Online Appendix

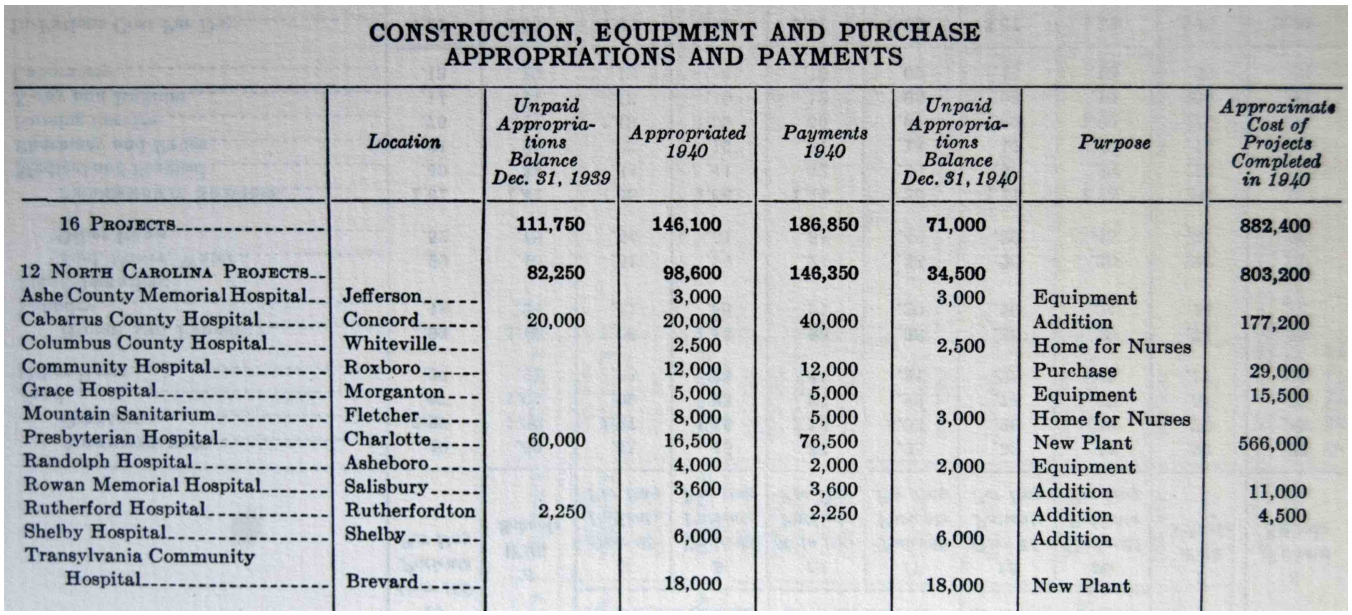
The Gift of a Lifetime: The Hospital, Modern Medicine, and Mortality

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

A Sample images of data sources

Figure A1. Sample images

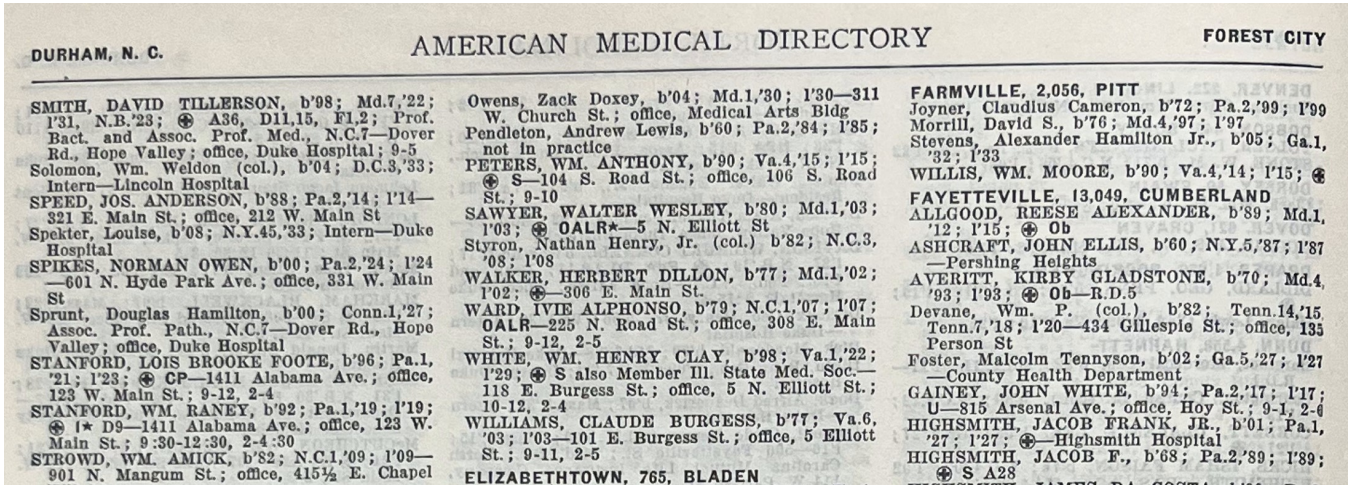
(a) The Duke Endowment's *Annual Report of the Hospital Section*



CONSTRUCTION, EQUIPMENT AND PURCHASE APPROPRIATIONS AND PAYMENTS

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

(b) *American Medical Directory*



DURHAM, N. C. — AMERICAN MEDICAL DIRECTORY — FOREST CITY

SMITH, DAVID TILLERSON, b'98; Md.7,'22; I'31, N.B.'23; ⊕ A36, D11,15, F1,2; Prof. Bact. and Assoc. Prof. Med., N.C.7—Dover Rd., Hope Valley; office, Duke Hospital; 9-5
Solomon, Wm. Weldon (col.), b'04; D.C.3,'33; Intern—Lincoln Hospital
SPEED, JOS. ANDERSON, b'88; Pa.2,'14; I'14—321 E. Main St.; office, 212 W. Main St
Spekter, Louise, b'08; N.Y.45,'33; Intern—Duke Hospital
SPIKES, NORMAN OWEN, b'00; Pa.2,'24; I'24—601 N. Hyde Park Ave.; office, 331 W. Main St
Sprunt, Douglas Hamilton, b'00; Conn.1,'27; Assoc. Prof. Path., N.C.7—Dover Rd., Hope Valley; office, Duke Hospital
STANFORD, LOIS BROOKE FOOTE, b'96; Pa.1, '21; I'23; ⊕ CP—1411 Alabama Ave.; office, 123 W. Main St.; 9-12, 2-4
STANFORD, WM. RANEY, b'92; Pa.1,'19; I'19; ⊕ I★ D9—1411 Alabama Ave.; office, 123 W. Main St.; 9:30-12:30, 2-4:30
STROWD, WM. AMICK, b'82; N.C.1,'09; I'09—901 N. Mangum St.; office, 415½ E. Chapel
Owens, Zack Doxey, b'04; Md.1,'30; I'30—311 W. Church St.; office, Medical Arts Bldg
Pendleton, Andrew Lewis, b'60; Pa.2,'84; I'85; not in practice
PETERS, WM. ANTHONY, b'90; Va.4,'15; I'15; ⊕ S—104 S. Road St.; office, 106 S. Road St.; 9-10
SAWYER, WALTER WESLEY, b'80; Md.1,'03; I'03; ⊕ OALR★—5 N. Elliott St
Styron, Nathan Henry, Jr. (col.) b'82; N.C.3, '08; I'08
WALKER, HERBERT DILLON, b'77; Md.1,'02; I'02; ⊕—306 E. Main St.
WARD, IVIE ALPHONSO, b'79; N.C.1,'07; I'07; OALR—225 N. Road St.; office, 308 E. Main St.; 9-12, 2-5
WHITE, WM. HENRY CLAY, b'98; Va.1,'22; I'29; ⊕ S also Member Ill. State Med. Soc.—118 E. Burgess St.; office, 5 N. Elliott St.; 10-12, 2-4
WILLIAMS, CLAUDE BURGESS, b'77; Va.6, '03; I'03—101 E. Burgess St.; office, 5 Elliott St.; 9-11, 2-5

ELIZABETHTOWN, 765, BLADEN

FARMVILLE, 2,056, PITT
Joyner, Claudius Cameron, b'72; Pa.2,'99; I'99
Morrill, David S., b'76; Md.4,'97; I'97
Stevens, Alexander Hamilton Jr., b'05; Ga.1, '32; I'33
WILLIS, WM. MOORE, b'90; Va.4,'14; I'15; ⊕

FAYETTEVILLE, 13,049, CUMBERLAND
ALLGOOD, REESE ALEXANDER, b'89; Md.1, '12; I'15; ⊕ Ob
ASHCRAFT, JOHN ELLIS, b'60; N.Y.5,'87; I'87—Pershing Heights
AVERITT, KIRBY GLADSTONE, b'70; Md.4, '93; I'93; ⊕ Ob—R.D.5
Devane, Wm. P. (col.), b'82; Tenn.14,'15, Tenn.7,'18; I'20—434 Gillespie St.; office, 135 Person St
Foster, Malcolm Tennyson, b'02; Ga.5,'27; I'27—County Health Department
GAINEY, JOHN WHITE, b'94; Pa.2,'17; I'17; U—815 Arsenal Ave.; office, Hoy St.; 9-1, 2-6
HIGHSMITH, JACOB FRANK, JR., b'01; Pa.1, '27; I'27; ⊕—Highsmith Hospital
HIGHSMITH, JACOB F., b'68; Pa.2,'89; I'89; ⊕ S A28

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

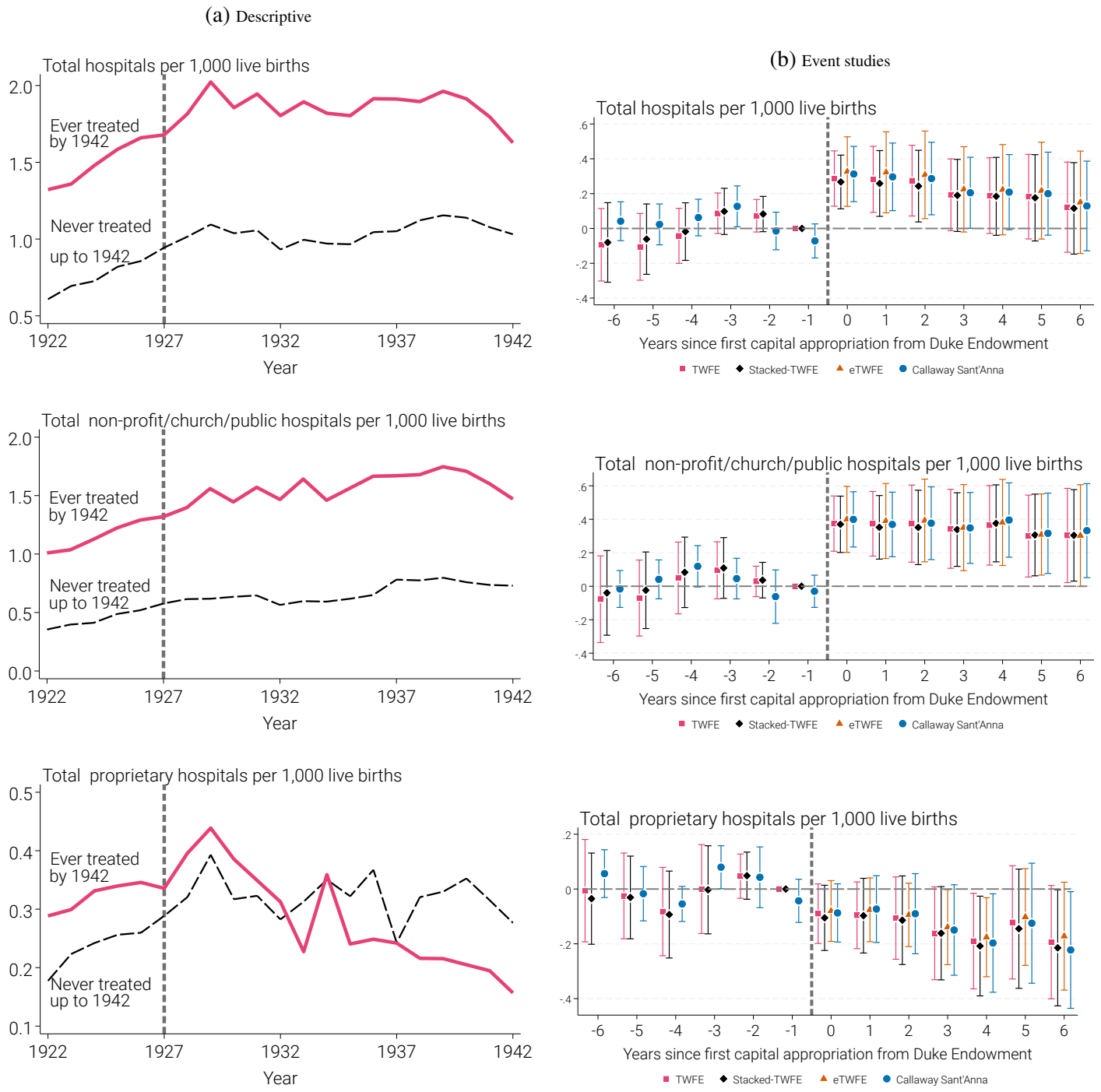
B First stage for hospitals and hospital beds: Additional results

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

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

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

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



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

C First stage results for doctors and other health care professionals

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

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

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

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

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

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

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

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

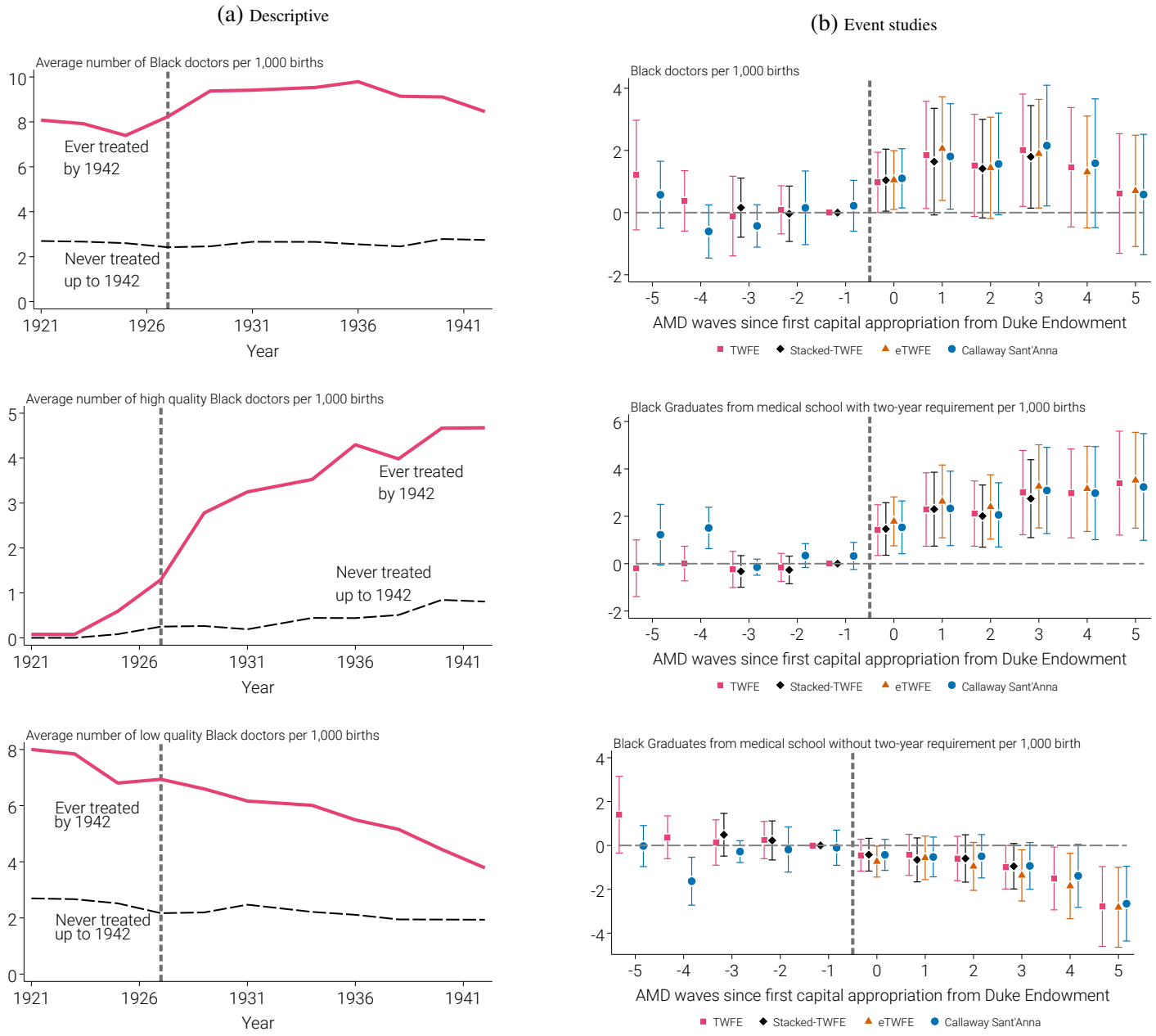
Notes: Each coefficient comes from a separate regression estimated by two-way fixed effects. In column 1, the dependent variable is the number of doctors (Panel A), the number of Black doctors (Panel B), or the number of White doctors (Panel C) in a county and publication year of the *American Medical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

Table C4. Effects on doctors: Additional heterogeneity analyses

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

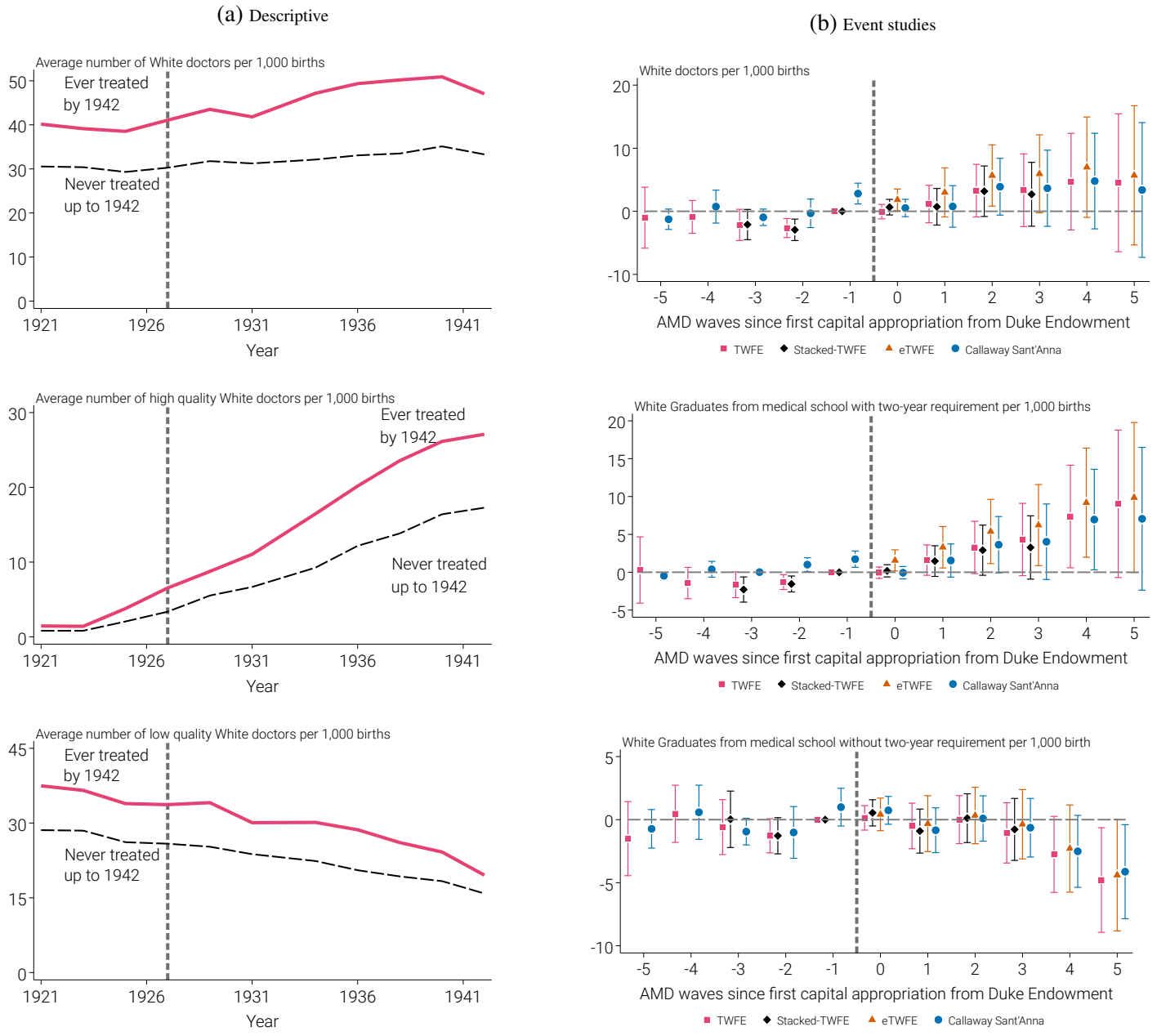
Notes: Each coefficient comes from a separate regression estimated by two-way fixed effects. The unit of observation is a county and publication year of the *American Medical Directory* (AMD wave). During the sample period, the AMD was published in 1921, 1923, 1925, 1927, 1929, 1931, 1934, 1936, 1938,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

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



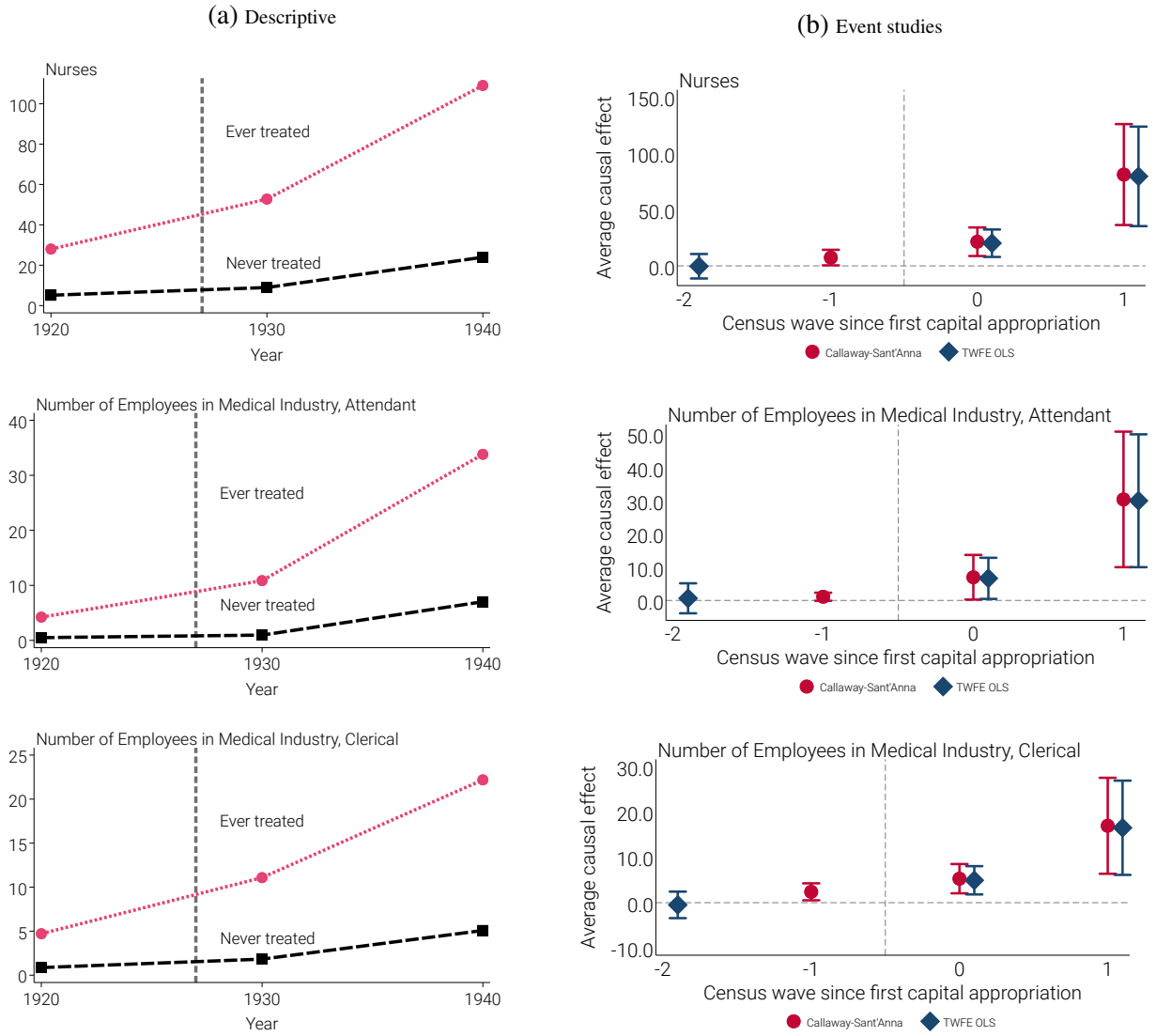
Notes: Column (a) plots plots shows the average annual number of doctors in a county by Duke treatment status. Column (b) plots event study estimates of coefficient values and 95% confidence intervals for the lead and lag indicator variables for relative time periods from $t = -5$ to $t = 5$ around the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

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



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

Figure C3. Employment of other health care professionals



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

D Infant mortality: Additional results

Table D1. Robustness of infant mortality results to alternate estimators

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

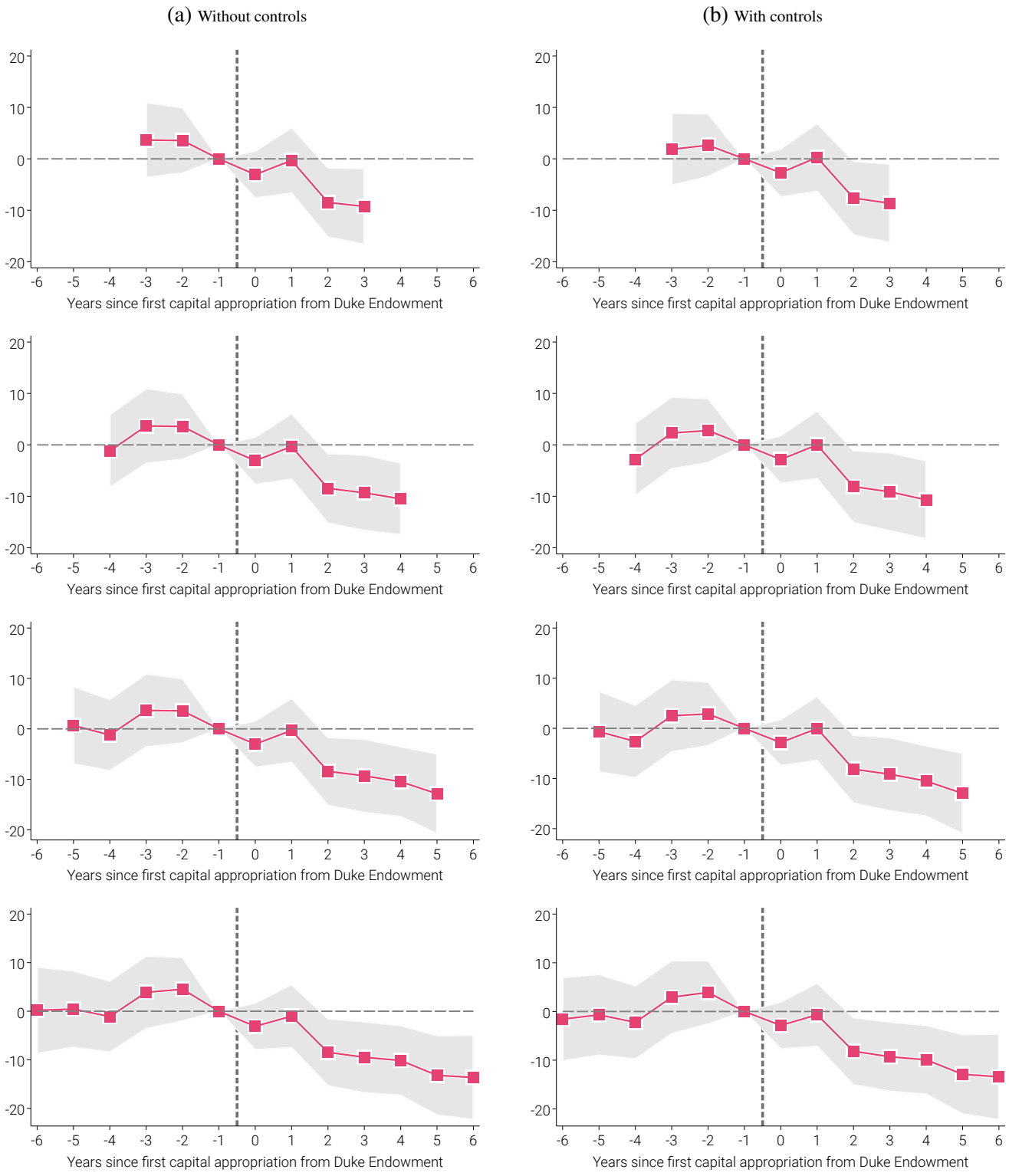
Notes: Each coefficient comes from a separate regression estimated by Poisson (columns 1 and 5), stacked Poisson with a balanced panel (columns 2 and 6), the extended two-way fixed effects estimator by Wooldridge (2021) (columns 3 and 7), or the Callaway and Sant'Anna (2021) estimator (columns 4 and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

Table D2. Robustness of infant mortality results to log specification

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

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

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



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

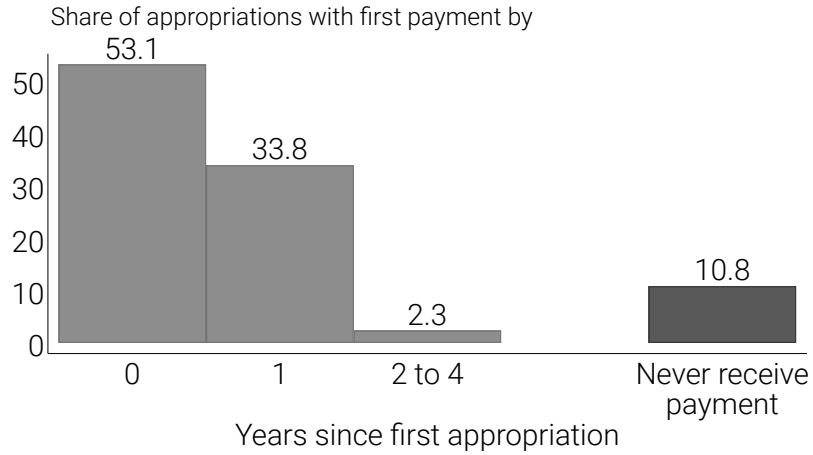
E Duke funding: Heterogeneity by project type

Table E1. Appropriation and payment details by project type

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

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

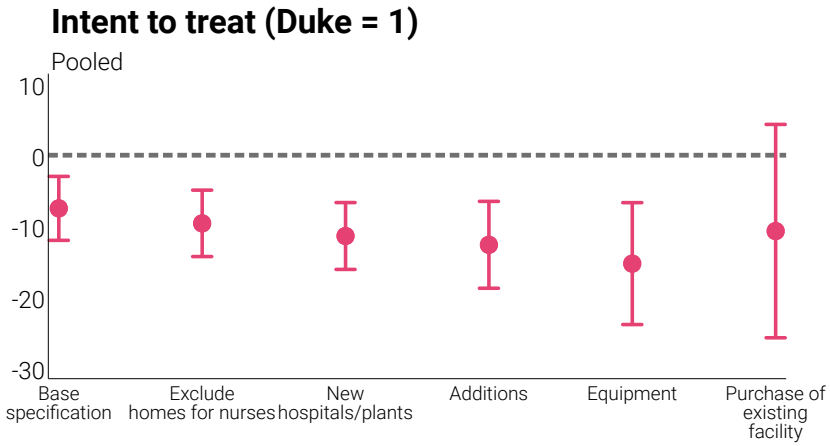
Figure E1. Time from appropriation to payment



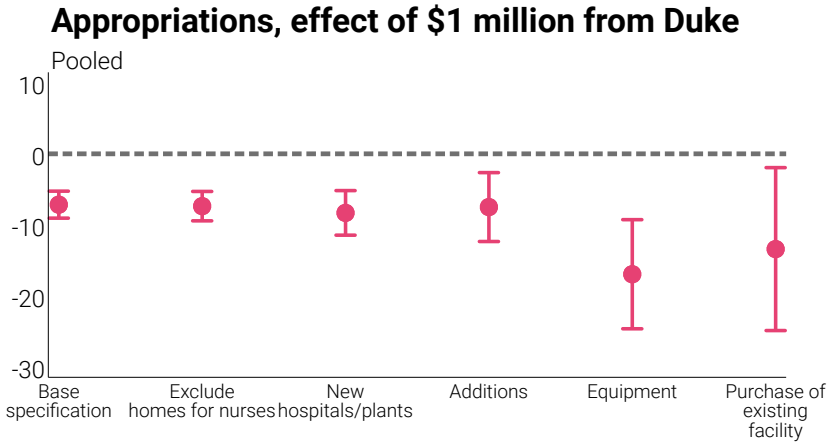
Notes: This figure plots the share of unique appropriations by the number of years after the initial appropriation when the first payment was received by the hospital.

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

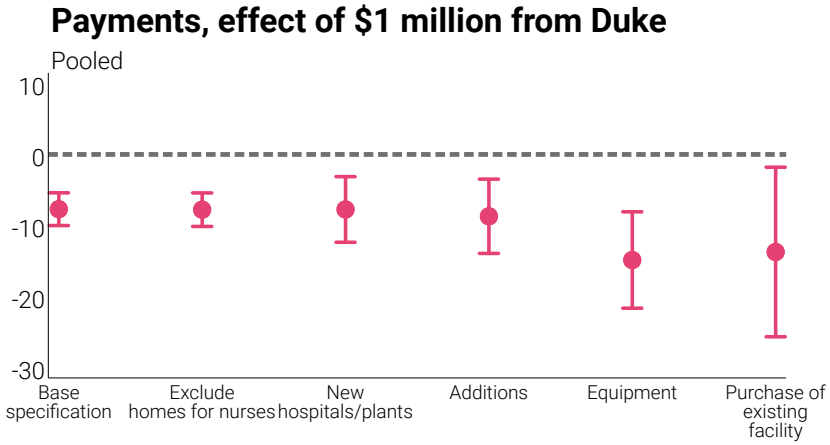
(a) Intent to treat (binary)



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



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



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

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

F.1 Event studies

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

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

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

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

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

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

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

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

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

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

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

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

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

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

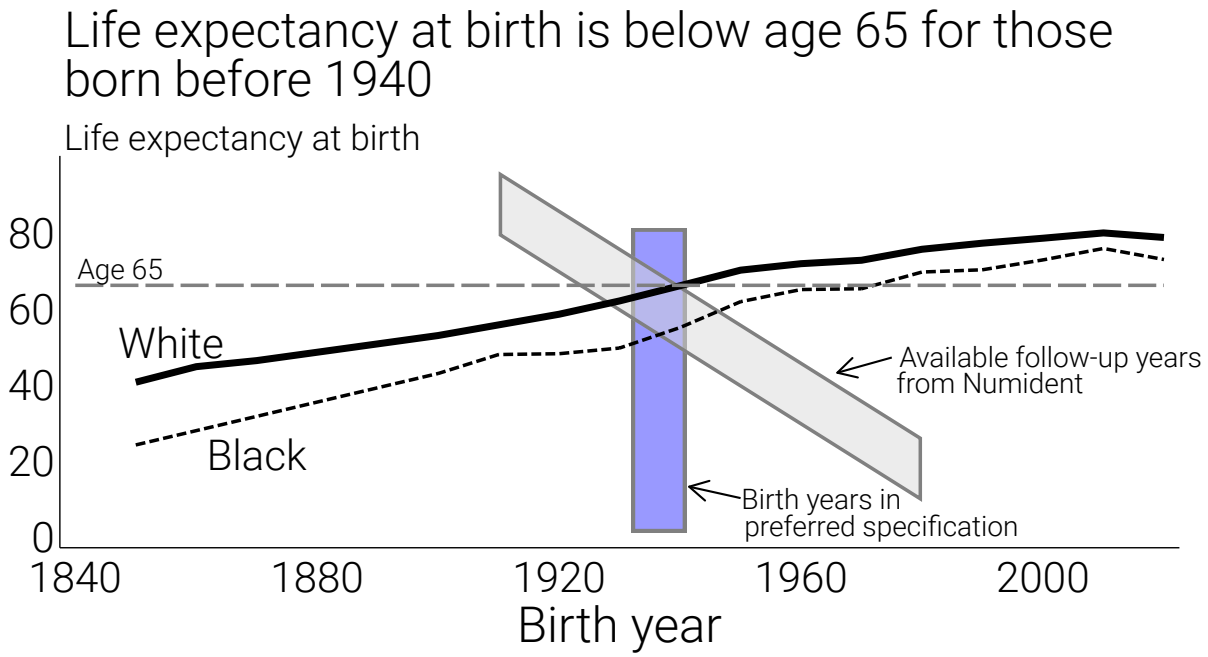
Notes: Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. In columns 1 to 3, the unit of observation is a county-by-year of birth cell. In columns 4 to 6, the unit of observation is a birth county by birth year by follow-up age triplet. Birth cohorts are restricted to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

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

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

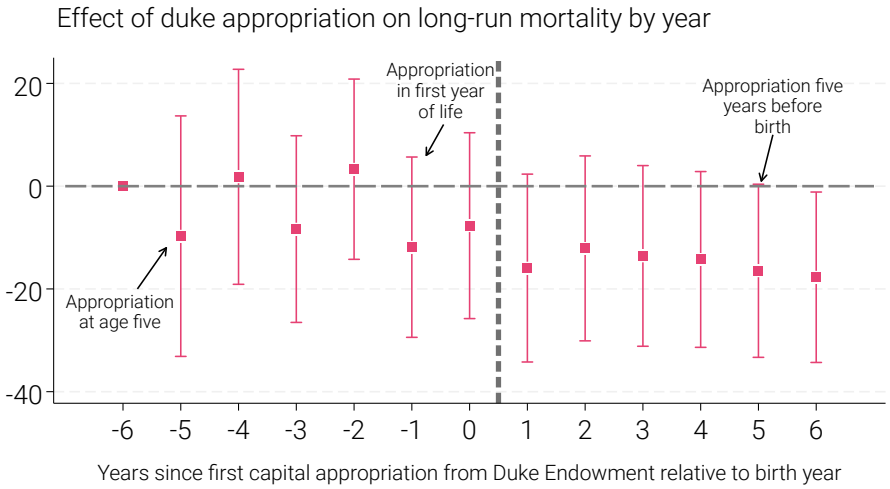
Notes: Each coefficient comes from a separate Poisson pseudo maximum likelihood regression. The unit of observation is a birth county by birth year by follow-up age triplet. Birth cohorts are restricted to 1932 to 1941. Deaths are restricted to ages 56 to 64 and years 1988 to 2005. The dependent variable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

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



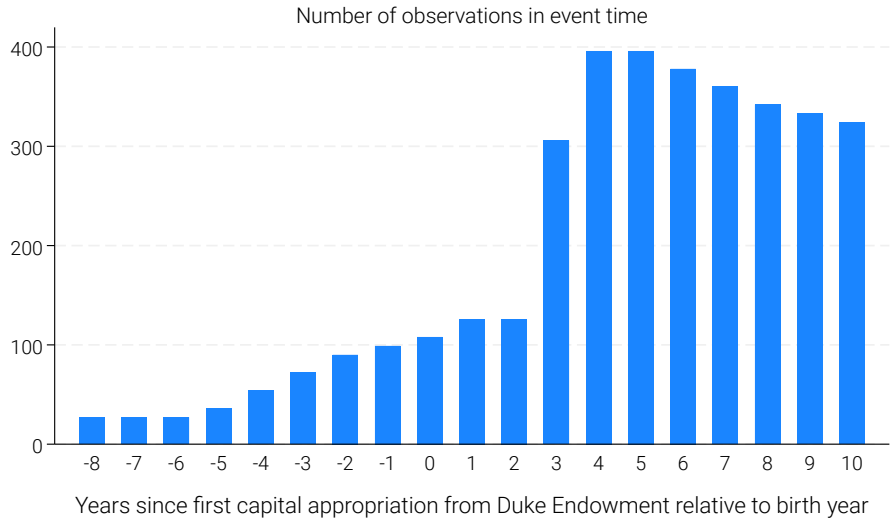
Notes: Life expectancy by birth cohort for White (solid line) and Black (dashed line) individuals from Haines (2008); Arias (2014); Arias et al. (2021). The parallelogram shaded in grey represents the set of birth cohorts and ages at death available in the Numident. The rectangle shaded in blue represents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

Figure F2. Long-run analysis: Event study estimates



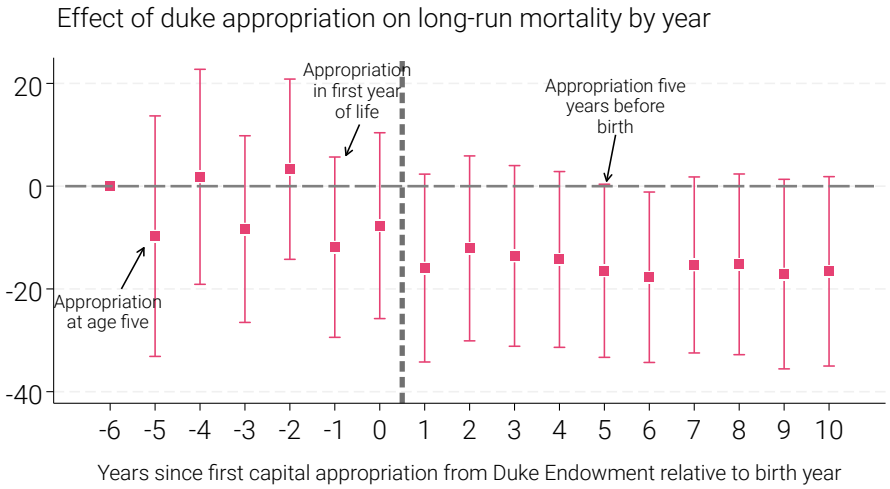
Notes: Plot contains long-run event-study estimates the effect of the *same* treatment as the short-run analysis – Duke support *around the time of birth*. In all event studies, event time is defined as the year of birth minus the year of first capital appropriation from The Duke Endowment. For example,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

Figure F3. Unbalanced composition of treated observations across event times



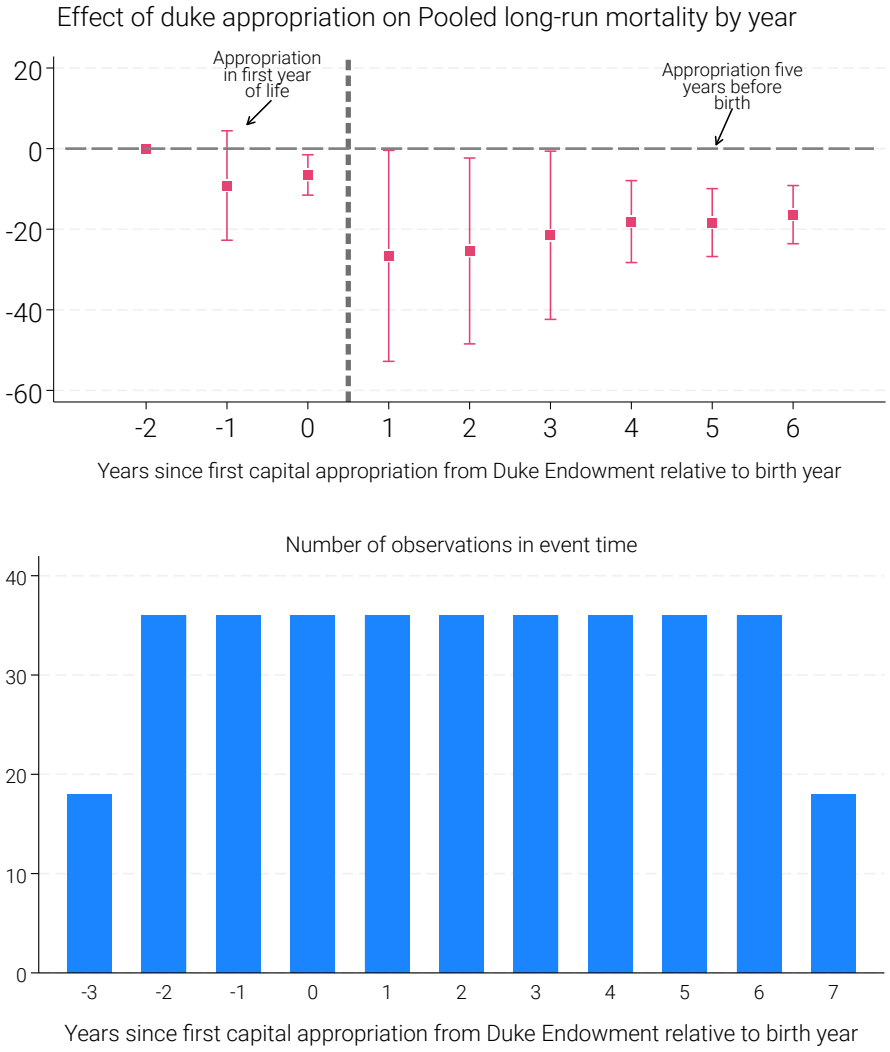
Notes: This Figure displays the treated number of observations for each event time that are included in the event studies reported in Figures F2 and F4.

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



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

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



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

G Other results

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

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

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

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

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

[35] 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.

[36] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

H Panel construction and sample size

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

1. There are 12 observations among 4 counties for which the number of deaths of Black infants exceeds the number of Black births. For these observations we replace the birth count with the mortality count.

2. There are 5 counties and 24 observations with zero Black births. In our main specifications for infant mortality rates by race (panels B and C of Table 3), we exclude these 5 counties in order to maintain a balanced panel. The observations with zero Black births will drop out since the weights are undefined for these observations.

3. After dropping all observations for the 5 counties that ever have zero Black births, there are 19 counties and 101 observations with zero Black deaths and non-zero Black births. There are 2 other counties and 4 observations with zero White deaths and non-zero White births. In specifications with the log of the infant mortality rate as the dependent variable we drop these 21 counties from the sample, in addition to the 5 counties already dropped (for a total of 26 counties). Our main specification is estimated using Poisson pseudo-maximum likelihood which can handle the presence of zero values for the dependent variable, unlike a log-level specification that would drop these observations.

4. We exclude the aforementioned 26 counties from estimation samples in specifications with the log of the infant mortality rate as the dependent variables in the following exhibits: Figure 5 where noted, Online Appendix Tables D2 and I1, and Online Appendix Figure I1.

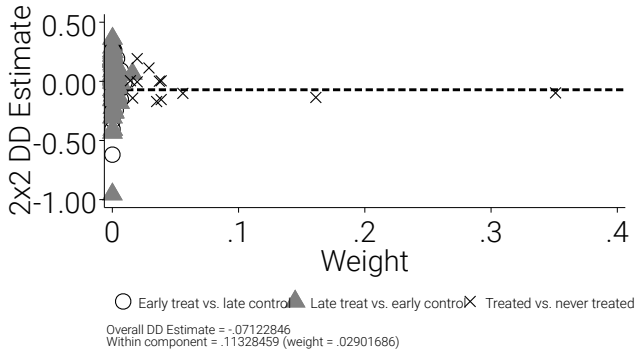
I Event study diagnostics

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

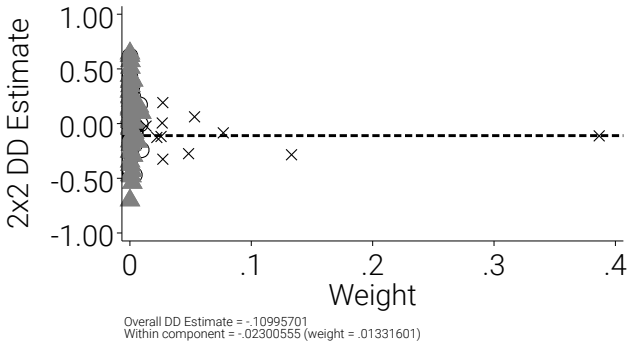
Type of DD comparison	Average Estimate	Number of 2x2 Comparisons	Total Weight
A. Pooled infant mortality rate			
Earlier Treated vs. Later Treated Controls	–0.017	66	0.092
Later Treated vs. Earlier Treated Controls	–0.009	66	0.091
Treated vs. Untreated Controls	–0.084	12	0.817
Average DD estimate	–0.071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

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

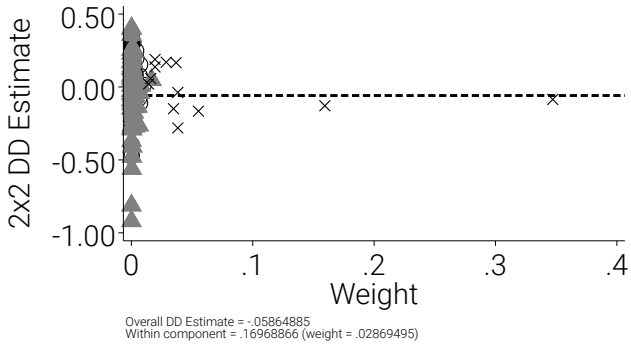
Figure I1. Goodman-Bacon (2021a) decomposition diagnostic



(a) Pooled infant mortality rate



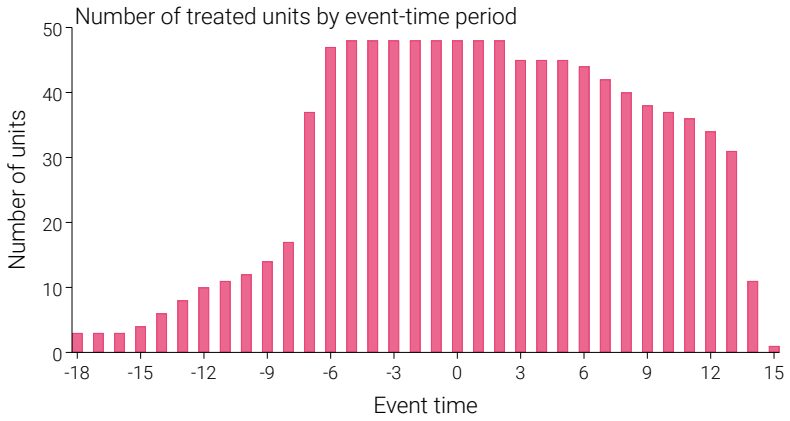
(b) Black infant mortality rate



(c) White infant mortality rate

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

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



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

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

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

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

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

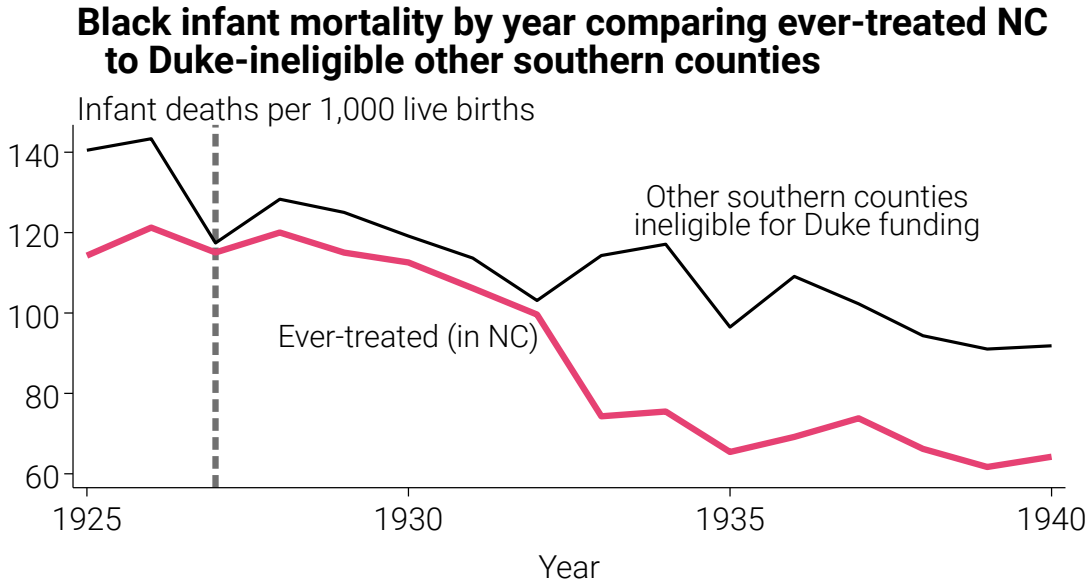
Table J1. Effect of Duke support on pooled infant mortality rate: Intensive margin instrumental variables estimates

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

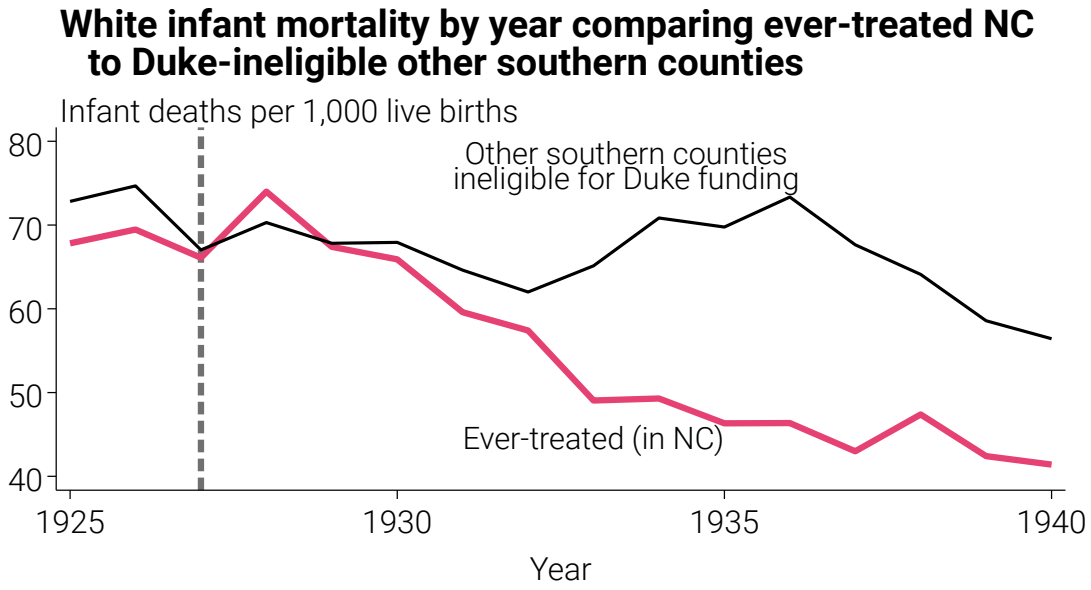
Notes: Each coefficient comes from a separate regression. Columns 1 through 5 focus on cumulative appropriations from The Duke Endowment in 2017 $. Columns 6 through 10 focus on cumulative payments from The Duke Endowment in 2017 $. All specifications in the top panel A, use the same sample that includes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

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

(a) Black infant mortality



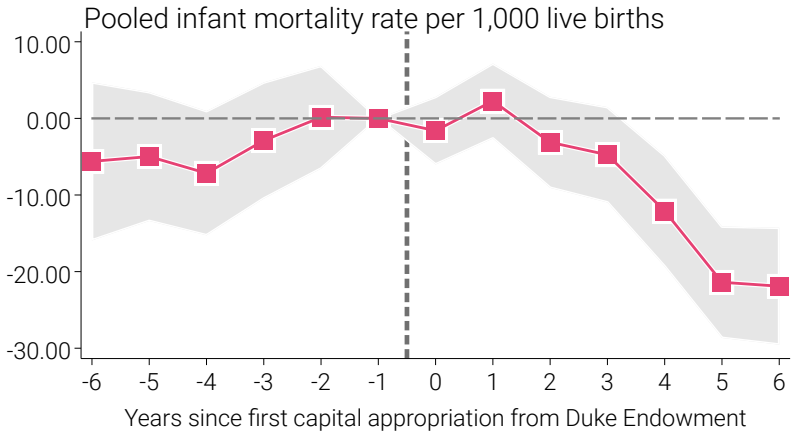
(b) White infant mortality



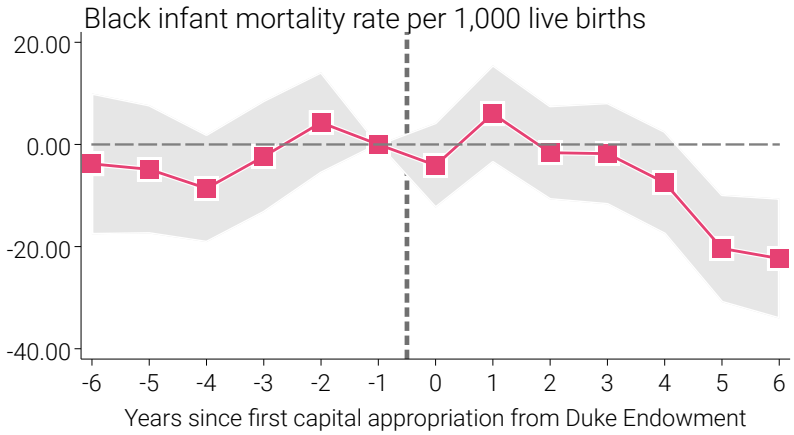
Notes: This figure compares county-year infant mortality by race for counties in North Carolina that are "Ever treated" during our sample (pink solid line) to infant mortality rates from from other Southern counties ineligible for Duke funding (thick black line). These data come from Fishback et al.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

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

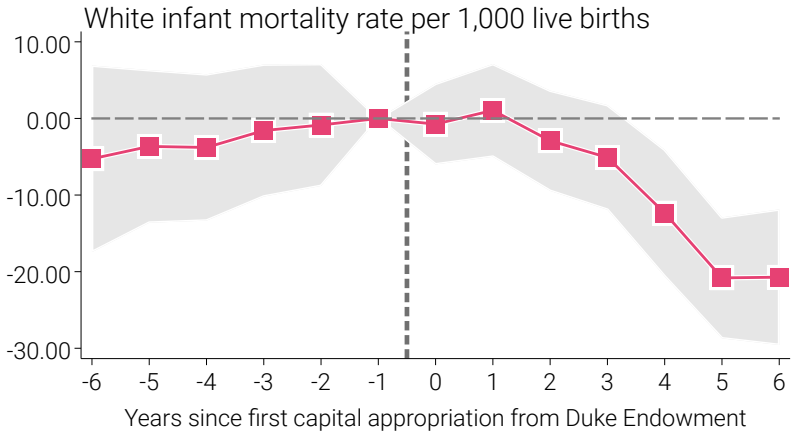
(a) Pooled



(b) Black



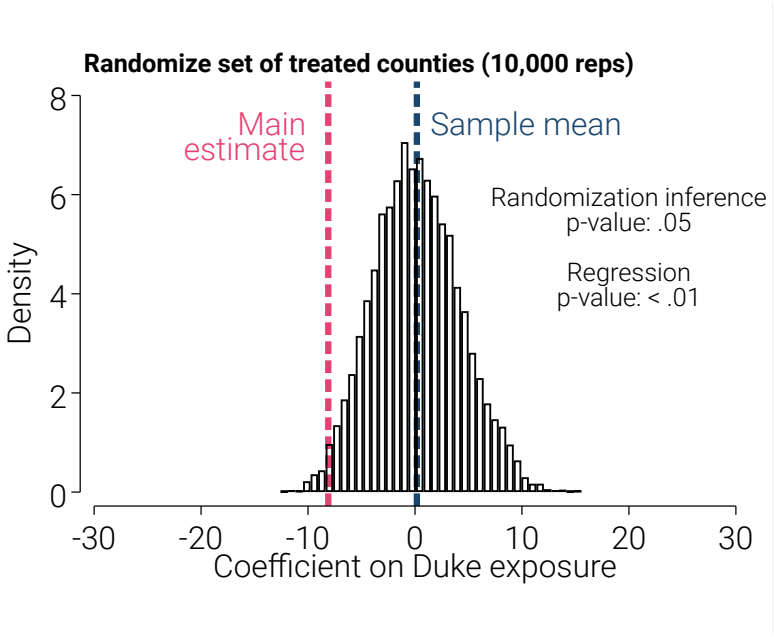
(c) White



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

K Randomization of the treatment

Figure K1. Randomization of Duke support for infant mortality rate



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

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

Table L1. First-stage hospital analysis extended to 1950

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

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

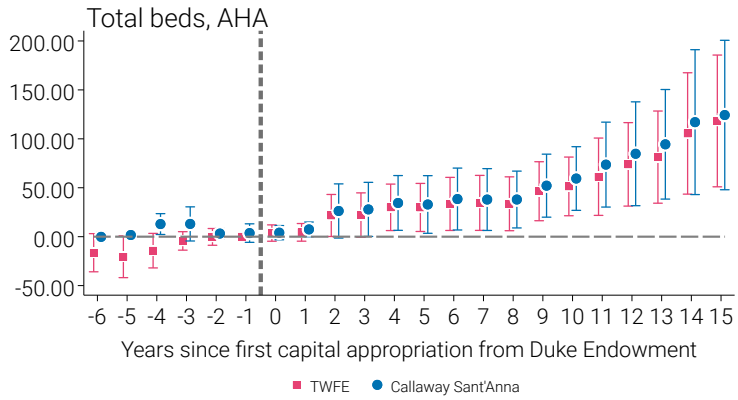
Table L2. Effects on infant mortality: Non-Carolina controls extended to 1962

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

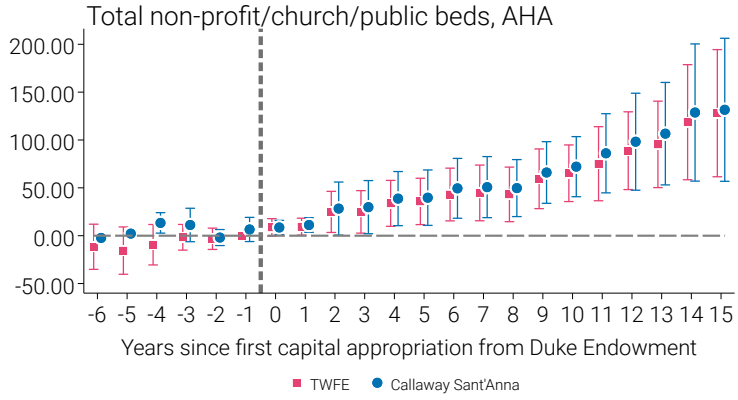
Notes: The specifications reported in this table extend the analysis with non-Carolina control counties by including data up to 1962. All specifications use our main infant mortality measure for North Carolina counties for our main sample years (1922 to 1942). For the non-Carolina controls, columns 1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

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

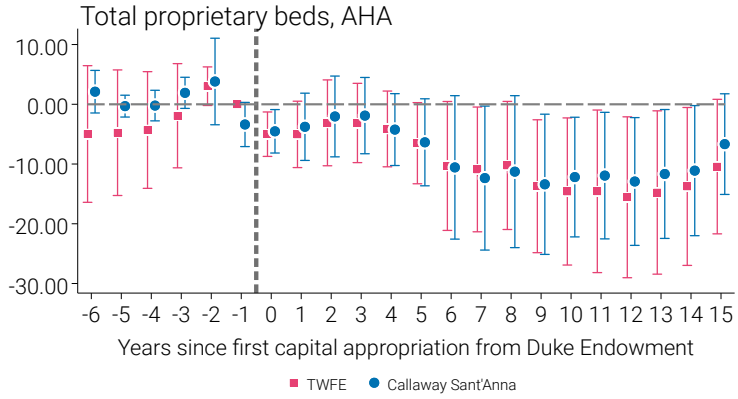
(a) Total Beds



(b) Church/Non-Profit/Public Beds



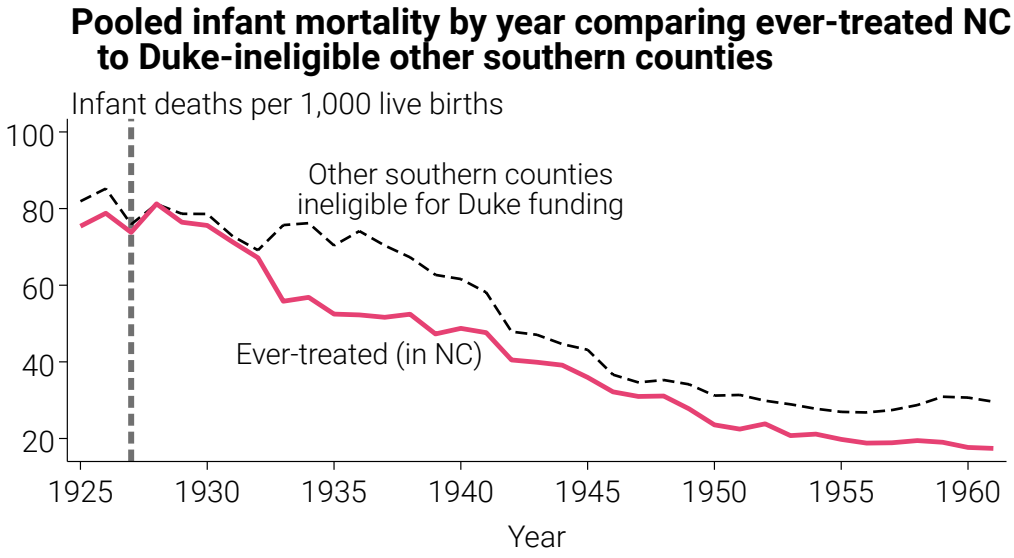
(c) Proprietary Beds



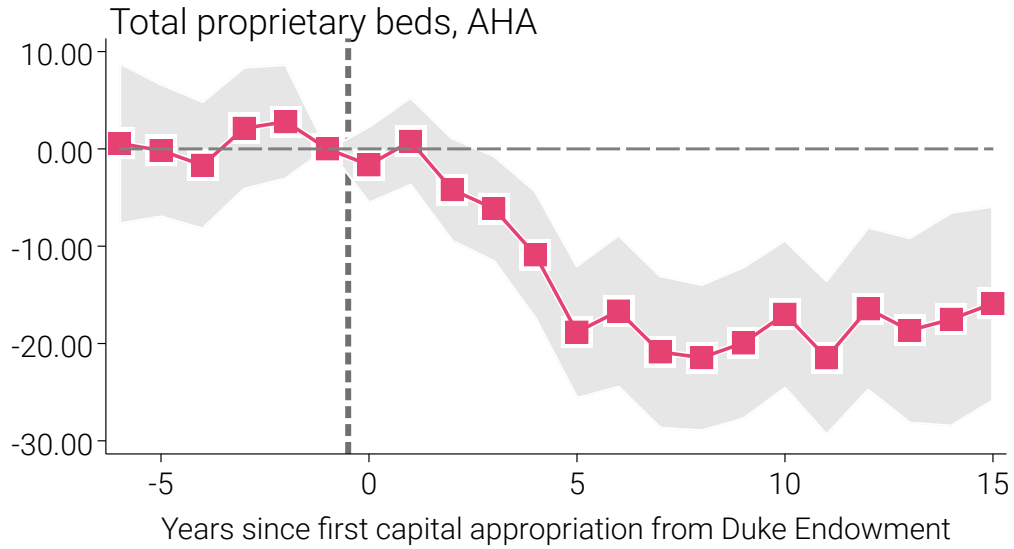
Notes: The event studies correspond to the specifications reported in column 2 of panel A in Online Appendix Table L1.

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

(a) Infant mortality rate across time



(b) Event study



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

M Propensity score matching: Alternative control group and falsification test

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

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

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

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

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

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

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

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

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

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

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

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

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

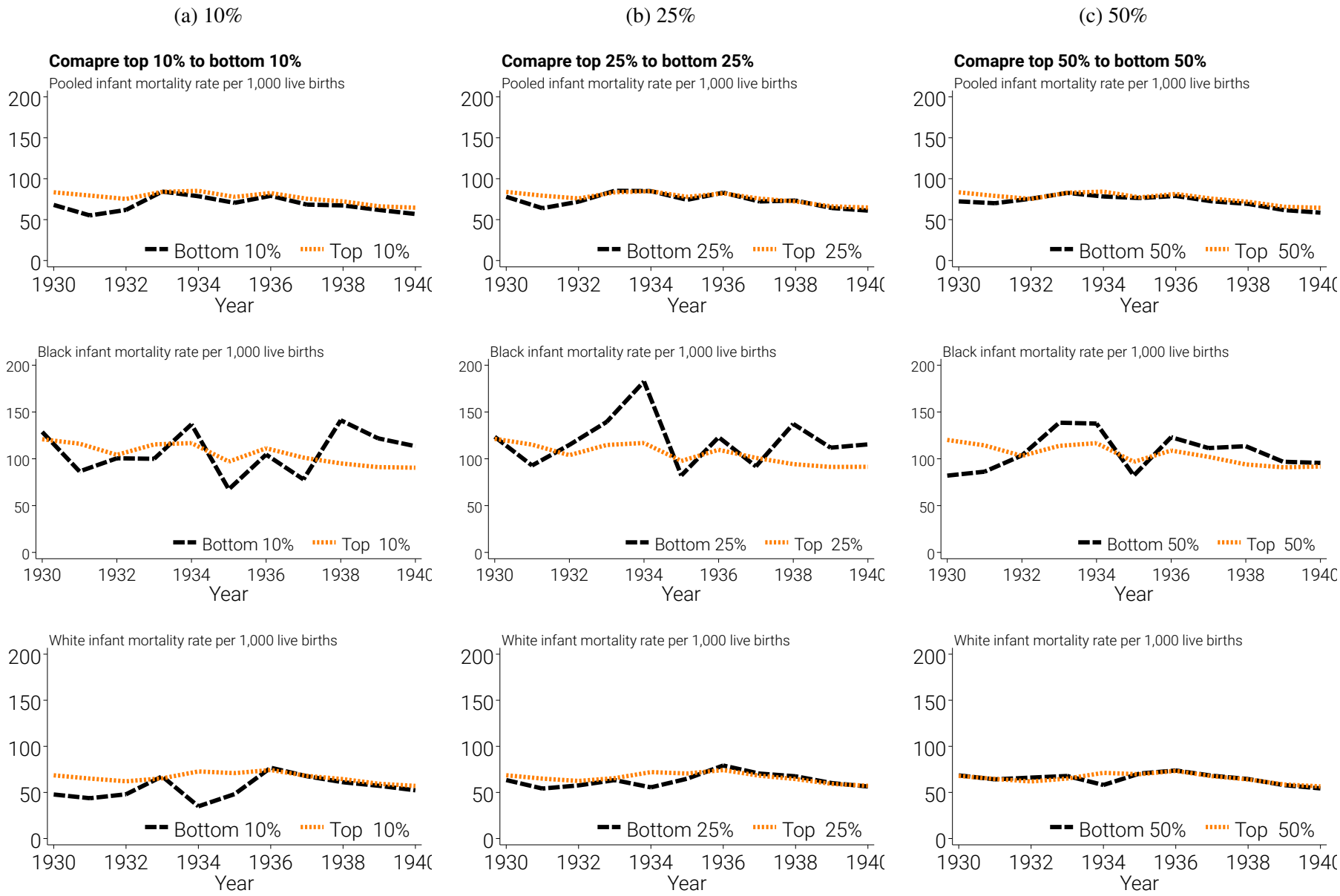


Table M2. Propensity score match - top 10%

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

Table M3. Propensity score match - top 25%

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

Table M4. Propensity score match - top 50%

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

N Summary statistics

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

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

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

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

Appendix References

Anderson, T. W. and Herman Rubin (1949) "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, Vol. 20, No. 1, pp. 46 – 63.

Arias, E. (2014) "United States life tables, 2010," *National vital statistics reports*, Vol. 63, No. 7.

Arias, E., B. Tejada-Vera, F. Ahmad, and K.D. Kochanek (2021) "Provisional life expectancy estimates for 2020," *Vital Statistics Rapid Release*, No. 15, pp. 50–61.

Bailey, Martha, Karen Clay, Price Fishback, Michael Haines, Shawn Kantor, Edson Severnini, and Anna Wentz (2015) *U.S. County-Level Natality and Mortality Data, 1915-2007*: Inter-university Consortium for Political and Social Research [distributor].

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

Callaway, Brantly and Pedro H.C. Sant'Anna (2021) "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, Vol. 225, No. 2, pp. 200–230.

Feigenbaum, James, Lauren Hoehn-Velasco, and Sophie Li (2023) "Germ Theory at Home: The Role of Private Action in Reducing Child Mortality during the Epidemiological Transition," Working Paper.

Fishback, Price V., Michael R. Haines, and Shawn Kantor (2007) "Births, Deaths, and New Deal Relief during the Great Depression," *The Review of Economics and Statistics*, Vol. 89, No. 1, pp. 1–14.

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

Goodman-Bacon, Andrew (2021) "The Long-Run Effects of Childhood Insurance Coverage: Medicaid Implementation, Adult Health, and Labor Market Outcomes," *American Economic Review*, Vol. 111, No. 8, pp. 2550–2593.

Haines, Michael (2008) "Fertility and Mortality in the United States," in Robert Whaples ed. *EH.Net Encyclopedia*: Economic History Association.

Jayachandran, Seema, Adriana Lleras-Muney, and Kimberly Smith (2010) "Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs," *American Economic Journal: Applied Economics*, Vol. 2, No. 2, pp. 118–146.

Lazuka, Volha (2020) "Infant Health and Later-life Labor Market Outcomes: Evidence from the Introduction of Sulfa Antibiotics in Sweden," *Journal of Human Resources*, Vol. 55, No. 2, pp. 660–698.

Lee, David S., Justin McCrary, Marcelo J. Moreira, and Jack Porter (2022) "Valid t-Ratio Inference for IV," *American Economic Review*, Vol. 112, No. 10, pp. 3260–90.

Pei, Zhuan, Jörn-Steffen Pischke, and Hannes Schwandt (2019) "Poorly Measured Confounders Are More Useful on the Left than on the Right," *Journal of Business & Economic Statistics*, Vol. 37, No. 2, pp. 205–216.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Anne Chen, Grace Cooper, Stephanie Richards, and Megan Schouweiler (2023) *IPUMS USA: Version 14.0 [dataset]*, Minneapolis, MN: IPUMS.

Wooldridge, Jeffrey M. (2021) "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators," Working Paper.