above, at — m.
The principal cause of death and related causes of importance in order of onset were as follows:

	Date of onset
Encephalitis	

Contributory causes of importance not related to principal cause:

Otitis Media	

Name of operation — Date of
What test confirmed diagnosis? — Was there an autopsy?
23. If death was due to external causes (violence) fill in also the following:
Accident, suicide, or homicide? — Date of injury — 19
Where did injury occur? — (Specify city or town, county, and State)
Specify whether injury occurred in industry, in home, or in public place.
Manner of injury
Nature of injury
24. Was disease or injury in any way related to occupation of deceased?
If so, specify
(Signed) C. K. McLean M. D.
(Address) Charlotte

Robinson, Harvey, b'65; Pa.11,'89; l'10
Thacker, Jos. H., b'68; Pa.11,'89; l'89
Watkins, James W. (col.), b'70; N.C.3,'01; l'02

RESACA (PINK HILL P.O.), 35, DUPLIN
MAXWELL, JOHN FLAVIUS, b'48; ◊; (1)

RICHLANDS, 548, ONSLOW
McCUISTON, ALLEN MASTEN, b'87; N.C.4, '11; l'11
SUTTON, CARL W., b'81; La.1,'05; l'05
BOLTON, MAHLON, b'63; Pa.2,'85; l'85

RICH SQUARE, 475, NORTHHAMPTON
COOKE, QUINTON H., b'79; N.C.1,'05; l'05
VAUGHAN, JOSEPH CLINTON, b'88; Va.4, '15; l'15

RIDGEWAY, 250, WARREN
Williams, Thos. Barker, b'55; Md.1,'77; l'84

ROANOKE RAPIDS, 3,369, HALIFAX
JARMAN, F. GRAHAM, b'87; Va.6,'11; l'14
LONG, THOS. W. M., b'86; Va.6,'08; l'09
MARTIN, JOHN WM., b'91; Va.4,'16; l'19
PATCHIN, DANL. FRANK, b'90; N.Y.19, '13; l'16

ROARING RIVER, 100, WILKES
Douthirt, Cranford Haywood, b'86; Md.1, '14; l'16

ROBBINSVILLE, 119, GRAHAM
Hooper, L. D.; ◊; l'85; not in practice—R.D
Howell, Swinfield F., b'60; ◊; l'93
Maxwell, Martin Tillman, b'60; Ga.10,'85; l'85

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); a North Carolina Certificate of Death, the source of our infant mortality measure by race and county of birth (bottom left); and the *American Medical Directory*, the source for the number of physicians by county and year (bottom right).

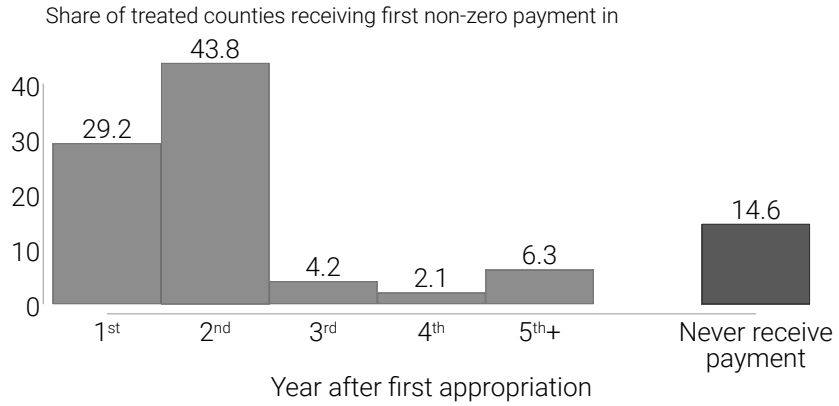
B Heterogeneity by project type

Table B1. Appropriation and payment details by project type.

	Mean	S.D.	Min.	Max.	N
All projects					
Appropriations, millions	0.47	0.52	0.00	4.31	271
Payments, millions	0.40	0.53	0.00	4.39	164
New hospitals or plants					
Appropriations, millions	0.75	0.87	0.04	2.59	15
Payments, millions	0.85	0.75	0.12	2.16	9
Additions					
Appropriations, millions	0.59	0.52	0.01	4.31	131
Payments, millions	0.53	0.63	0.01	4.39	69
Equipment					
Appropriations, millions	0.15	0.39	0.00	1.69	34
Payments, millions	0.09	0.20	0.00	1.13	31
Purchases of existing facilities					
Appropriations, millions	0.64	0.32	0.12	1.79	26
Payments, millions	0.56	0.20	0.21	0.97	14
Not stated					
Appropriations, millions	0.23	0.36	0.01	1.89	68
Payments, millions	0.25	0.38	0.01	1.89	44

Notes: Summary statistics for non-zero appropriations and payments from The Duke Endowment in millions of 2017 dollars. The sample includes all appropriations and payments for hospitals in North Carolina between 1927 and 1942, excluding support for homes for nurses. Some appropriations or payments may apply to more than one project type.

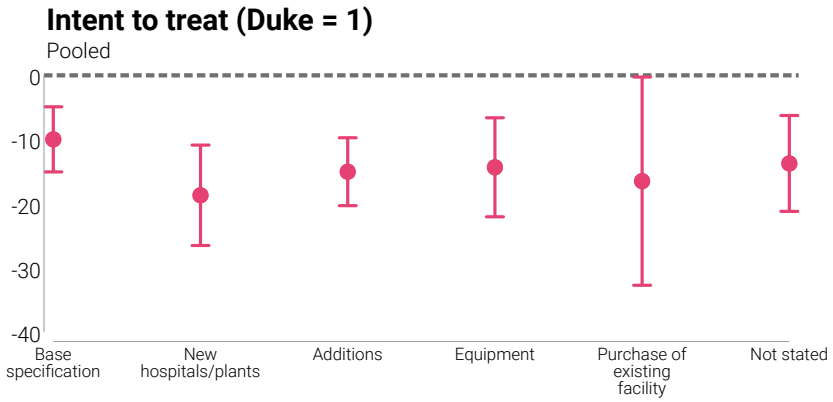
Figure B1. Time from appropriation to payment



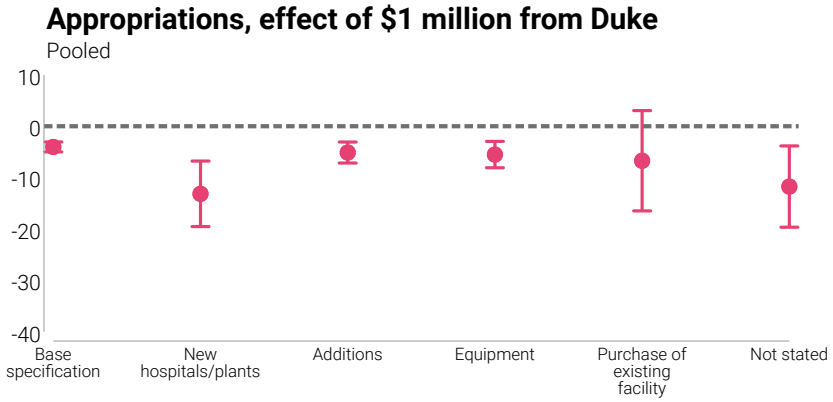
Notes: This figure plots the share of ever-treated counties by event-time, i.e. the number of years after the first appropriation when the first payment was received by the county.

Figure B2. Differential effects of Duke support by project type: Extensive and intensive margin estimates

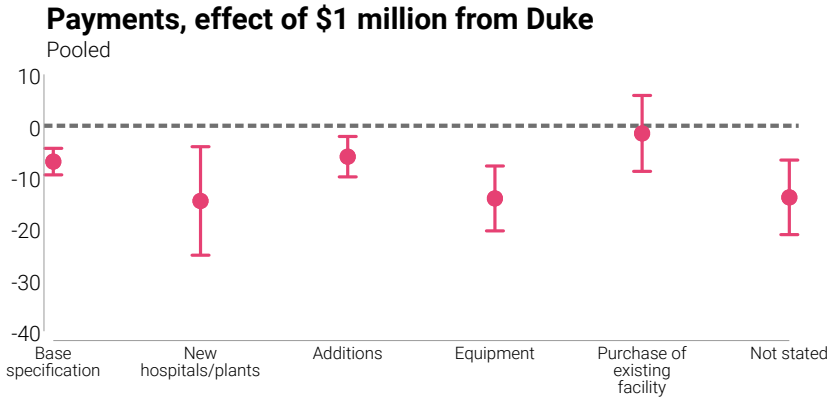
(a) Intent to treat (binary)



(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 and excludes funding for other 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 are corrected for bias 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.

C 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 C5, 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 both the numerator and denominator 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 3).

Shift-share DiD: In Equation 5, we estimate the effects of sulfa drugs on infant mortality in our main estimation sample using a within-North-Carolina shift-share design:[36]

$$Y^{R}_{ct} = exp(\delta_0 + \delta_1 \text{Pneumonia mortality}_c \times \text{Post sulfa}_t + \zeta_c + \eta_t + \mathbf{\Theta X_{ct}})\epsilon_{ct} \tag{5}$$

where Y^{R}_{ct}, ζ_c, η_t, and $\mathbf{X_{ct}}$ are defined as in Equation 1. 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).[37] Since 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).[38] We cluster the standard errors (ε_{ct}) by county of birth to account for correlated errors within a county.

[36] This 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.

[37] We 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.

[38] We 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 C6 presents the results of estimating Equation 5 for pooled infant mortality (panel A), Black infant mortality (panel B) and White infant mortality (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 between 6.5% and 10.5%. Similarly to the results in Table 1, estimates by race are larger for Black compared with White infants and this difference is statistically significant at conventional levels. 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 1.

Table C1. Balancing test: Effects of Duke support on control variables

	(1) Mean % difference	(2) p-value
% Illiterate	–3.97	0.20
% Black	–0.89	0.26
% Urban	–3.63	0.17
Retail sales per capita	0.14	0.96
County health department open	2.78	0.78

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. This analysis follows Pei et al. (2019), where a % difference of -10 would indicate that Duke support is associated with a change in outcome of 10% in treated versus control counties. Coefficients corrected for bias in the following way: $100 \times (exp(\beta) - 1)$. Standard errors are estimated using the delta method and are clustered at the county-level.

Table C2. Extensive margin intent-to-treat effect of Duke support on maternal mortality

	Number of Deaths			Death rate per 100,000		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Maternal mortality						
Percent effect from Duke (=1)	–14.11	–17.93**	–19.98**	1.69	–7.79	–9.66
	(8.91)	(8.76)	(8.64)	(15.54)	(8.87)	(7.90)
Observations	1,100	1,100	1,100	1,100	1,100	1,100
B. Maternal mortality, resident						
Percent effect from Duke (=1)	–9.53	–11.26	–13.55**	1.16	–4.50	–5.76
	(7.47)	(6.87)	(6.75)	(14.12)	(7.36)	(6.51)
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 corrected for bias 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 C3. 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)
A. Pooled births						
Percent effect from Duke (=1)	1.68	–0.40	–1.86	–0.53	–2.97*	–2.45
	(1.93)	(2.13)	(2.08)	(1.55)	(1.52)	(1.53)
Observations	2,100	2,100	2,100	2,100	2,100	2,100
B. Black births						
Percent effect from Duke (=1)	2.74	2.25	1.46	2.74	–1.23	–0.50
	(2.79)	(2.92)	(2.61)	(2.79)	(2.23)	(1.97)
Observations	2,100	2,100	2,100	2,100	2,100	2,100
C. White births						
Percent effect from Duke (=1)	1.41	–2.04	–3.19	1.68	–2.69	–2.26
	(2.22)	(2.42)	(2.42)	(1.87)	(1.69)	(1.74)
Observations	2,100	2,100	2,100	2,100	2,100	2,100
P-value for difference by race	0.54	0.13	0.13	0.15	0.41	0.38
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 corrected for bias 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. 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 C4. Effect of Duke support on infant mortality rate by timing of death

	Deaths on:	Exlcuding those who died in the first:		
	day 0	day	week	month
	(1)	(2)	(3)	(4)
A. Pooled infant mortality rate				
Percent effect from Duke (=1)	–12.93***	–7.91***	–9.47***	–11.13***
	(4.72)	(2.33)	(2.50)	(2.77)
Observations	2,100	2,100	2,100	2,100
B. Black infant mortality rate				
Percent effect from Duke (=1)	–22.51***	–13.60***	–15.12***	–17.26***
	(5.67)	(3.68)	(3.87)	(3.80)
Observations	1,974	1,995	1,995	1,995
C. White infant mortality rate				
Percent effect from Duke (=1)	–8.95	–5.19*	–6.13*	–7.25**
	(5.73)	(2.89)	(3.24)	(3.50)
Observations	1,974	1,995	1,995	1,995
P-value for difference by race	0.05	0.07	0.08	0.05
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 1. 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 corrected for bias 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 C5. 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)
A. Pooled infant mortality rate							
Percent effect from Duke (=1)	–8.72***	–10.45***	–9.76***	–8.70***	–19.32***	–16.72***	5.33
	(2.78)	(2.90)	(2.78)	(2.69)	(5.90)	(6.34)	(8.73)
Observations	2,100	2,100	2,100	2,100	2,100	1,100	1,100
B. Black infant mortality rate							
Percent effect from Duke (=1)	–14.37***	–14.90***	–13.27***	–12.98***	–19.86**	–17.73*	–0.77
	(4.09)	(4.07)	(3.81)	(3.81)	(9.15)	(10.18)	(10.18)
Observations	2,058	2,058	2,058	2,058	2,037	1,078	1,056
C. White infant mortality rate							
Percent effect from Duke (=1)	–6.88**	–6.89**	–5.91**	–6.09**	–18.86***	–16.45**	11.04
	(2.67)	(2.95)	(2.88)	(2.63)	(6.27)	(7.12)	(13.77)
Observations	2,058	2,058	2,058	2,058	2,058	1,089	1,056
P-value for difference by race	0.05	0.05	0.05	0.05	0.92	0.91	0.44
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 1 by adding stillbirths to the numerator (death count) and denominator (live plus stillbirths). 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 1 for a description of the specifications. Standard errors are estimated using the delta method and are clustered at the county-level.

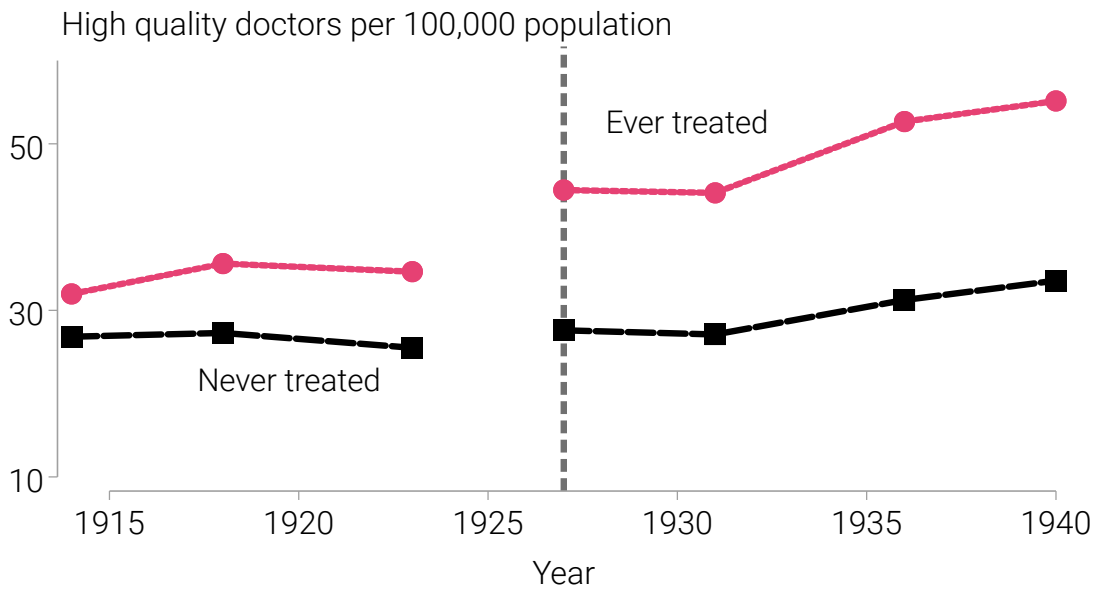
Table C6. Effects of sulfa drugs on infant mortality: Shift-share difference-in-differences

	(1)	(2)	(3)
A. Pooled infant mortality rate			
Percent effect from $\text{Pneumonia}_{\text{IQR}} \times$ Post sulfa	–5.20**	–9.56***	–8.32**
	(2.59)	(3.12)	(3.40)
Observations	2,100	2,100	2,100
B. Black infant mortality rate			
Percent effect from $\text{Pneumonia}_{\text{IQR}} \times$ Post sulfa	–21.86***	–12.08**	–10.49**
	(6.61)	(4.64)	(4.93)
Observations	1,995	1,995	1,995
C. White infant mortality rate			
Percent effect from $\text{Pneumonia}_{\text{IQR}} \times$ Post sulfa	–3.51	–7.68**	–6.46*
	(3.45)	(3.28)	(3.55)
Observations	1,995	1,995	1,995
P-value for difference by race	0.88	0.14	0.03
County of birth FE	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes
Controls	No	No	Yes
Weights	No	Yes	Yes

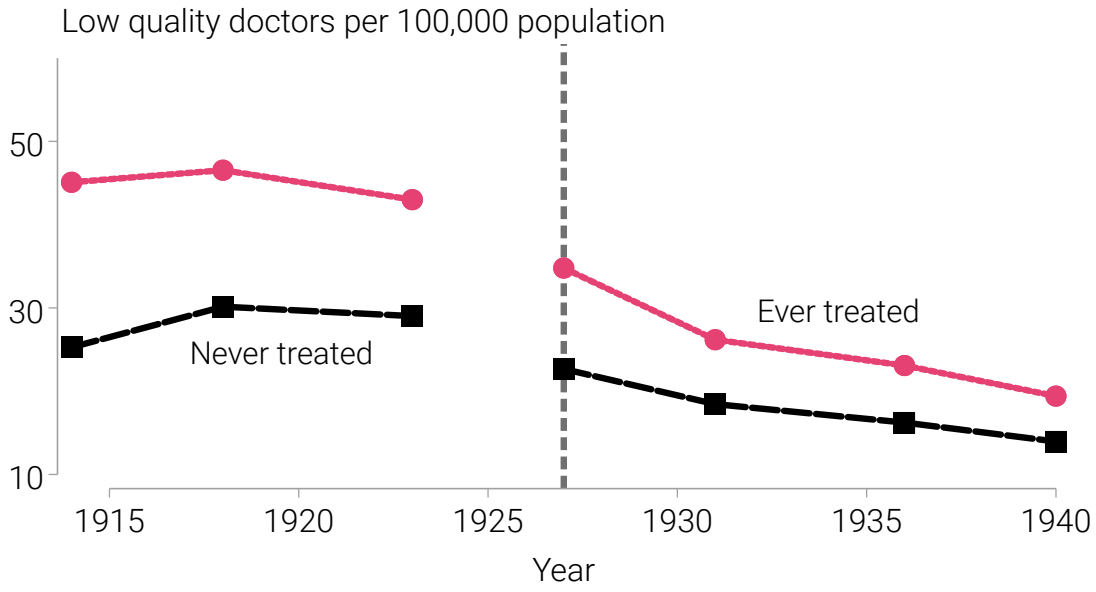
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The dependent variable is the infant mortality rate per 1,000 live births. Years 1937 and onward are considered post sulfa (shift factor). The baseline shares are county-level 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 5 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 corrected for bias 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 the five 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.

Figure C1. Physician labor supply across time and treatment status

(a) High quality doctors

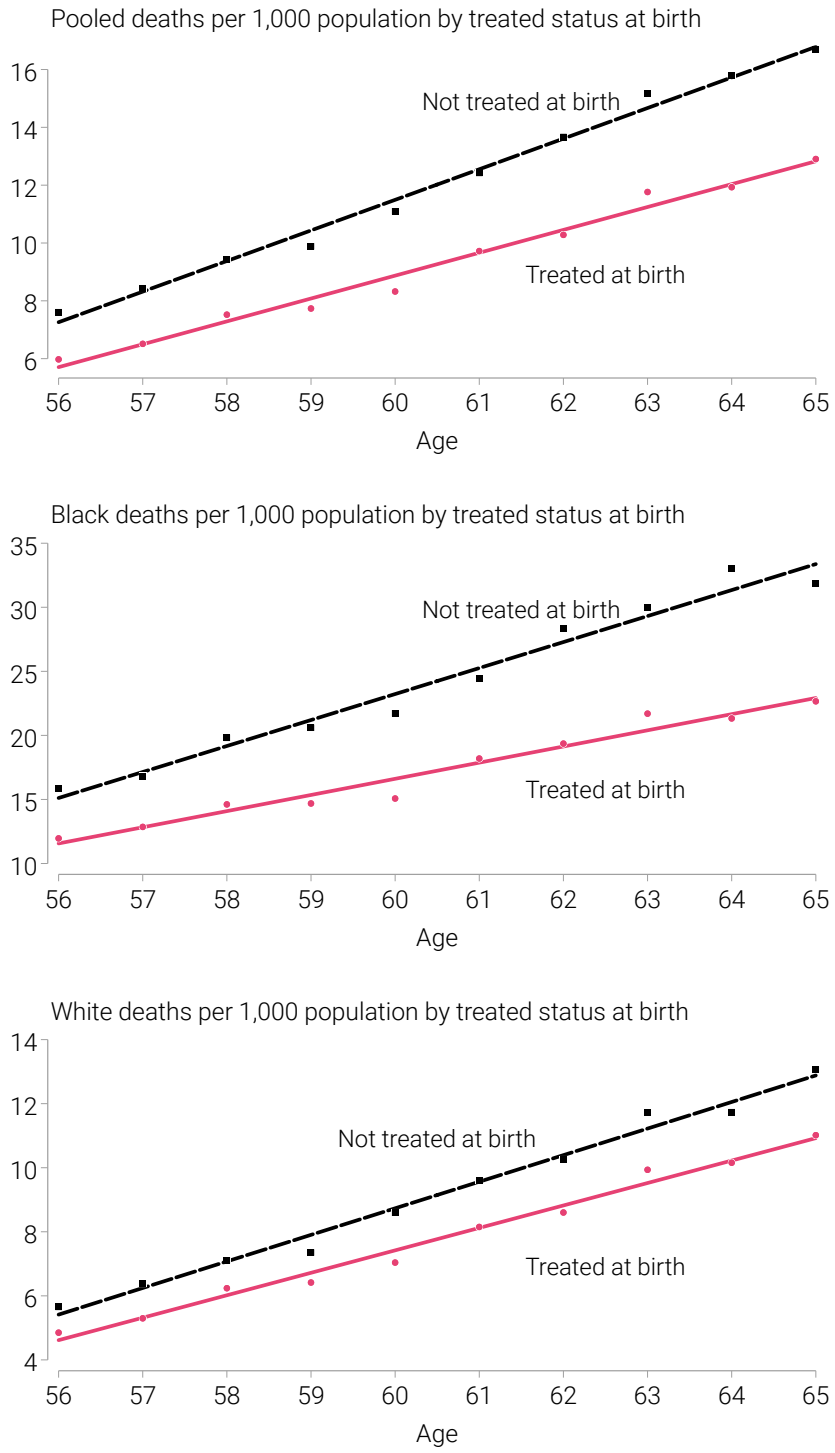


(b) Low quality doctors



Notes: This figure plots the average number of high-quality (panel A) and low-quality (panel B) doctors per 100,000 population listed in the *American Medical Directory* for the years 1914, 1918, 1923, 1927, 1931, 1936, and 1940. Doctor are considered high quality if they were licensed more than two years after the medical school they attended introduced a two-year degree requirement as an admission pre-requisite. All other doctors are considered low quality. "Ever treated" counties (pink solid line) first received Duke funding between 1927 and 1942 while "Never treated" counties (black dashed line) did not.

Figure C2. Later-life mortality rate by treated status and age



Notes: This figure shows a scatter plot of the average number of deaths per 1,000 population at ages 56 to 65 in a county-by-year-of-birth cohort and a linear regression line through these points, separately by treatment status at birth. Cohorts are treated at birth (pink solid line) if the first capital appropriation received by the county of birth occurred no later than the year of birth. All other cohorts in the sample are considered untreated (dashed black line). Data include the 1932 to 1942 birth cohorts and deaths that occurred between 1988 and 2007. Data are tabulated by the authors from the public-use version of the 2007 Numident file from the National Archives and Records Administration.

D Panel construction and sample size

In our main specification for short-run outcomes in panel A of Table 1, 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 11 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 18 observations with zero Black births. In our main specifications for infant mortality rates by race (panels B and C of Table 1), 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 4, Online Appendix Table F2, panels B and C in column 4 of and Online Appendix Table G1, and Online Appendix Figure F1.

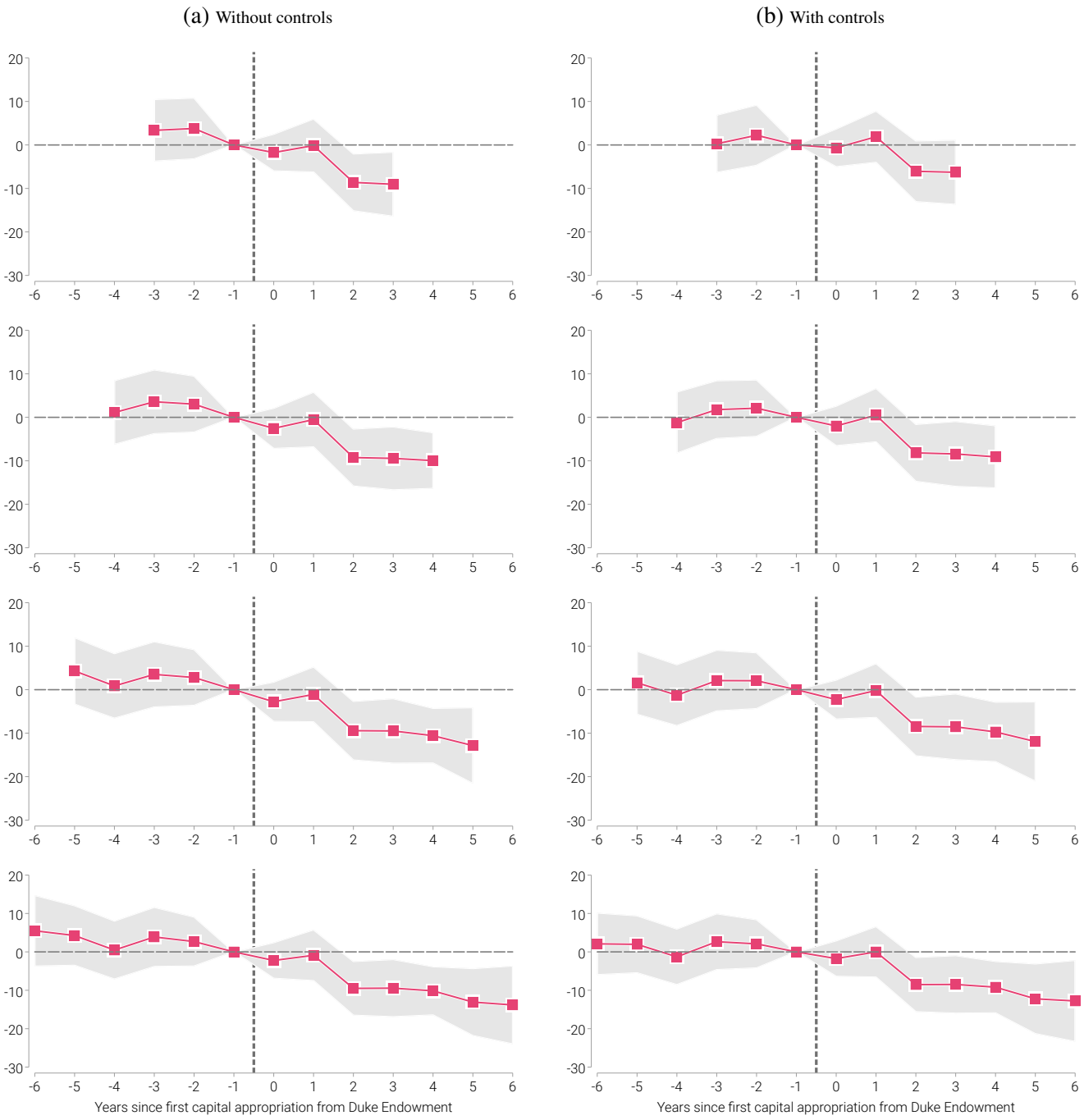
E Stacked regressions

Table E1. Effects of Duke support: Stacked Poisson and OLS infant mortality rate estimates

	Poisson IMR			OLS IMR	OLS ln(IMR)	OLS ln(IMR + 1)
	(1)	(2)	(3)	(4)	(5)	(6)
A. Pooled infant mortality rate						
Percent effect from Duke (=1)	–8.50***	–9.52***	–7.16***	–9.42***	–8.59***	–8.44***
	(2.48)	(2.61)	(2.35)	(3.44)	(2.71)	(2.67)
Observations	7,370	7,370	7,370	7,370	7,370	7,370
B. Black infant mortality rate						
Percent effect from Duke (=1)	–10.27***	–13.01***	–8.79***	–15.70***	–10.61**	–10.43**
	(3.68)	(3.58)	(2.96)	(5.38)	(4.05)	(3.99)
Observations	6,963	6,963	6,963	6,963	6,692	6,963
C. White infant mortality rate						
Percent effect from Duke (=1)	–7.29**	–7.70**	–6.05**	–7.22*	–7.51**	–7.51**
	(3.16)	(3.13)	(2.84)	(3.76)	(3.26)	(3.21)
Observations	6,963	6,963	6,963	6,963	6,945	6,963
P-value for difference by race	0.53	0.22	0.46	0.05	0.53	0.55
County of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	No	No	No
Weights	No	Yes	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 stacked regression. Each stack includes treated counties from a single treatment timing group and control counties that are not treated within the event time window of $\kappa = \pm 5$. The sample includes 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. In columns 1 to 4, the dependent variable is the infant mortality rate per 1,000 live births. In columns 5 and 6, it is the natural log of the infant mortality rate and the natural log of one plus the infant mortality rate, respectively. Each coefficient represents the percent reduction in infant mortality from receiving a capital appropriation from The Duke Endowment. A coefficient of -10 would mean that infant mortality rate declined by 10%. We present the results in this manner so that results from specifications using natural log transformation of infant mortality rate are comparable to those that simply use the infant mortality rate. Coefficients from specifications with Poisson maximum likelihood and a natural log transform of the dependent variable are corrected for bias in the following way: $100 \times (exp(\beta) - 1)$. Coefficients from specifications with a level dependent variable are transformed in percent effects by dividing the coefficient by the mean for untreated observations: $100 \times (\beta / \bar{Y}_U)$. Control variables in column 3 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. Panels B to C drop counties that ever have zero births. Column 5 additionally drops counties that ever have zero deaths. P-values reported in the bottom row come from tests for the difference in coefficients presented in panels B and C. They are estimated using a model that fully interacts all variables with race. Standard errors are estimated using the delta method and are clustered at the county-level.

Figure E1. 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. See Table E1 for a description of the stacked regression specifications. 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.

F Event study diagnostics

Table F1. Average treatment effect estimates from alternative difference-in-differences estimators

	(1) TWFE OLS	(2) Callaway and Sant'Anna (2020)	(3) 2-stage DiD Gardner (2021)
A. Natural log of pooled infant mortality rate + 1			
Effect from Duke (=1)	–11.62***	–14.63***	–14.51***
	(3.47)	(4.29)	(3.91)
Observations	2,100	2,100	2,100
B. Natural log of Black infant mortality rate + 1			
Effect from Duke (=1)	–17.08***	–20.36***	–20.36***
	(5.81)	(5.71)	(6.23)
Observations	1,995	1,995	1,995
C. Natural log of White infant mortality rate + 1			
Effect from Duke (=1)	–9.18**	–12.99***	–13.22***
	(3.60)	(4.90)	(3.97)
Observations	1,995	1,995	1,995

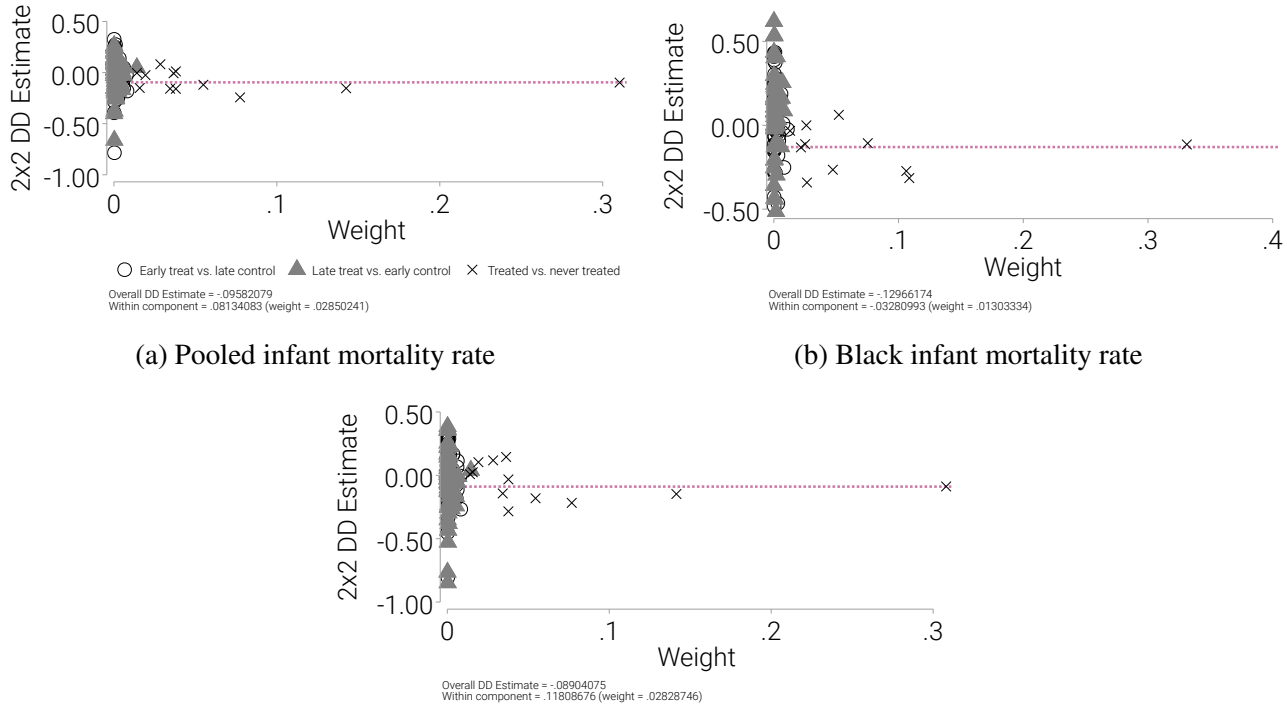
Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Extensive margin intent-to-treat estimates. The dependent variable is the natural log of one plus the infant mortality rate per 1,000 live births. In column 1 we report results from a two-way fixed effect regression estimated by OLS. Column 2 reports the ATT across all groups and periods from Callaway and Sant'Anna (2021). Column 3 reports the two-stage DiD estimator from Gardner (2021). In the first stage we regress the dependent variable on county and year-of-birth fixed effects, and in the second stage we regress the residuals from the first stage on the Duke support indicator. Panels B and C drop the five counties that ever have zero race-specific births between 1922 and 1942. We multiple the coefficients by 100 to approximately represent the percent change in infant mortality rate due to receiving a capital appropriation from The Duke Endowment for the relevant comparison group. Regressions include fixed effects for county and year of birth but no other controls. Estimates are weighted by the number of births in a county and year. Standard errors are clustered at the county-level

Table F2. Goodman-Bacon (2021a) decomposition diagnostic

Type of DD comparison	Average Estimate	Number of 2x2 Comparisons	Total Weight
A. Pooled infant mortality rate			
Earlier Treated vs. Later Treated Controls	–0.024	66	0.095
Later Treated vs. Earlier Treated Controls	–0.042	66	0.095
Treated vs. Untreated Controls	–0.111	12	0.810
Average DD estimate	–0.096	144	1.000
B. Black infant mortality rate			
Earlier Treated vs. Later Treated Controls	0.067	55	0.080
Later Treated vs. Earlier Treated Controls	–0.019	55	0.087
Treated vs. Untreated Controls	–0.160	11	0.833
Average DD estimate	–0.130	121	1.000
C. White infant mortality rate			
Earlier Treated vs. Later Treated Controls	–0.073	66	0.098
Later Treated vs. Earlier Treated Controls	–0.018	66	0.098
Treated vs. Untreated Controls	–0.100	12	0.804
Average DD estimate	–0.089	144	1.000

Notes: The table decomposes the static DiD two-way fixed effects estimate corresponding to the OLS event study plotted in Figure 4 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 F1. Goodman-Bacon (2021a) decomposition diagnostic



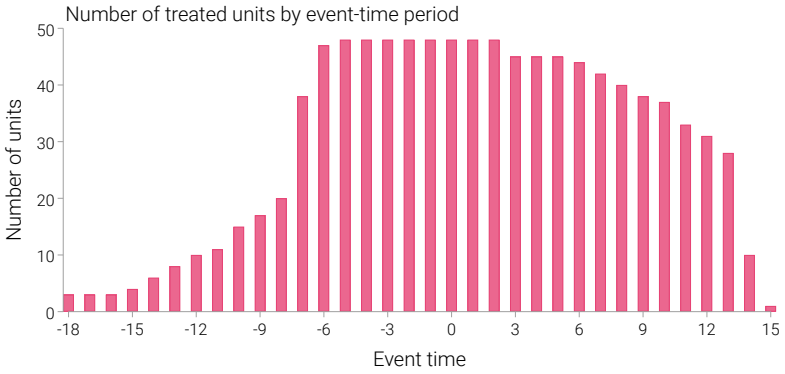
(a) Pooled infant mortality rate

(b) Black infant mortality rate

(c) White infant mortality rate

Notes: Figure F1 decomposes the static DiD two-way fixed effects estimate corresponding to the OLS event study plotted in Figure 4 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 F2. 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 that could be included in the event studies plotted in the bottom row of Figure 2. Figure 2 does not display coefficients binned at $t = -7$ and $t = 7$.

G OLS specifications as alternative to Poisson

Table G1. Effect of Duke support on infant mortality rate: Robustness to alternate OLS specifications

	(1) Poisson IMR	(2) OLS IMR	(3) OLS ln(IMR)	(4) OLS ln(IMR + 1)
A. Pooled infant mortality rate				
Percent effect from Duke (=1)	–9.97***	–9.97***	–8.61**	–10.43***
	(2.59)	(2.59)	(3.36)	(2.99)
Observations	2,100	2,100	2,100	2,100
B. Black infant mortality rate				
Percent effect from Duke (=1)	–16.37***	–19.36***	–17.44***	–17.17***
	(3.72)	(5.75)	(4.28)	(4.20)
Observations	1,995	1,995	1,894	1,995
C. White infant mortality rate				
Percent effect from Duke (=1)	–7.06**	–5.15	–7.72**	–7.62**
	(2.98)	(3.62)	(3.37)	(3.29)
Observations	1,995	1,995	1,991	1,995

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Extensive margin intent-to-treat estimates. Each coefficient comes from a separate regression. Column 1 reproduces estimates from column 6 of Table 1. Columns 2 to 4 report estimates from OLS regressions with the following dependent variables: the number of infant deaths in a county and year, the natural log of the infant mortality rate per 1,000 live births (dropping counties that ever had years with zero infant deaths), and the natural log of one plus the infant mortality rate, 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%. Across panels A to C in column 2, coefficients are scaled by the mean infant mortality rate – pooled, Black only, and White only, respectively – in always-untreated counties. Coefficients in columns 1, 3, and 4 are corrected for bias in the following way: $100 \times (\exp(\beta) - 1)$. Control variables include county and year fixed effects, % 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. Panels B and C drop the five counties that ever had zero births between 1922 and 1942. P-values reported in the bottom row come from tests for the difference in coefficients presented in panels B and C. They are estimated using a model that fully interacts all variables with race. Standard errors in Poisson regressions are estimated using the delta method. All standard errors are clustered at the county-level.

Table G2. Robustness of physician labor supply estimates to alternate OLS specifications

	Y_{ct} = Doctors per 100,000			
	(1) Poisson Y_{ct}	(2) OLS Y_{ct}	(3) OLS $\ln(Y_{ct})$	(4) OLS $\ln(Y_{ct} + 1)$
A. Number of doctors per 100,000 population				
Percent effect from Duke (=1)	6.60	2.83	7.11	6.94*
	(4.29)	(4.36)	(4.30)	(4.17)
Observations	700	700	700	700
B. Number of high quality doctors per 100,000 population				
Percent effect from Duke (=1)	10.74*	19.77***	13.51**	10.49
	(5.67)	(6.27)	(6.77)	(6.42)
Observations	700	700	665	700
C. Number of low quality doctors per 100,000 population				
Percent effect from Duke (=1)	–9.03**	–18.11***	–6.64	–13.14**
	(3.79)	(4.71)	(5.71)	(5.37)
Observations	665	700	588	700
County of birth FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression. The dependent variables are the number of doctors (panel A), high quality doctors (panel B), and low quality doctors (panel C), respectively, per 100,000 population. The sample includes the number of doctors listed in the *American Medical Directory* for the years 1914, 1918, 1923, 1927, 1931, 1936, and 1940. Doctor are considered high quality if they were licensed more than two years after the medical school they attended introduced a two-year degree requirement as an admission pre-requisite. All other doctors are considered low quality. Control variables are listed in Table 1. Regressions are weighted by total population in county c and year t. Column 1 reproduces estimates from column 6 of Table 3. Column 2 replaces Poisson pseudo maximum likelihood estimation with OLS and adds 35 observations with zero low quality doctors in panel C. Column 3 uses the natural log of the number of doctors (dropping counties that ever had zero doctors of specific quality). Column 4 uses the natural log of the number of doctors plus one. Standard errors in Poisson regressions are estimated using the delta method. All standard errors are clustered at the county-level.

Table G3. Robustness of later-life mortality results to alternate OLS specifications

	(1) Poisson LLMR	(2) Poisson LLMR	(3) OLS LLMR	(4) OLS ln(LLMR)	(5) OLS ln(LLMR + 1)
A. Pooled later-life mortality rate					
Percent effect from Duke (=1)	–8.81***	–9.00***	–5.64	–8.53***	–6.51
	(3.15)	(3.33)	(3.61)	(3.12)	(5.60)
Observations	11,000	10,787	11,000	10,355	11,000
B. Black later-life mortality rate					
Percent effect from Duke (=1)	–8.50***	–5.94**	–5.99**	–11.03***	–8.29
	(2.61)	(2.47)	(3.01)	(3.87)	(13.43)
Observations	9,570	9,073	9,570	7,166	9,570
C. White later-life mortality rate					
Percent effect from Duke (=1)	–8.65**	–10.35***	–7.43*	–7.74**	0.82
	(3.53)	(3.82)	(4.37)	(3.22)	(9.20)
Observations	9,570	9,329	9,570	8,291	9,570
P-value for difference by race	0.96	0.24	0.27	0.44	0.61
County of birth FE	Yes	Yes	Yes	Yes	Yes
Year of birth FE	Yes	Yes	Yes	Yes	Yes
Weights	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
County and year of birth FEs X follow-up year FEs	No	Yes	No	No	No

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Each coefficient comes from a separate regression estimated by Poisson pseudo maximum likelihood (columns 1 to 2) or OLS (columns 3 to 6). The unit of observation is a birth county by birth year by follow-up year triplet. Birth cohorts are restricted to 1932 to 1942. Deaths are restricted to ages 56 to 65 and years 1988 to 2007. 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%. Across panels A to C in column 3, coefficients are scaled by the mean later-life mortality rate – pooled, Black only, and White only, respectively – in always-untreated counties. Coefficients in columns 1, 2, 4, and 5 are corrected for bias in the following way: $100 \times (\exp(\beta) - 1)$. Control variables include flexible interactions of follow-up year FEs with % illiterate, % Black, % other race, % urban, retail sales per capita, and presence of a county health department. The weights are county-by-year birth cohort size. Panels B and C drop the 13 counties that include cohorts with zero births in any year between 1932 and 1942 or zero surviving population at any age between 56 and 65 from 1988 to 2007. P-values reported in the bottom row come from tests for the difference in coefficients presented in panels B and C. They are estimated using a model that fully interacts all variables with race. Standard errors in Poisson regressions are estimated using the delta method. All standard errors are clustered at the county-level.

H Alternate samples and non-Carolina control counties

Table H1. Robustness to using non-North Carolina counties as untreated controls

	Y^R_{ct} = Infant mortality rate			Y^R_{ct} = Later-life mortality rate		
	(1)	(2)	(3)	(4)	(5)	(6)
A. Pooled						
Percent effect from Duke (=1)	–11.04**	–12.10***	–11.95***	–11.29***	–11.36***	–11.33***
	(4.32)	(4.36)	(4.39)	(3.34)	(3.27)	(3.28)
Observations	2,788	2,216	2,200	23,144	17,924	17,764
B. Black						
Percent effect from Duke (=1)	–14.73**	–16.41***	–16.47***	–6.66**	–6.23*	–6.23*
	(6.06)	(6.17)	(6.16)	(3.06)	(3.39)	(3.39)
Observations	2,516	1,977	1,968	20,060	15,650	15,620
C. White						
Percent effect from Duke (=1)	–8.97**	–9.71**	–9.32**	–12.34***	–12.82***	–12.79***
	(4.34)	(4.33)	(4.38)	(4.03)	(4.09)	(4.10)
Observations	2,531	1,992	1,980	20,080	15,670	15,620
P-value for difference by race	0.33	0.25	0.22	0.18	0.16	0.17
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	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. The dependent variable is the infant mortality rate in columns 1 to 3, and the later-life mortality rate in columns 4 to 6. Column 1 adds county-year observations from other southern counties that had cities, based on Bailey et al. (2015). Column 4 adds the analagous county-birth-year observations from the same counties added in Column 1, but for the later-life specifications. The Bailey et al. (2015) data are unbalanced before 1930 and after 1940. Thus, to ensure balance across years, the data in this table are restricted to birth years 1930 to 1940. Columns 2 and 5 drop all North Carolina counties that were not treated by Duke (i.e., did not receive a capital appropriation) before 1942. Columns 3 and 6 then remove control observations that did not have a non-profit hospital at any time from 1920 to 1940. Each coefficient represents the percent reduction in infant or later-life mortality rates due to receiving a capital appropriation from The Duke Endowment. A coefficient of -10 would mean that mortality declines by 10%. Coefficients are corrected for bias 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. Columns 4 to 6 also include flexible interactions of these control variables with follow-up year fixed effects. The weights are the number of births in a county and year. P-values reported in the bottom row come from tests for the difference in coefficients presented in panels B and C. They are estimated using 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 H2. Effect of Duke support on pooled infant mortality rate: Intensive margin instrumental variables estimates

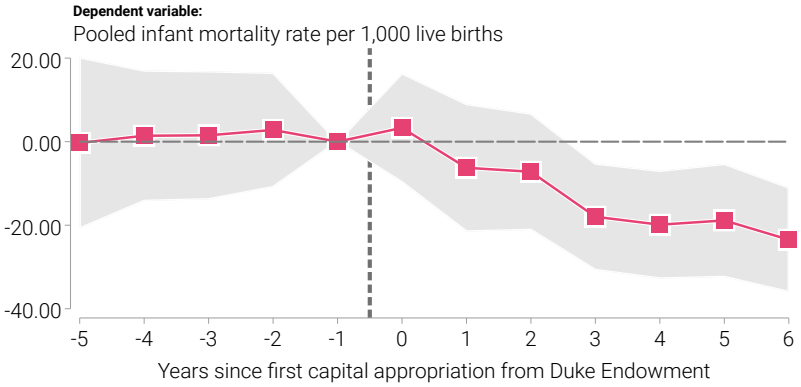
	Appropriations					Payments				
Specification:	Poisson (1)	OLS (2)	First stage (3)	Reduced form (4)	IV (5)	Poisson (6)	OLS (7)	First stage (8)	Reduced form (9)	IV (10)
Y^R_{ct}:	IMR	ln(IMR)	Appropriations	ln(IMR)	ln(IMR)	IMR	ln(IMR)	Payments	ln(IMR)	ln(IMR)
A. Southern counties with non-profit hospital (1930-1940)										
Percent effect from $1 million of Duke support	-6.08***	-6.35***			-9.05***	-9.31***	-9.44***			-14.93***
	(1.30)	(1.55)			(2.79)	(2.51)	(2.60)			(4.61)
Anderson-Rubin 95% Confidence Set					[-16.93, -3.27]**					[-29.96, -5.55]**
tF 95% Confidence Interval					[-16.69, -0.72]**					[-28.31, 0.94]*
(Endowment returns, billions) X 1(Non-profit hospital in NC before Duke)			0.35***	-2.80***				0.19***	-2.80***	
			(0.11)	(0.96)				(0.06)	(0.96)	
Observations	2,279	2,279	2,279	2,279	2,279	2,279	2,279	2,279	2,279	2,279
B. All NC counties (1922-1942)										
Percent effect from $1 million of Duke support	-4.03***	-4.63***			-11.07***	-6.98***	-8.37***			-16.82***
	(0.49)	(0.84)			(3.21)	(1.32)	(2.15)			(4.25)
Anderson-Rubin 95% Confidence Set					[-20.46, -5.24]***					[-27.83, -8.25]***
tF 95% Confidence Interval					[-19.47, -1.80]**					[-27.38, -4.72]***
(Endowment returns, billions) X 1(Non-profit hospital in NC before Duke)			0.38***	-3.69***				0.23***	-3.69***	
			(0.11)	(1.10)				(0.06)	(1.10)	
Observations	2,100	2,100	2,100	2,100	2,100	2,100	2,100	2,100	2,100	2,100

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 table 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 1, 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 2, expect that the top panel examines the outcomes with the non-NC control group of other southern counties with not-for-profit hospitals. 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 are corrected for bias in the following way: $100 \times (\exp(\beta) - 1)$. Standard errors are estimated using the delta method and are clustered at the county-level.

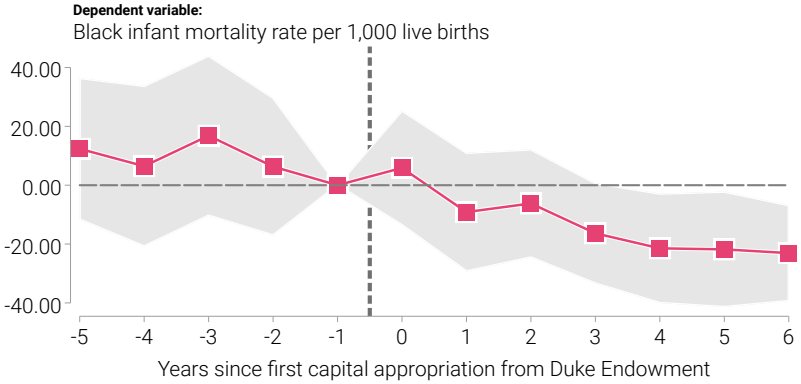
A23

Figure H1. Event studies: Replacing untreated North Carolina counties with other ineligible Southern counties

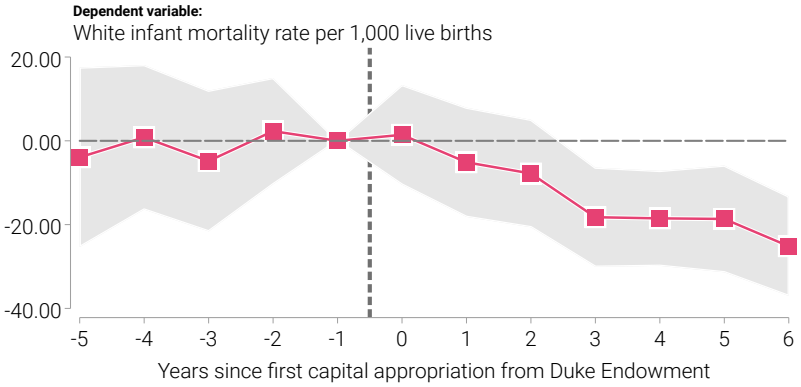
(a) Pooled



(b) Black



(c) White

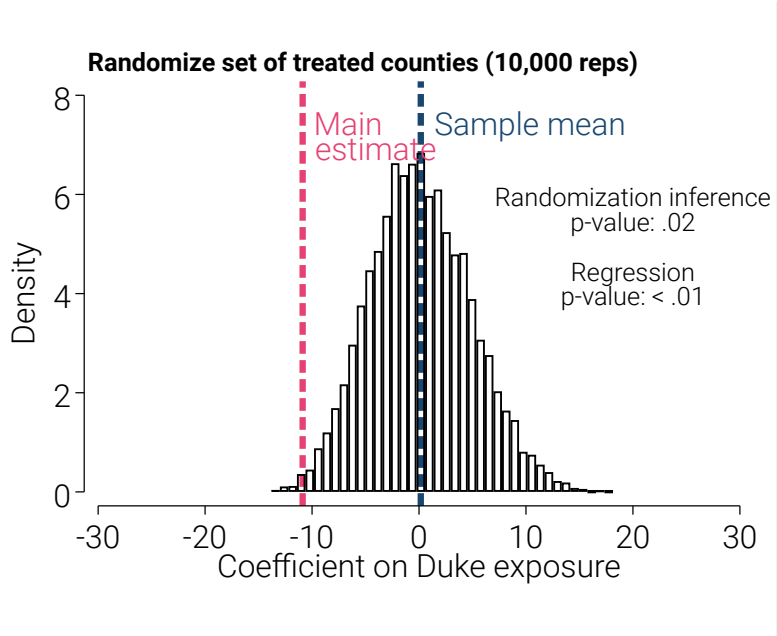


Notes: Each panel is an event study that corresponds to the regression from column 2 of Table H1. Panel A pools Black and White infants together, panel B examines the Black infant mortality rate, and panel C examines the White infant mortality rate. Each regression drops all North Carolina counties that 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 table 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.

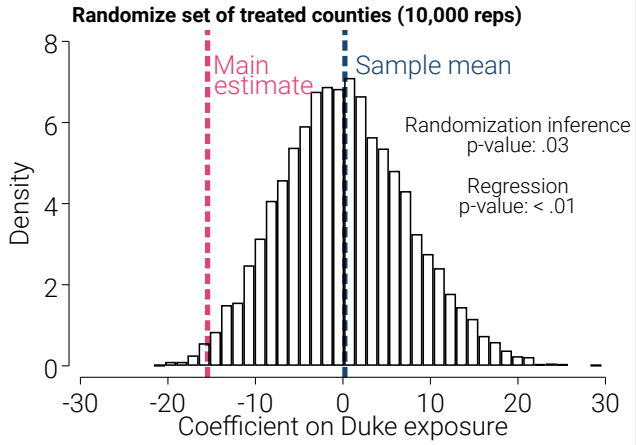
I Randomization of treatment

Figure I1. Randomization of Duke support for infant mortality rate results

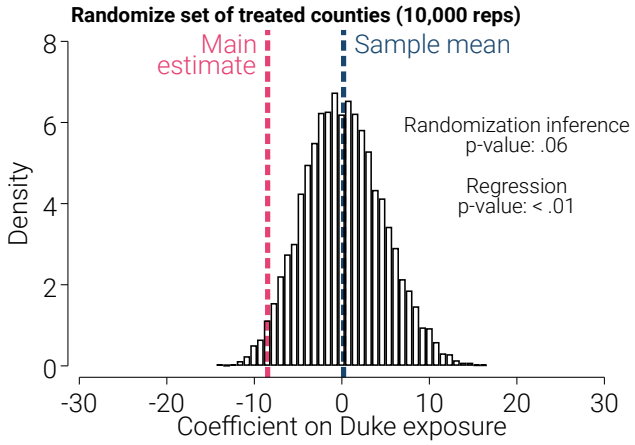
(a) Pooled



(b) Black



(c) White



Notes: This figure presents a histogram of coefficient estimates from 10,000 iterations of modifying the regression specification in column 5 of Table 1. The dependent variable is the pooled infant mortality rate in panel A, the Black infant mortality rate in panel B, or the White infant mortality rate in panel C. 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 1. 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 are corrected for bias 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.