

Supplemental Appendix

Medical Technology and Life Expectancy:
Evidence from the Antitoxin Treatment of Diphtheria

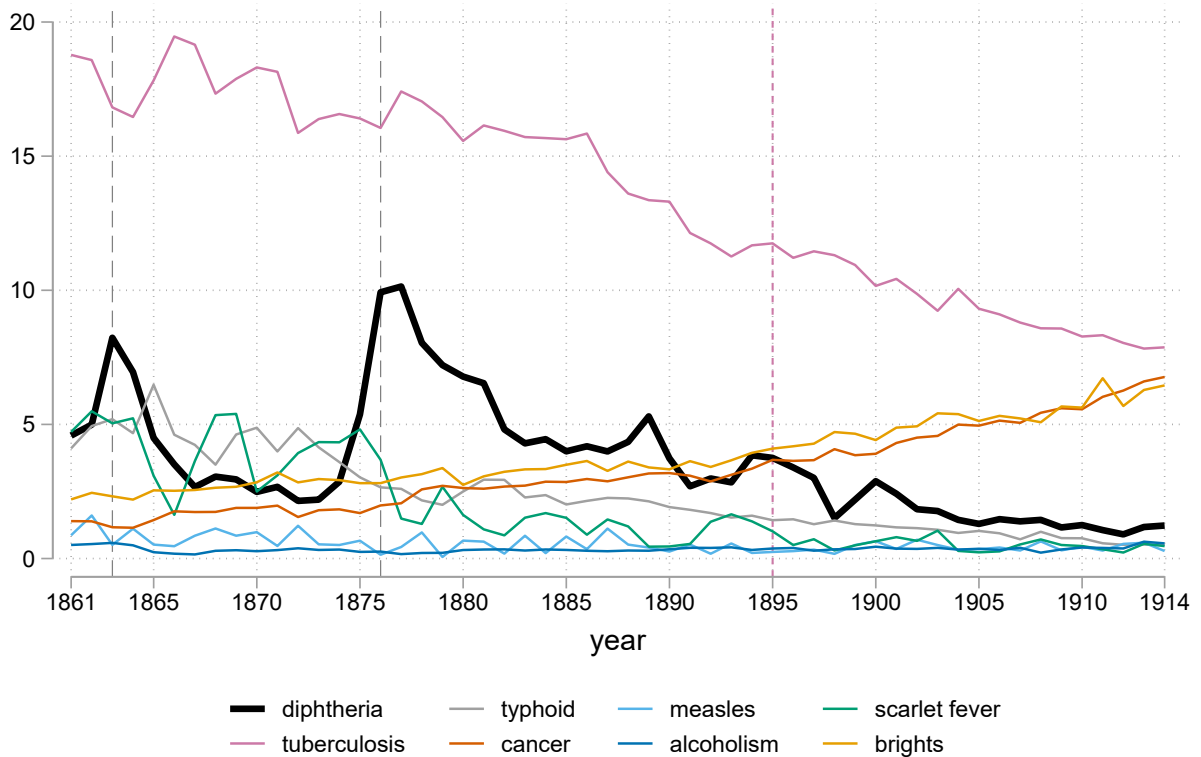
Philipp Ager Casper W. Hansen Peter Z. Lin*

*Ager: University of Mannheim and CEPR, philipp.ager@gmail.com; Lin: Western Kentucky University, zhixian.lin@wku.edu; Hansen: University of Copenhagen and CEPR, casper.worm.hansen@econ.ku.dk

A Appendix

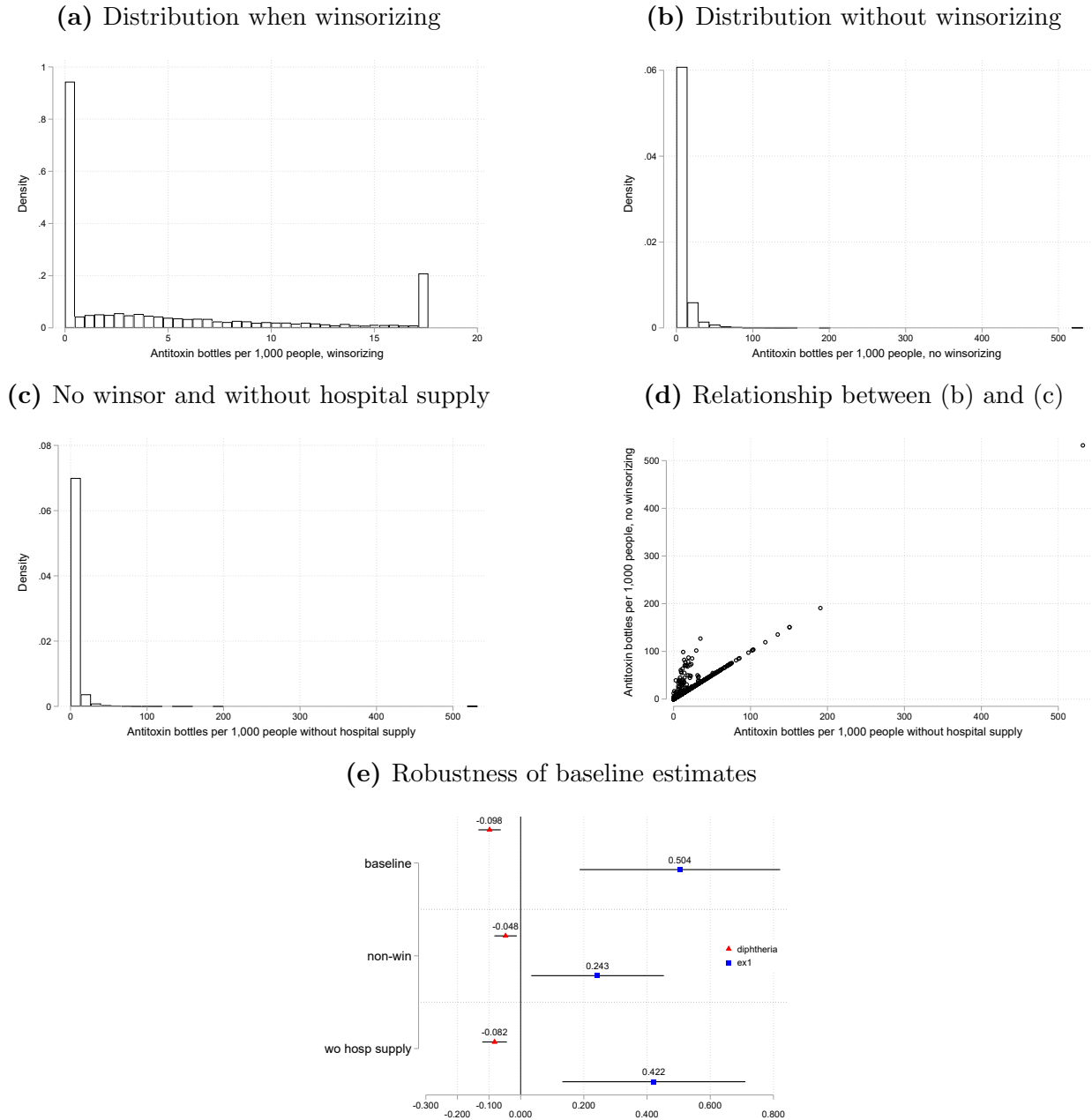
Additional Figures

Figure A.1: Cause-specific death ratios



Notes: This figure illustrates the trends in various major causes of death, as categorized in “the Annual Report on Births, Marriages, and Deaths in Massachusetts” (1914), shown as a percentage of the total number of deaths in the state. The bold black curve represents the percentage of deaths caused by diphtheria, which, for instance, approaches 10 percent in 1875.

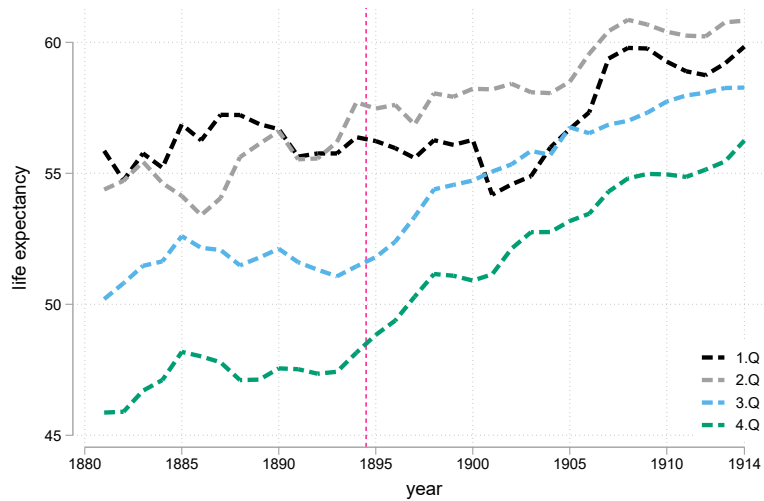
Figure A.2: Bottles distribution and winsorizing



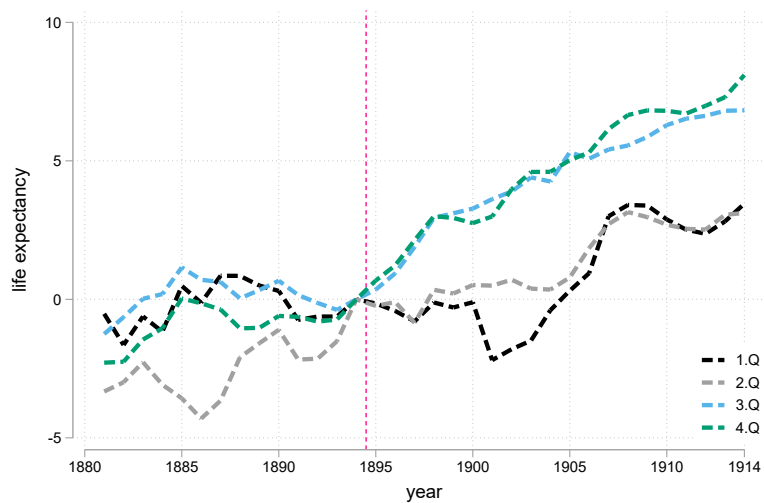
Notes: This figure illustrates the distribution of antitoxin bottles per 1,000 people from 1895 onward. Panel (a) shows the distribution with winsorization applied (i.e., baseline), while Panel (b) presents the distribution without winsorization. Panel (c) displays the unwinsorized distribution, excluding antitoxin bottles supplied to hospitals. Panel (d) compares the distributions shown in Panels (b) and (c). Finally, Panel (e) reports the 2SLS estimates for diphtheria and life expectancy at age 1 using these three measures: the winsorized baseline, the unwinsorized, and the unwinsorized excluding hospital-supplied bottles, as the endogenous explanatory variables. The KP-F-Stats are for diphtheria 35.53 (baseline); 7.35; 30.25, respectively, and for life expectancy 33.69 (baseline); 6.35; 30.03, respectively.

Figure A.3: Development of life expectancy at age 1 by treatment intensity

(a) Average by group



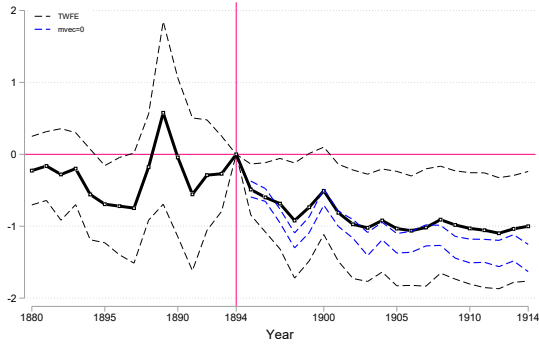
(b) And deviation from 1894 values



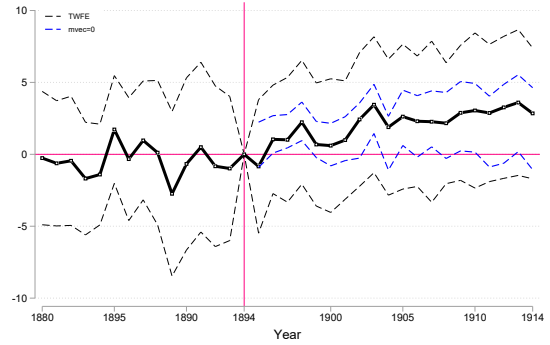
Notes: This figure shows the development of life expectancy at age 1 by groups according to quartiles of treatment intensity. In particular, we have collapsed the baseline sample of municipalities into four regions (according to their level of treatment) and then for each region calculated life expectancy at age 1 for each year. This avoids the problem of small populations when deriving the life tables and calculating life expectancy. Panel A shows the three-year moving average by group, while Panel B additionally take the deviation from 1894 values for each group.

Figure A.4: Annual event-study estimates

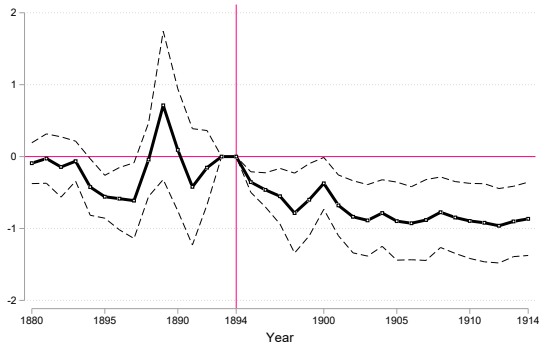
(a) Diphtheria, 1894 omitted



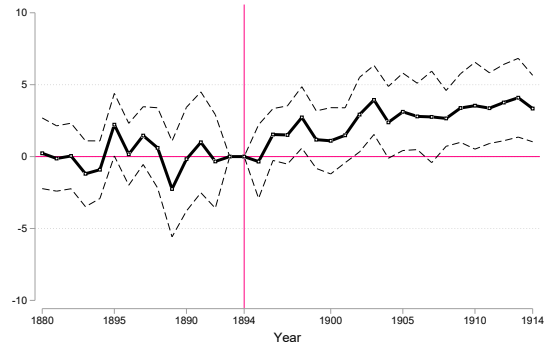
(b) Life expectancy at age 1, 1894 omitted



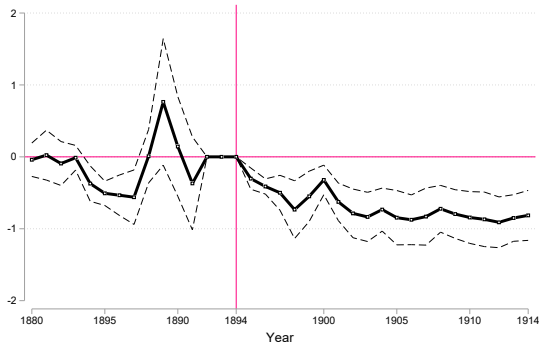
(c) Diphtheria, 1893 & 1894 omitted



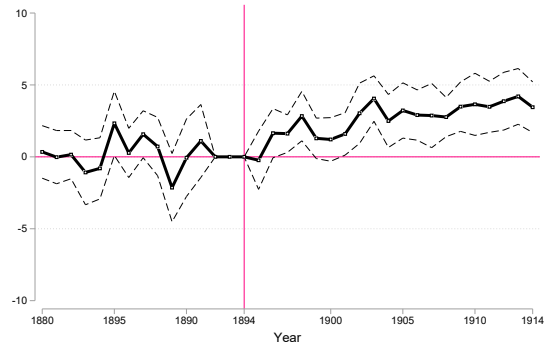
(d) Life expectancy at age 1, 1893 & 1894 omitted



(e) Diphtheria, 1892-1894 omitted



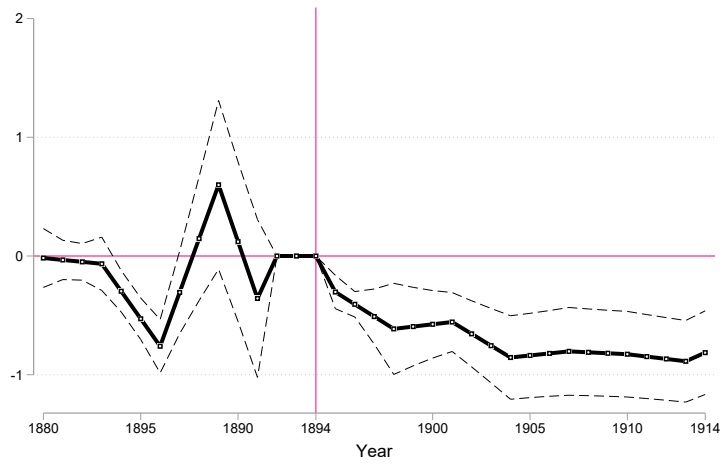
(f) Life expectancy at age 1, 1892-1894 omitted



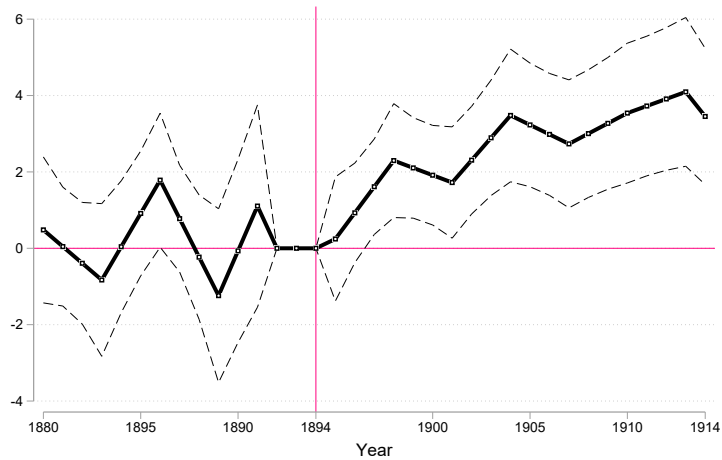
Notes: This figure shows reduced-form event study estimates for the diphtheria mortality rate, in Panels (a), (c) and (e), and life expectancy at age 1 in Panel (b), (d), (f). The event dummies are annual, contrast to the baseline, which uses a three-year pooling. The omitted years are 1894 (in Panels a and b), 1893 and 1894 (in Panels c and d), and 1892-1893 (in Panels e and f). All regressions control for municipality and year fixed effects and weighted by the 1895 municipality population size. The dashed curves are 95 percent confidence bands based standard errors clustered at the municipality level. The blue dashed curves (in Panels a and b) are confidence bands, allowing for linear smoothness (i.e., linear pre-trends) according to Rambachan and Roth (2023).

Figure A.5: Spline event-study estimates

(a) Diphtheria

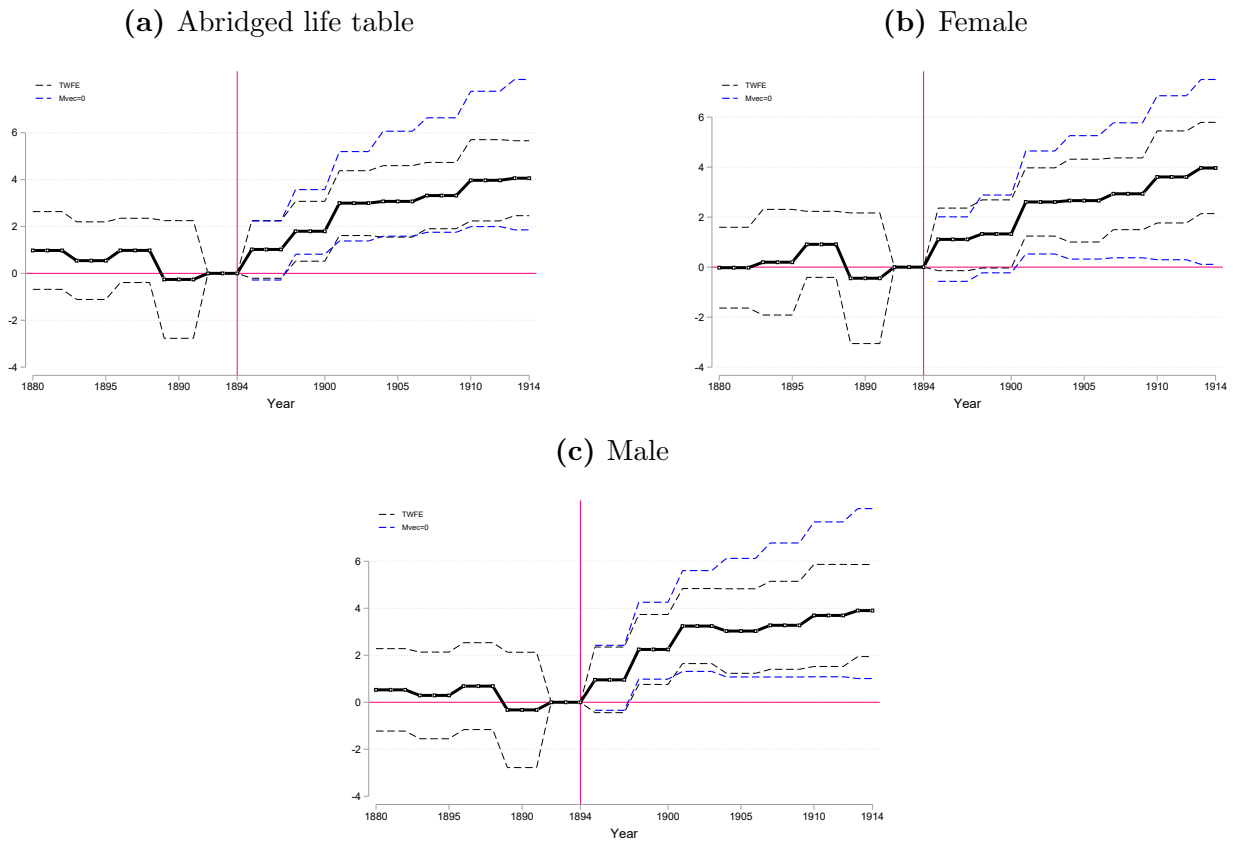


(b) Life expectancy



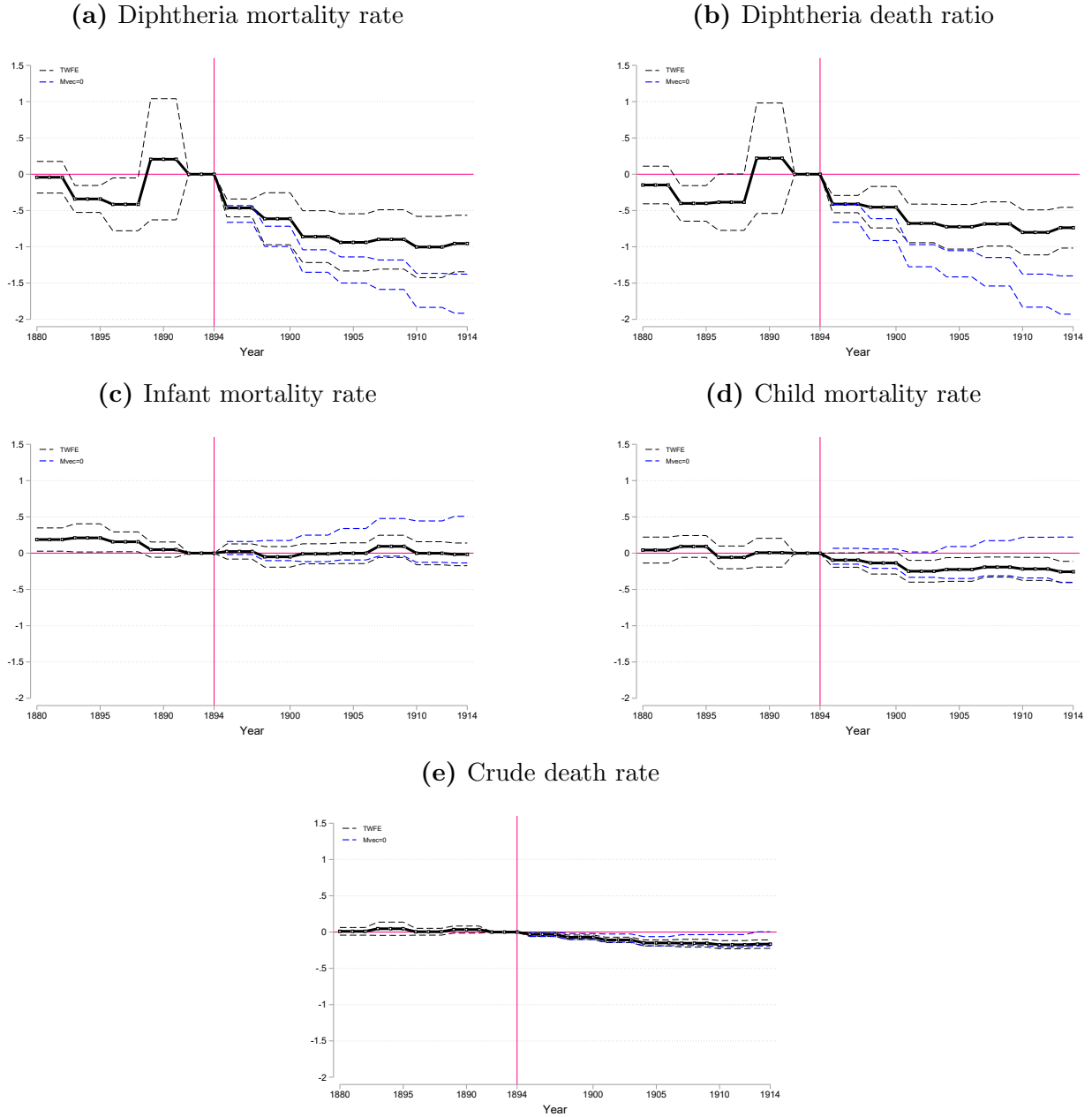
Notes: This figure shows reduced-form event-study estimates for the diphtheria mortality rate, in Panel (a), and life expectancy at age 1 in Panel (b). We use a three-year spline model. The regressions control for municipality and year fixed effects and weighted by the 1895 municipality population size. The dashed lines are 95 percent confidence bands based standard errors clustered at the municipality level.

Figure A.6: Event-study estimates, additional life expectancy outcomes



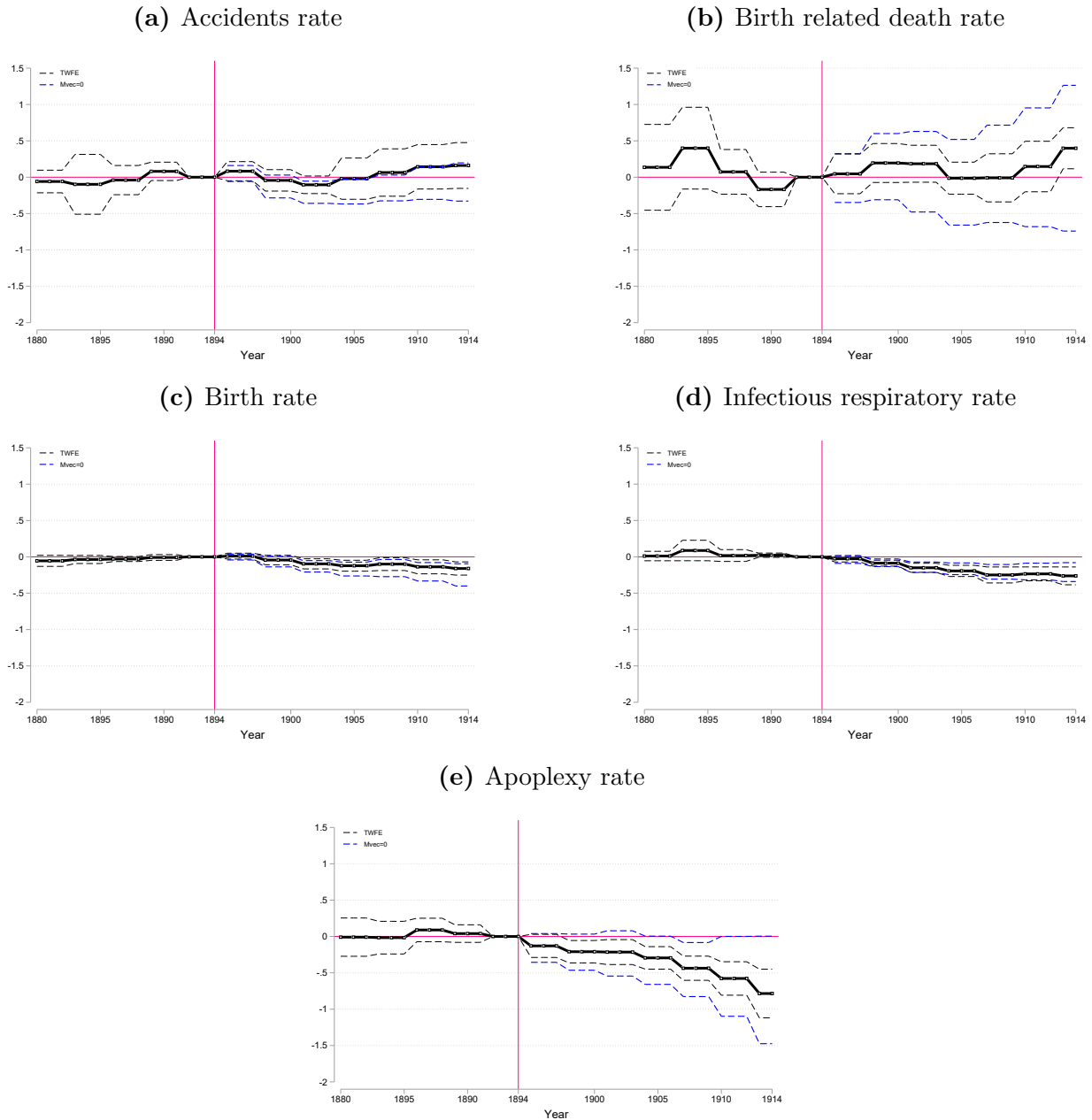
Notes: This figure shows reduced-form event-study estimates for additional life expectancy outcomes. In Panel (a), the outcome is life expectancy at age 1 derived from the abridged life table; in Panel (b) the outcome is female life expectancy at age 1; in Panel (c) the outcome is male life expectancy at age 1. In all regressions, we use a three-year pooling model, except for 1913 and 1914, which is two years. The omitted three years are 1892-1894. The regressions control for municipality and year fixed effects and weighted by the 1895 municipality population size. The dashed curves are 95 percent confidence bands based standard errors clustered at the municipality level. The blue dashed curves are confidence bands, allowing for linear smoothness (i.e., linear pre-trends) according to Rambachan and Roth (2023).

Figure A.7: Event-study estimates (relative to pre-antitoxin outcome), additional mortality outcomes I



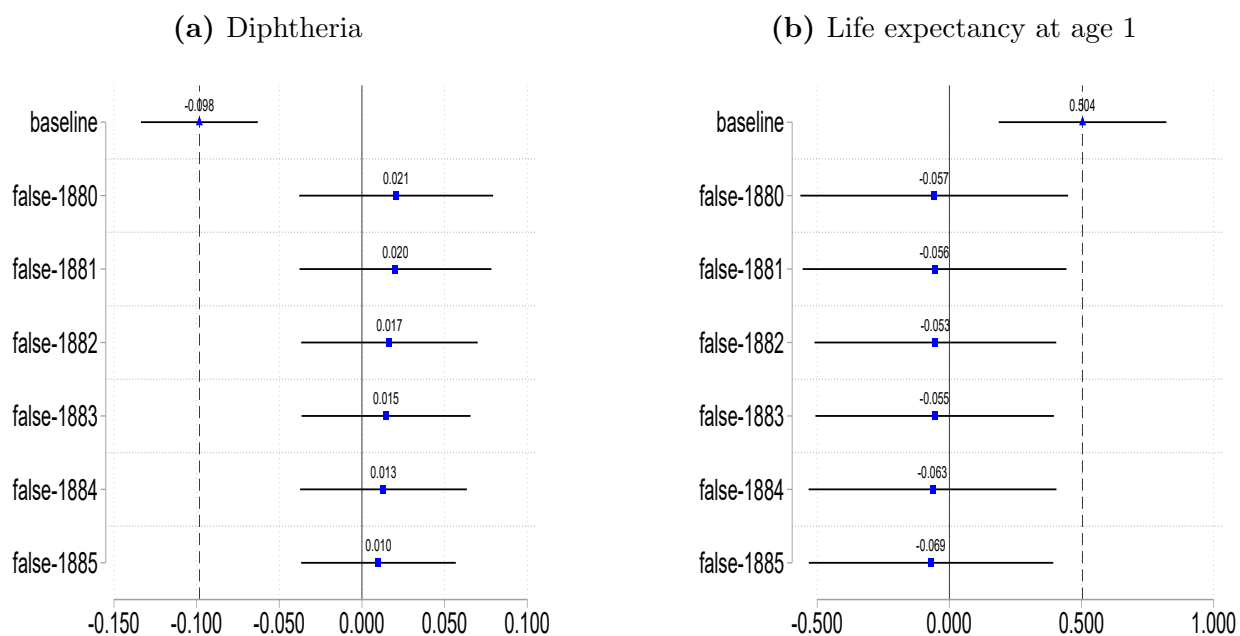
Notes: This figure shows reduced-form event-study estimates for additional mortality outcomes. All estimates have been rescaled by the size of the pre-antitoxin outcome (i.e., the average from 1880 to 1894). Thus, one can interpret the estimates as percent changes. For comparison, Panel (a) shows the percent reductions in the diphtheria mortality rate. The different outcomes are described in the sub-figure title. In all regressions, we use a three-year pooling model, except for 1913 and 1914, which is two years. The omitted three years are 1892-1894. The regressions control for municipality and year fixed effects and weighted by the 1895 municipality population size. The dashed curves are 95 percent confidence bands based standard errors clustered at the municipality level.

Figure A.8: Event-study estimates (relative to pre-antitoxin outcome), additional mortality outcomes II



Notes: This figure shows reduced-form event-study estimates for additional mortality outcomes. All estimates have been rescaled by the size of the pre-antitoxin outcome (i.e., the average from 1880 to 1894). Thus, one can interpret the estimates as percent changes. For comparison, Panel (a) shows the percent reductions in the diphtheria mortality rate. The different outcomes are described in the sub-figure title. In all regressions, we use a three-year pooling model, except for 1913 and 1914, which is two years. The omitted three years are 1892-1894. The regressions control for municipality and year fixed effects and weighted by the 1895 municipality population size. The black dashed curves are 95 percent confidence bands based standard errors clustered at the municipality level. The blue dashed curves are confidence bands, allowing for linear smoothness (i.e., linear pre-trends) according to Rambachan and Roth (2023).

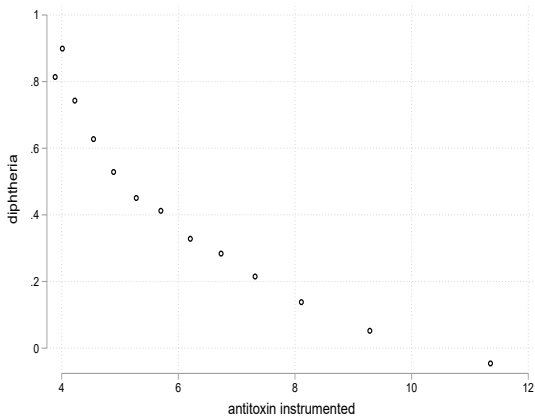
Figure A.9: False antitoxin start dates



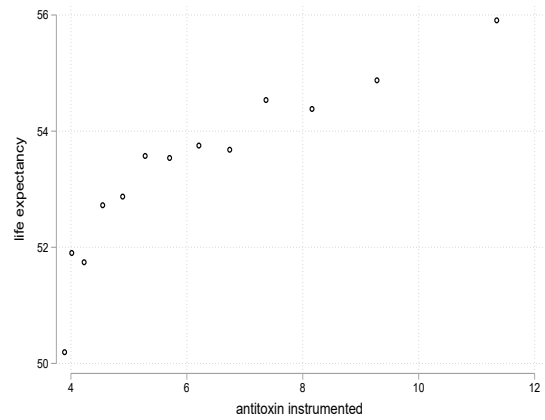
Notes: This figure reports 2SLS estimates when assuming false start dates for the free distribution of antitoxin. In Panel (a), the outcome is the diphtheria mortality rate, and in Panel (b), the outcome is life expectancy at age 1. Both outcomes are measured for the years 1880 to 1896, and in 'false-1880', we assume that the distribution of antitoxin starts in 1880 and in 'false-1881' the start year is 1881, etc. For comparison, the baseline estimates are also reported as 'baseline'. All regressions are weighted by the municipality population size in 1895 and control for municipality and county-by-year fixed effects. Standard errors account for arbitrary heteroskedasticity and are clustered at the municipality level. The horizontal black solid lines are 95 percent confidence bands.

Figure A.10: Binscatter plots

(a) Diphtheria

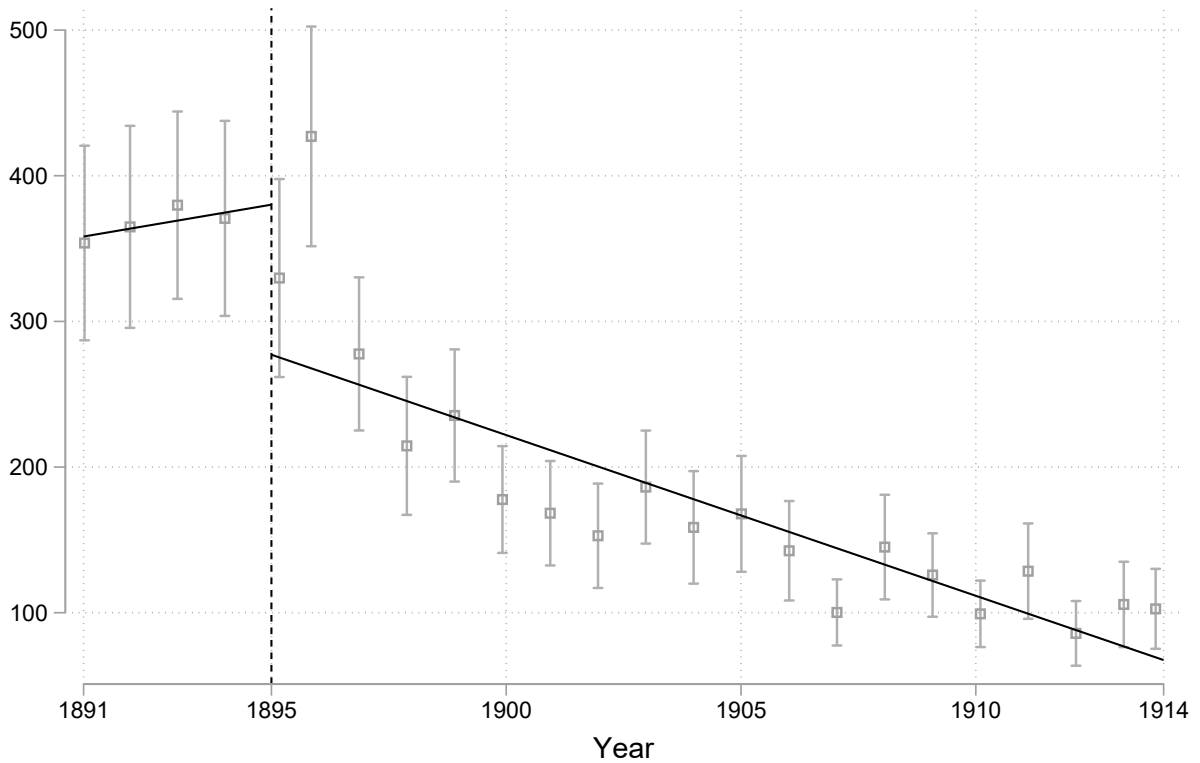


(b) Life expectancy at age 1



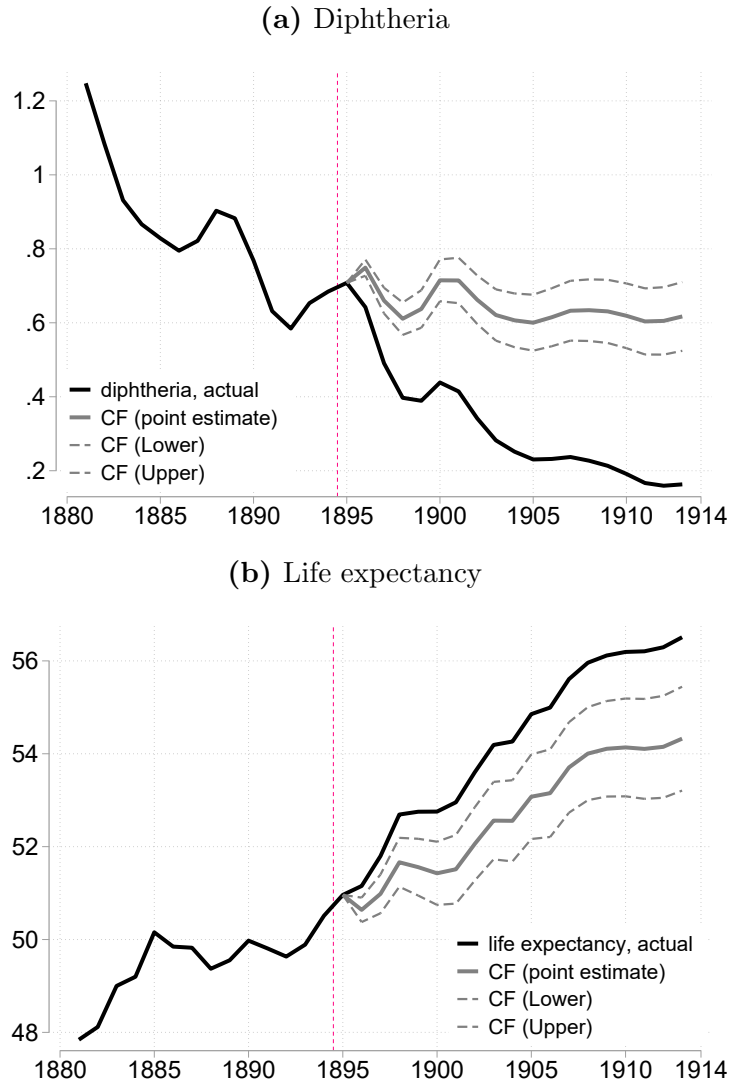
Notes: This figure shows the relationship between the two outcomes and antitoxin instrumented in binned scatterplots, using the residualization procedure as described in Cattaneo et al. (2024) to control for the baseline fixed effects.

Figure A.11: Trends in the diphtheria case-fatality rate



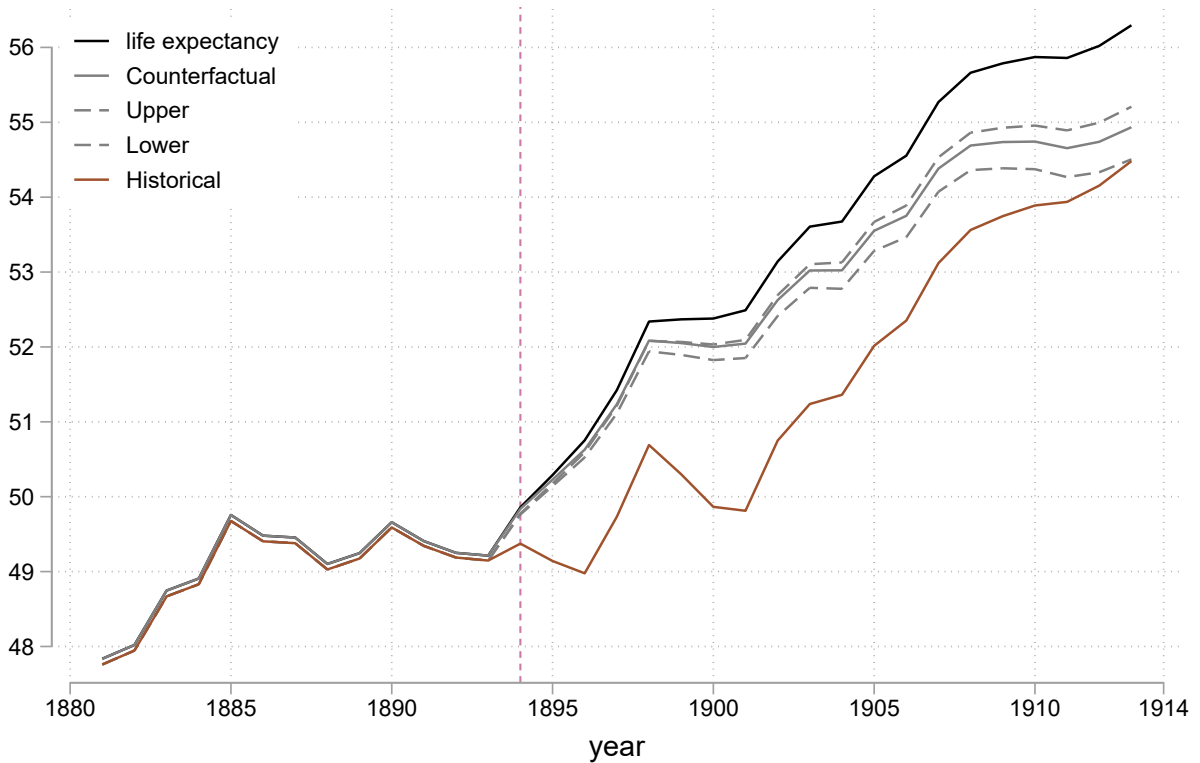
Notes: This figure shows the development of the case-fatality rate, defined as the number of diphtheria deaths per 1,000 cases. Case data are available only from 1891, and the panel is unbalanced. Observations exceeding 1,000 have been capped at 1,000. The gray vertical lines represent 95% confidence intervals.

Figure A.12: Robustness to counterfactual when delaying the policy



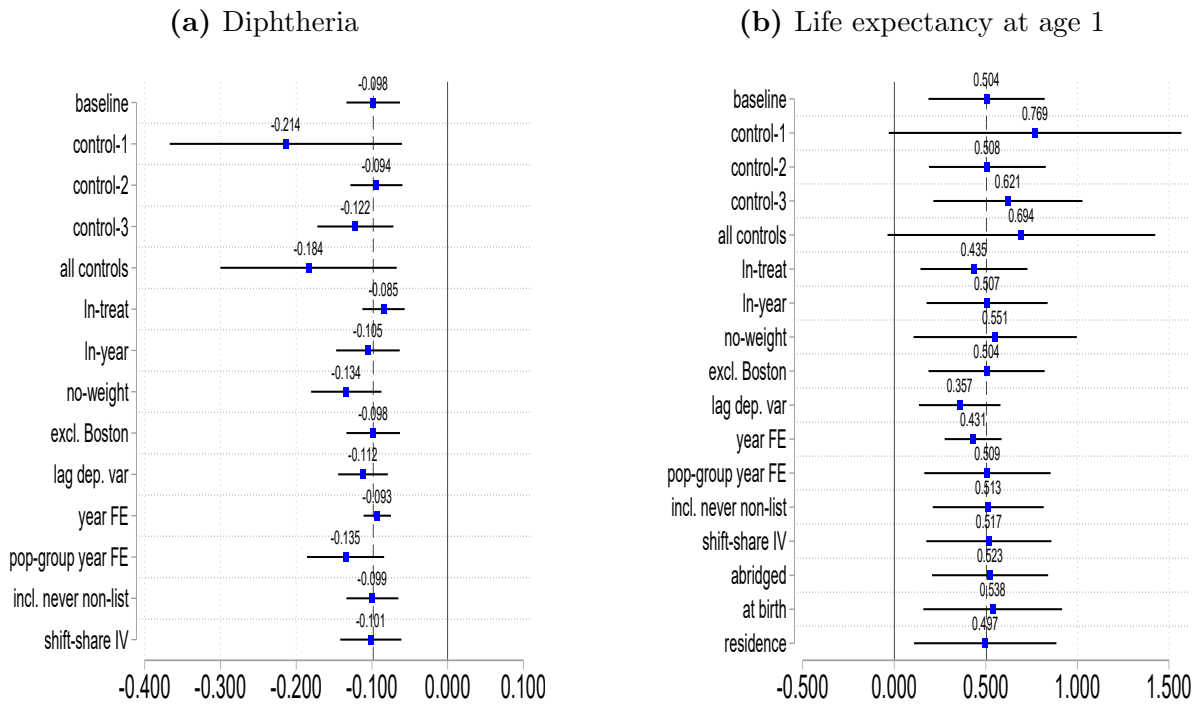
Notes: This figure presents reduced-form estimates using logged year instead of linear year to calculate the counterfactual (CF) trends for the diphtheria mortality rate (Panel A) and life expectancy at age 1 (Panel B). The CF trends, based on annual delays in average treatment intensity, are shown as gray solid lines, with dashed lines representing 95% confidence intervals. The solid black lines depict the observed population-weighted averages of the outcomes. All curves are smoothed using three-year moving averages.

Figure A.13: Plausibility of magnitudes



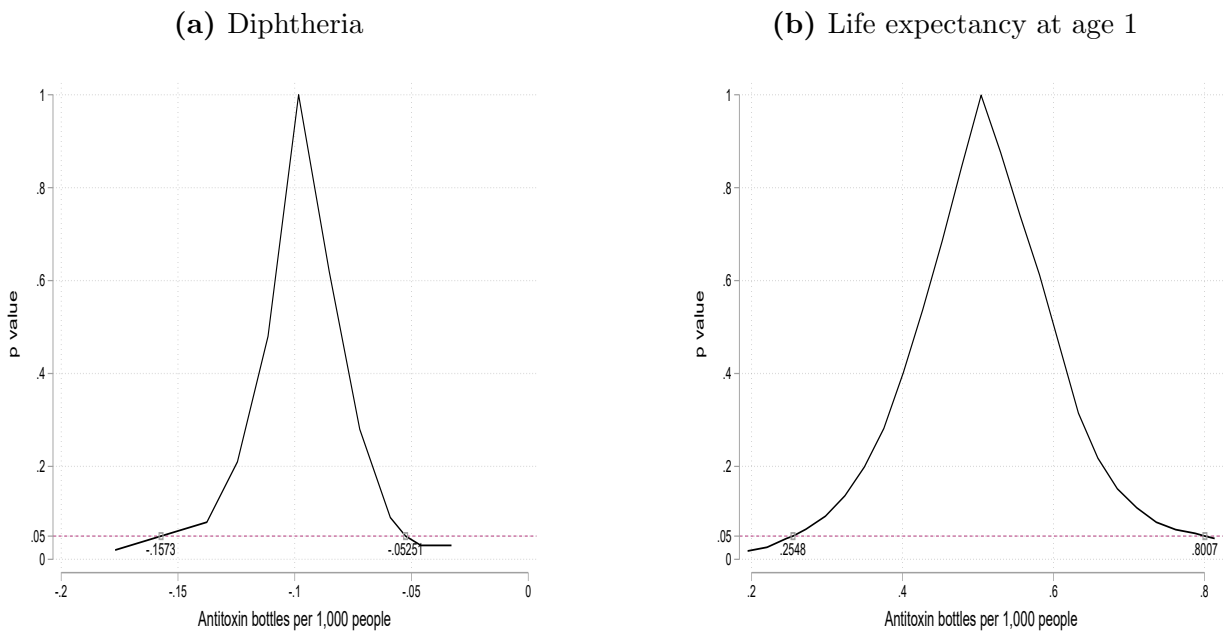
Notes: This figure performs plausibility checks based on the counterfactual analysis for diphtheria shown in Panel (a) of Figure 5. We calculate the number of lives saved by comparing the observed diphtheria mortality rate to the counterfactual rate. We then assume these saved lives were evenly distributed across ages 1 to 5 and subtract these figures from the aggregate life table for Massachusetts. The historical counterfactual is based on the number of lives saved according to SBH report from 1901. This report provides the number of lives saved due to antitoxin for the years 1895 and 1901. For the remaining years, we use the average of lives saved for the years 1899, 1900, and 1901.

Figure A.14: Robustness checks



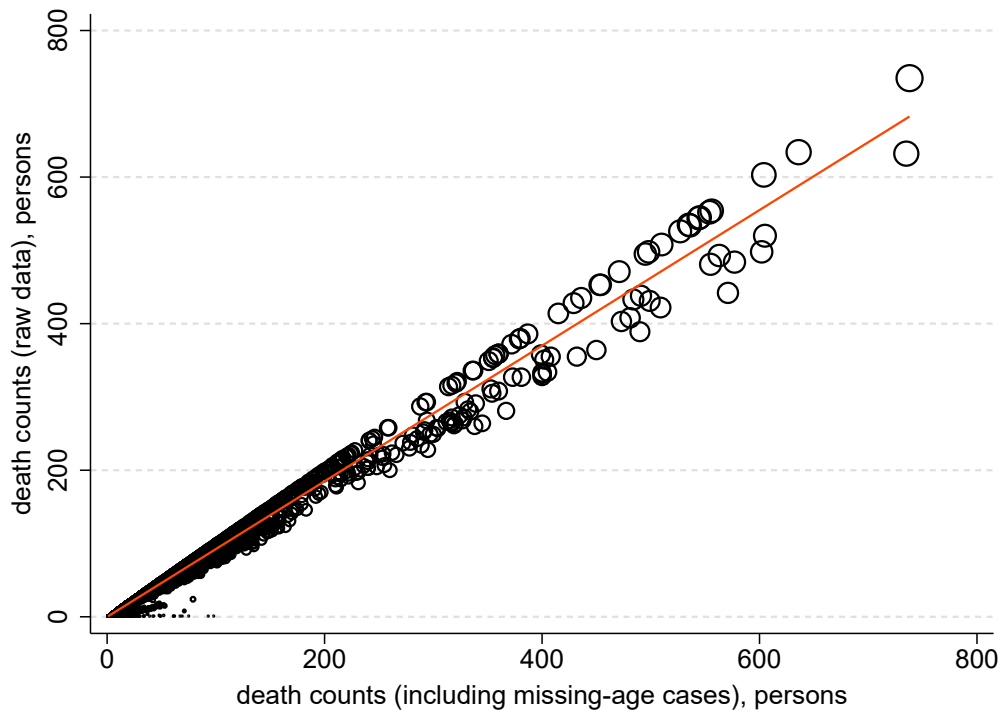
Notes: This figure presents 2SLS estimates of the baseline model subjected to various robustness checks. The top row highlights the baseline estimates for the diphtheria mortality rate (Panel a) and life expectancy at age 1 (Panel b). The specification 'control-1' includes controls for all pre-antitoxin variables listed in Table 1, along with year fixed effects. 'Control-2' adds population shares by age groups (0, 1-4, 5-9, and 10-14), while 'control-3' incorporates controls for the roll-out of public water works, infectious disease hospitals, and general hospitals. The 'all controls' specification combines all these controls in one model. The 'ln-treat' specification uses logged treatment levels instead of levels in the baseline model. The 'ln-year' specification replaces the linear year variable with the natural logarithm of years since the policy was implemented. The 'no-weight' specification removes the weighting by running unweighted regressions. The 'excl. Boston' specification excludes the municipality of Boston. The 'lag. dep. var' specification includes up to five lags of the dependent variable. The 'year FE' specification replaces county-year fixed effects with year fixed effects, while the 'pop-group year FE' specification includes interactions between population size quartile groups from 1895 and year fixed effects. The 'incl. never non-list' specification includes never non-listed municipalities in terms of antitoxin bottles as zeros instead of dropping them from the sample. The 'shift-share IV' specification employs the shift-share instrument instead of the baseline IV. The 'abridged' and 'at birth' specifications provide estimates using life expectancy at age 1 from the abridged life table and life expectancy at birth, respectively. The 'residence' specification uses place of residence rather than place of occurrence (due to data limitations, the sample here ends in 1905). All regressions are weighted by the municipality population size in 1895 (except for the 'no-weight' specification) and include controls for municipality and county-by-year or year fixed effects. Standard errors are robust to heteroskedasticity and clustered at the municipality level. Horizontal lines represent 95% confidence intervals

Figure A.15: Confidence intervals when clustering at the county level



Notes: This figure presents the distribution of coefficient estimates from the baseline model, using wild-cluster bootstrapping with 100 replications and clustering at the county level. Panel (a) shows results with diphtheria mortality as the outcome, and Panel (b) uses life expectancy at age 1. Given that Massachusetts has only 14 counties, standard errors clustered at the county level may be unreliable, and inference is therefore based on wild-cluster bootstrapped confidence intervals. The graphs plot the empirical distribution of the bootstrapped coefficients for the treatment variable (antitoxin bottles per 1,000 people). These figures demonstrate that the estimates are statistically significant even under conservative inference with clustered standard errors at the county level.

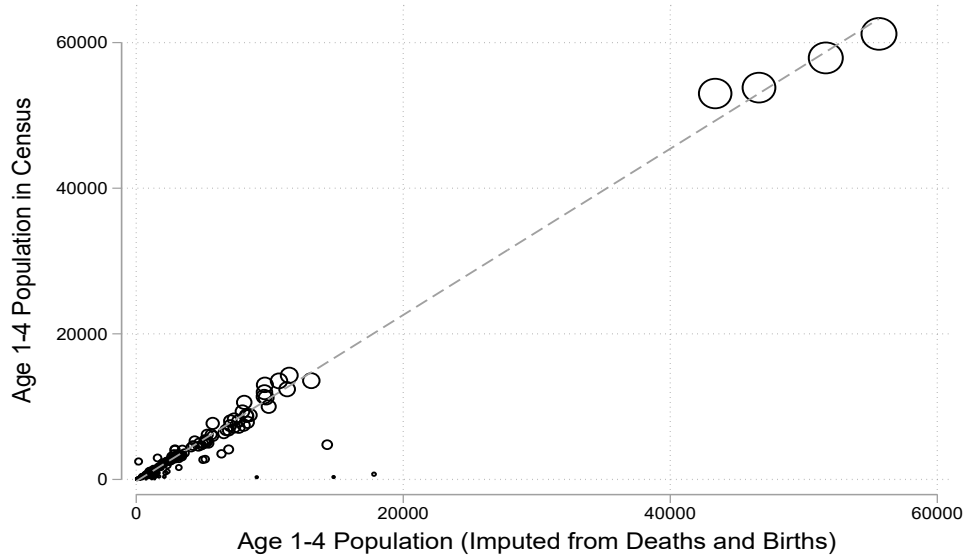
Figure A.16: Imputed and Original Death Counts



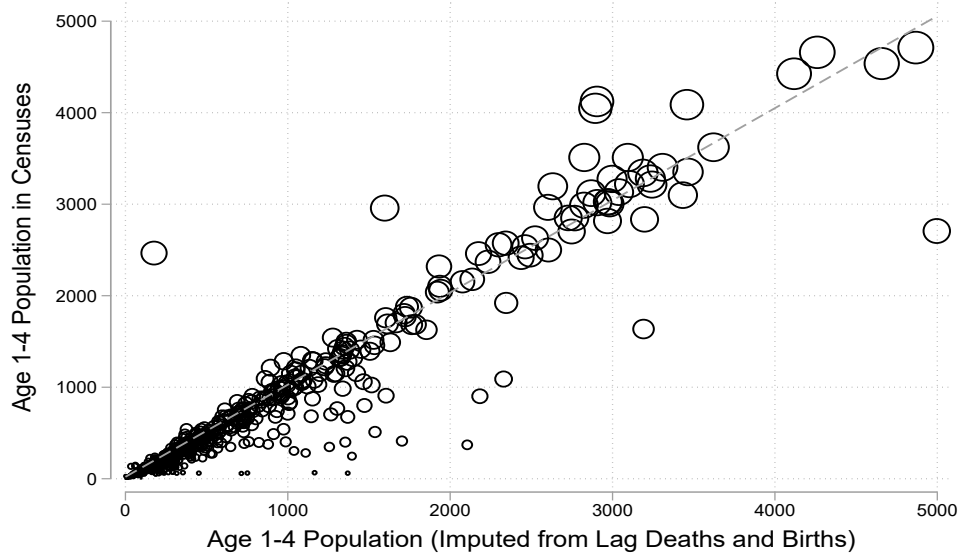
Notes: This figure presents scatterplots of imputed death counts, which include redistributed death cases with missing-age (*x-axis*) and the death counts excluding redistributed death cases (*y-axis*). Each scatterplot represents a year-municipality observation between 1895 and 1915. Size of scatterplots represent the imputed death counts, and the fitted line comes from a bivariate regression weighted by the imputed death counts.

Figure A.17: Imputed and Census Reported Population Age 1-4

(a) Age 1-4 Population by Source



(b) Age 1-4 Population by Source (under 5,000 People)



Notes: This figures present scatterplots of imputed population aged 1 to 4 (x-axis) and census-reported population aged 1 to 4 (y-axis) in the census years of 1895, 1900, 1905, 1910, and 1915. Census-reported populations come from Federal censuses 1900 and 1910; and Massachusetts State census in 1895, 1905, and 1915. Each scatter-plot represents an observation of municipality and census year. Size of scatterplots represents the population size, and the fitted line is from a regression weighted by imputed population size

Additional Tables

Table A.1: Summary Statistics by Pre- and Post-antitoxin Periods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	<i>Pre-antitoxin period: 1880-1894</i>					<i>Post-antitoxin period: 1895-1914</i>				
	N	mean	p25	p50	p75	N	mean	p25	p50	p75
diphtheria rate	4,230	0.87	0.35	0.74	1.25	5,640	0.32	0.13	0.24	0.39
life exp. at age 1	4,200	49.35	44.37	48.71	52.90	5,600	54.33	50.92	53.88	57.68
life exp. at age 5	4,200	50.86	47.08	50.46	53.15	5,600	54.24	51.80	53.98	56.79
life exp. at age 1 (abridged)	4,230	49.27	44.54	48.74	52.99	5,639	53.99	50.49	53.65	57.40
infant mortality rate	4,207	169.97	139.65	170.28	191.49	5,604	158.79	134.48	153.06	177.92
child mortality rate	4,210	26.56	18.00	26.22	33.89	5,626	16.08	10.82	15.61	20.38
antitoxin p.c.	4,230	0.00	0.00	0.00	0.00	5,640	9.50	2.01	9.31	17.62
treatment	4,230	0.74	0.48	0.69	1.16	5,640	0.74	0.48	0.69	1.16
population	4,230	7,223	1,334	2,452	4,763	5,640	10,610	1,450	2,695	6,882

Notes: This table reports summary statistics for selected key variables averaged over the pre-antitoxin period (1880 to 1894) in columns (1)-(5) and the post-antitoxin periods (1895 to 1914) in columns (6)-(10), using the 1895 municipality population size as weight.

Table A.2: Antitoxin adoption by municipality characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
treatment \times I \times (t-1894)	0.349*** (0.058)	0.283*** (0.058)	0.263*** (0.056)	0.153*** (0.049)	0.162*** (0.046)	0.173*** (0.046)
infec. rate, 88-94 \times I \times (t-1894)		0.026** (0.010)	0.032*** (0.011)	-0.021 (0.015)	-0.022 (0.015)	-0.022 (0.015)
apoplexy rate, 88-94 \times I \times (t-1894)		-0.088 (0.058)	-0.120** (0.059)	-0.005 (0.060)	0.026 (0.059)	0.023 (0.060)
doctors pr. capita, 95 \times I \times (t-1894)			0.098*** (0.032)	0.107*** (0.032)	0.093*** (0.030)	0.097*** (0.030)
dist Boston \times I \times (t-1894)				-0.004*** (0.001)	-0.003* (0.001)	-0.003** (0.001)
persons pr. 1k sqm, 95 \times I \times (t-1894)				0.028*** (0.007)	0.020** (0.009)	0.019** (0.009)
persons pr. dwell., 95 \times I \times (t-1894)				0.066** (0.033)	0.023 (0.030)	0.025 (0.030)
persons pr. room, 95 \times I \times (t-1894)				0.164 (0.384)	0.416 (0.360)	0.428 (0.348)
fb share, 95 \times I \times (t-1894)				0.023 (0.359)	0.275 (0.362)	0.258 (0.370)
water intervention					-0.504* (0.289)	-0.408 (0.281)
infec. hosp. intervention					2.102** (0.912)	2.131** (0.906)
general hosp. intervention					0.683** (0.331)	0.660* (0.343)
age composition	No	No	No	No	No	Yes
$N \times T$	9870	9870	9800	9800	9800	9800
N	282	282	280	280	280	280

Notes: This table reports the correlation between the number of antitoxin bottles per 1,000 people and selected municipality characteristics. The key variable, “treatment”, is the average diphtheria mortality rate over 1889–1894. “Infect. rate” is the total mortality rates for eight infectious diseases (see data appendix), and “apoplexy rate” is the apoplexy (sudden death) mortality rate, both averaged over 1889–1894. “doctors pc.” is doctors per 1,000 people in 1895; “dist. Boston” is the aerial distance to Boston; “pop. density” is population per 1,000 sq. miles in 1895; “dwell. size” and “room size” are the number of people per dwelling and per room in 1895, respectively; and “immigrant” is the share of foreign-born individuals in 1895. These pre-antitoxin period characteristics are interacted with a post-1895 dummy (I) and a linear trend (t-1894). The vector of age composition includes percentage of population at age 0, 1-4, 5-9, and 10-14. All regressions control for municipality and county-by-year fixed effects, and are weighted by the 1895 population. Standard errors (reported in parentheses) are robust and clustered at the municipality level. ***, **, and * indicate significance at 1%, 5%, and 10% levels.

Table A.3: Decomposing the change of antitoxin adoption by treatment into municipality covariates

	(1)	(2)	(3)	(4)	(5)
	Base	Model-1	Explained	Model-2	Explained
treatment \times I \times (t-1894)	0.349 (0.058)	0.153 (0.049)	0.195 (0.044)	0.173 (0.046)	0.176 (0.044)
Municipality Covariates:					
infect. rate \times I \times (t-1894)	No	Yes	-0.040 (0.031)	Yes	-0.042 (0.030)
apoplexy rate \times I \times (t-1894)	No	Yes	0.001 (0.011)	Yes	-0.004 (0.011)
doctors pc. \times I \times (t-1894)	No	Yes	0.002 (0.016)	Yes	0.002 (0.014)
dist. Boston \times I \times (t-1894)	No	Yes	0.016 (0.013)	Yes	0.012 (0.010)
pop. density \times I \times (t-1894)	No	Yes	0.083 (0.031)	Yes	0.057 (0.028)
dwel. size \times I \times (t-1894)	No	Yes	0.111 (0.063)	Yes	0.043 (0.051)
room size \times I \times (t-1894)	No	Yes	0.020 (0.046)	Yes	0.053 (0.042)
immigrant \times I \times (t-1894)	No	Yes	0.003 (0.040)	Yes	0.029 (0.042)
water intervention	No	No		Yes	0.004 (0.003)
infectious hospital	No	No		Yes	0.029 (0.024)
general hospital	No	No		Yes	0.005 (0.005)
age composition	No	No		Yes	-0.012 (0.006)

Notes: This table decomposes the changes in the coefficients of the “treatment”, which is interacted with a post-1895 dummy and a linear time trend (t-1894), as reported in Table A.2. Column (1) reports the coefficient on the treatment variable in a base model without any municipality-level covariates (Table A.2, column 1). Column (2) shows the coefficient on the treatment variable in the model including pre-antitoxin municipality covariates, interacted with a post-1895 dummy and a linear time trend (Table A.2, column 4). Column (3) decomposes the coefficient change between column (1) and column (2) into municipality covariates, based on Gelbach (2016). Column (4) provides the coefficient on the treatment variable in a full model controlling for additional municipality covariates, including the presence of a city sanitation system, infectious disease hospitals, general hospitals, and the age composition of the population (Table A.2, column 6). Column (5) decomposes the coefficient change between the base model (Column 1) and the full model (Column 4). Standard errors are clustered at the municipality level and are reported in parentheses.

Table A.4: Decomposing the change of estimated effect of antitoxin distribution on public health into municipality covariates

	Diphtheria Mortality Rate			Life Expectancy at Age 1		
	Base (1)	Full (2)	Explained (3)	Base (4)	Full (5)	Explained (6)
antitoxin p.c.	-0.098 (0.018)	-0.183 (0.058)	0.085 (0.045)	0.504 (0.161)	0.676 (0.364)	-0.172 (0.261)
Municipality Covariates:						
infect. rate \times I \times (t-1894)	No	Yes	-0.021 (0.017)	No	Yes	0.217 (0.114)
apoplexy rate \times I \times (t-1894)	No	Yes	-0.003 (0.006)	No	Yes	0.074 (0.042)
doctors pc. \times I \times (t-1894)	No	Yes	0.001 (0.009)	No	Yes	-0.011 (0.068)
dist. Boston \times I \times (t-1894)	No	Yes	0.005 (0.005)	No	Yes	-0.032 (0.034)
pop. density \times I \times (t-1894)	No	Yes	0.023 (0.019)	No	Yes	-0.135 (0.086)
dwel. size \times I \times (t-1894)	No	Yes	0.021 (0.028)	No	Yes	-0.038 (0.176)
room size \times I \times (t-1894)	No	Yes	0.030 (0.031)	No	Yes	-0.250 (0.206)
immigrant \times I \times (t-1894)	No	Yes	0.019 (0.023)	No	Yes	0.035 (0.122)
water intervention	No	Yes	0.002 (0.002)	No	Yes	-0.002 (0.013)
infectious hospital	No	Yes	0.016 (0.013)	No	Yes	-0.073 (0.068)
general hospital	No	Yes	0.002 (0.002)	No	Yes	0.015 (0.015)
age composition	No	Yes	-0.011 (0.005)	No	Yes	0.026 (0.028)

Notes: *This table decomposes the change in the causal effect of antitoxin distribution on public health into a list of municipality covariates, based on the approach by Gelbach (2016). The causal effects are estimated using the instrumental variable approach, as described in the main text. The estimates are visualized in Figure A.14. Column (1) reports the 2SLS estimate of the effect of antitoxin distribution on the diphtheria mortality rate in a model that excludes municipality covariates (see also Figure A.14, panel (a), baseline). Column (2) reports the estimated causal effect in a full specification that includes a comprehensive list of municipality covariates specified in Table A.2, column (6) (see also Figure A.14, panel (a), all controls). Column (3) decomposes the difference in estimates between columns (1) and (2) into all municipality covariates. Similarly, columns (4)-(6) decompose the change in estimated causal effects of antitoxin diffusion on life expectancy at age 1. Standard errors are reported in parentheses and are clustered at the municipality level.*

Table A.5: Effects on Age-specific Mortality Rates

	(1) age 0	(2) age 1	(3) age 2	(4) age 3	(5) age 4	(6) age 5	(7) age 6	(8) age 7	(9) age 8	(10) age 9	(11) age 10
antitoxin p.c.	-3.12* (1.73)	-1.49 (0.97)	-1.21*** (0.31)	-0.85*** (0.23)	-0.75*** (0.20)	-0.58*** (0.16)	-0.39*** (0.14)	-0.28** (0.12)	-0.16* (0.10)	-0.24** (0.11)	-0.07 (0.09)
Mean pre-y	188.90	55.42	24.77	17.21	13.14	10.29	8.12	6.93	5.63	4.75	3.96
$N \times T$	9800	9800	9800	9800	9800	9799	9800	9800	9800	9800	9800
N	280	280	280	280	280	280	280	280	280	280	280
KP-F-Stat	35.96	35.96	35.96	35.96	35.96	35.95	35.96	35.96	35.96	35.96	35.96
AR 95 CI	[-6.9, .2]	[-3.8, .2]	[-2.0, -.7]	[-1.4, -.5]	[-1.2, -.4]	[-1.0, -.3]	[-.7, -.1]	[-.6, -.1]	[-.4, .01]	[-.5, -.1]	[-.3, .1]
tf 95 CI	[-6.9, .7]	[-3.6, .6]	[-1.9, -.5]	[-1.4, -.4]	[-1.2, -.3]	[-.9, -.2]	[-.7, -.1]	[-.5, -.01]	[-.4, .1]	[-.5, -.0]	[-.3, .1]

Notes: This table reports the effects on q -type age-specific mortality rates (ages 0 to 10) using the linear trend-break model as outlined in Equation (2). The outcomes are expressed per 1,000 individuals of the relevant age group. The sample includes the years 1880 to 1914. All regressions are weighted by the municipality population size in 1895 and control for municipality and county-by-year fixed effects. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level. KP-F-Stat is Kleibergen-Paap F statistic, AR 95 CI is the Anderson-Rubin 95% confidence intervals (Anderson and Rubin 1949) and tf 95% confidence intervals (Lee et al. 2022).

Table A.6: Stacked mortality model

	(1)	(2)	(3)	(4)	(5)
	mortality	mortality	mortality	mortality	mortality
antitoxin p.c. \times I	-0.085*** (0.017)	-0.130*** (0.047)	-0.091*** (0.017)	-0.064*** (0.014)	-0.066*** (0.014)
Controls	all	exogenous	childhood	declining	waterborne
$N \times T$	128,310	19,740	39,480	59,220	29,610
N	282	282	282	282	282
KP-F-Stat	36.10	36.11	36.10	36.10	36.10

Notes: This table reports 2SLS estimated from a stacked model that resembles the baseline model, but the panel is now three-dimensional (municipality-year-disease). We interact the main RHS variables in Equations (1) and (2) with an indicator for diphtheria. Column 1 includes 12 control diseases (typhoid, tuberculosis, pneumonia, scarlet fever, measles, whooping cough, bronchitis, accidents, childbirth, meningitis, apoplexy, and digestive diseases). Column 2 only includes “exogenous” causes as controls (accidents). Column 3 only includes childhood diseases as controls (scarlet fever, whooping cough, measles). Column 4 only includes diseases where we also observe secular declines during the pre-antitoxin period as controls (typhoid, tuberculosis, scarlet fever, meningitis, and digestive diseases). Column 5 only includes waterborne diseases as controls (typhoid and digestive diseases). All regressions are weighted by the municipality population size in 1895 and control for municipality-by-year, disease-by-year-by-county, and municipality-by-disease fixed effects. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table A.7: Effects on the health-care sector

	(1)	(2)	(3)	(4)	(5)
age groups	all	20-29	30-39	40-49	50-59
Panel A: Doctors					
antitoxin p.c.	-0.021 (0.022)	-0.010 (0.010)	-0.011 (0.008)	-0.009 (0.010)	0.003 (0.008)
Mean pre-y	1.639	0.273	0.484	0.357	0.280
AR 95 CI	[-.07, .02]	[-.03, .01]	[-.03, .00]	[-.03, .01]	[-.01, .02]
tf 95 CI	[-.07, .03]	[-.03, .01]	[-.03, .01]	[-.03, .01]	[-.02, .02]
Panel B: Nurses					
antitoxin p.c.	0.056 (0.106)	0.064 (0.079)	0.009 (0.018)	-0.010 (0.008)	0.001 (0.004)
Mean pre-y	0.100	0.047	0.032	0.010	0.005
AR 95 CI	[-.15, .28]	[-.08, .23]	[-.03, .05]	[-.03, .01]	[-.01, .01]
tf 95 CI	[-.18, .29]	[-.11, .24]	[-.03, .05]	[-.03, .01]	[-.01, .01]
Panel C: Pharmacists					
antitoxin p.c.	-0.009 (0.014)	-0.001 (0.007)	-0.007 (0.007)	0.003 (0.005)	-0.007 (0.004)
Mean pre-y	0.744	0.222	0.241	0.137	0.079
AR 95 CI	[-.04, .02]	[-.01, .01]	[-.02, .01]	[-.01, .01]	[-.02, .00]
tf 95 CI	[-.04, .02]	[-.02, .01]	[-.02, .01]	[-.01, .01]	[-.02, .00]
$N \times T$	98,00	98,00	98,00	98,00	98,00
N	280	280	280	280	280
KP-F-Stat	35.96	35.96	35.96	35.96	35.96

Notes: This table reports the impact of the antitoxin treatment on the number of doctors per 1,000 people (Panel A), the number of nurses per 1,000 people (Panel B), and the number of pharmacists per 1,000 people (Panel C). The top row indicates the corresponding age group (e.g., column 2 provides the number of doctors/nurses/pharmacists in the ages 20-29). The method of estimation is 2SLS using the baseline annual linear trend-break model as outlined in Equation (2). The sample includes the years 1880 to 1914. All regressions are weighted by the municipality population size in 1895 and control for municipality and county-by-year fixed effects. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level. KP-F-Stat is Kleibergen-Paap F statistic, AR 95 CI is the Anderson-Rubin 95% confidence intervals (Anderson and Rubin 1949) and tf 95% confidence intervals (Lee et al. 2022).

Table A.8: Cost effectiveness analysis

(1) Year	(2) SBH Period	(3) Account ex- penses (\$)	(4) Bottles (Q)	(5) Cost per bot- tle (\$)	(6) Market price per bot- tle (\$)	(7) Saved con- sum (\$)	(8) Patients treated (Q)	(9) Total cost (\$)	(10) Cost per life saved (anti- toxin only) (\$)	(11) in 2023 (\$)	(12) Cost per life saved (to- tal) (\$)	(13) in 2023 (\$)
1895	1895 Mar.-1896 Mar.		1724	.15	1.5	2326	345	1984	4.18	157	31.95	1196
1896	1896 Mar.-1897 Mar.		3219	.15	1.5	4344	644	3704	3.82	142.8	29.2	1092
1897	1897 Mar.-1898 Mar.		4668	.15	1.5	6299	934	5371	3.60	137	27.6	1044
1898	1898 Mar.-1899 Mar.		12491	.15	1.5	16856	2498	14371	7.07	268	54.02	2047
1899	1899 Mar.-1900 Mar.		31997	.15	1.5	43179	6399	36814	14.1	536	108	4097
1900	1900 Mar.-1901 Mar.		40211	.15	1.5	54263	8042	46264	14.46	541	111	4138
1901	1901 Mar.-1902 Mar.		53389	.15	1.5	72046	10678	61426	16.24	601	124	4591
1902	1902 Mar.-1903 Mar.		40211	.15	1.5	55596	8042	46413	10.82	395	81	2959
1903	1903 Mar.-1904 Mar.		33475	.16	1.57	47338	6695	38756	8.08	289	59.2	2118
1904	1904 Mar.-1904 Sep.	5858	22255	.26	1.59	29534	4451	28113	7.95	281	38.2	1347
1905	1904 Sep.-1905 Sep.	9302	47387	.20	1.57	65185	9477	56689	11.31	404	68.96	2463
1906	1905 Sep.-1906 Nov.	13067	70424	.19	1.61	100096	14085	83491	14.25	498	91.03	3181
1907	1906 Nov.-1907 Nov.	14491	64087	.23	1.68	93095	12817	78578	14.26	477	77.30	2586
1908	1907 Nov.-1908 Nov.	15000	94645	.16	1.64	140571	18929	109645	13.39	457	97.87	3343
1909	1908 Nov.-1909 Nov.	18783	90131	.21	1.63	127707	18026	108914	15.29	528	88.63	3062
1910	1909 Nov.-1910 Nov.	19992	92623	.22	1.70	137205	18525	112615	14.89	493	83.90	2776
1911	1910 Nov.-1911 Nov.	18544	96522	.19	1.70	145271	19304	115066	12.77	422	79.23	2621
1912	1911 Nov.-1912 Nov.	19969	82085	.24	1.73	122218	16417	102054	12.75	413	65.16	2112
1913	1912 Nov.-1913 Nov.	20377	96891	.21	1.77	151027	19378	117268	12.10	384	69.63	2210
1914	1913 Nov.-1914 Nov.	20969	118561	.18	1.79	190737	23712	139530	11.61	365	77.24	2429

Notes: Column (1) shows the calendar year, while column (2) lists the SBH production period associated with it. Column (3) reports the expenditure for antitoxin production based on SBH accounts, available only from 1904 onward. Column (4) provides the total number of bottles produced. Column (5) shows the cost per bottle: \$0.15 for 1901, adjusted for inflation in 1902-1903, and calculated by dividing columns (3) and (4) from 1904 onward. Column (6) lists market prices per bottle, based on SBH (1901) and adjusted for inflation starting in 1902. Column (7) estimates total consumer savings under the free distribution policy, calculated as the difference between columns (6) and (5), multiplied by column (4). Column (8) reports the number of patients treated, assuming five bottles per patient as per SBH (1901). Column (9) calculates total costs, combining production costs and an assumed additional \$5 per patient. Column (10) gives the cost per averted diphtheria death, considering only production costs and using a baseline counterfactual. Column (11) adjusts this figure to 2023 dollar values. Columns (12) and (13) provide the corresponding 2023 values based on total costs instead of production costs alone.

Table A.9: Replication of Table 5 by Gender and Parental Background

	(1) Boys	(2) Girls	(3) Parents Foreign- born	(4) Parents U.S. born	(5) Father Low Skill	(6) Father High Skill
Panel A. Dep. Var.: School Attending \leq 3 months						
Exposure	-0.005*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)	-0.006*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)
Obs.	151,594	150,767	1371,93	111,048	121,124	181,251
R-squared	0.043	0.042	0.047	0.041	0.050	0.038
Mean pre-Y	0.107	0.109	0.114	0.100	0.111	0.105
Panel B. Dep. Var. is School Attending \geq 8 months						
Exposure	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.002)	0.005*** (0.001)	0.006*** (0.001)
Obs.	151,594	150,767	137,193	111,048	121,124	181,251
R-squared	0.057	0.055	0.060	0.056	0.066	0.051
Mean pre-Y	0.780	0.773	0.770	0.785	0.771	0.781
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports how antitoxin exposure affected school attendance in 1900 by gender (columns 1-2), parental birthplace (columns 3-4), and the father's socioeconomic status (columns 5-6). The dependent variable is a dummy of whether a child aged 5-15 attended school for no more than three months in Panel A and for at least eight months in Panel B. The variable of interest, "Exposure", denotes the average number of antitoxin bottles per 1,000 people that a child during the first nine years was exposed to. All regressions include fixed effects for municipality and year of birth and a set of individual controls (except excluding the gender dummies columns 1-2) as outlined on page 37. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table A.10: Replication of Table 5 based on the linked sample

	(1) attends any school	(2)	(3) attends ≤ 3 months	(4)	(5) attends ≥ 8 months	(6)
Exposure	0.003 (0.003)	0.003 (0.003)	-0.005*** (0.001)	-0.005*** (0.001)	0.007*** (0.001)	0.007*** (0.001)
Obs.	247,950	247,920	178,974	178,936	178,974	178,936
R-squared	0.145	0.158	0.033	0.039	0.045	0.052
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind. Controls	No	Yes	No	Yes	No	Yes
Mean pre-Y	0.682	0.682	0.104	0.104	0.784	0.784

Notes: This table replicates Table 5 using the linked sample (1900 to 1940). The sample includes 5 to 15-year-old children who lived in Massachusetts in 1900 and can be linked to 1940 using the crosswalks from the Census Tree Project. The dependent variable is a dummy of whether a child attended school at all (columns 1-2); for no more than three months (columns 3-4); and for at least nine months (columns 5-6). The variable of interest, “Exposure”, denotes the average number of antitoxin bottles per 1,000 people that a child during the first nine years was exposed to. The controls are the same as in Table 5. Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table A.11: Replication of Table 6 by Gender

	(1)	(2)	(3)	(4)	(5)	(6)
	years in school	low-skill occ.	blue-collar occ.	white- collar occ.	ln(occscore)	ln(wages)
Panel A. Men						
Exposure	0.0143** (0.0067)	0.0009 (0.0006)	-0.0010 (0.0008)	0.0001 (0.0006)	0.0011** (0.0005)	0.0021* (0.0012)
Obs.	139,294	144,724	144,724	144,724	134,692	109,278
R-squared	0.165	0.046	0.025	0.070	0.039	0.059
Mean pre-Y	9.952	0.307	0.206	0.398	3.325	7.236
Panel B. Women						
Exposure	-0.0019 (0.0047)	-0.0005 (0.0005)	-0.0002 (0.0002)	-0.0004 (0.0007)	-0.0024* (0.0013)	0.0015 (0.0026)
Obs.	99,430	103,142	103,142	103,142	31,249	26,020
R-squared	0.192	0.033	0.009	0.028	0.049	0.094
Mean pre-Y	10.289	0.112	0.007	0.183	3.061	6.715
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports how antitoxin exposure during childhood affected labor market outcomes as adults by gender. The sample includes 5 to 15-year-old children who lived in Massachusetts in 1900 and can be linked to 1940 using the crosswalks from the Census Tree Project. The following outcomes in 1940 are used as dependent variables: educational attainment (column 1), a dummy of whether the individual works in a low-skilled (column 2), blue-collar skilled (column 3), or white-collar occupation (column 4), the ln occupational income score (column 5), and ln wages (column 6). The variable of interest, “Exposure”, denotes the average number of antitoxin bottles per 1,000 people that a child during the first nine years was exposed to. All regressions control for municipality and year of birth fixed effects and a set of individual controls (see page 37 for details). Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table A.12: Replication of Table 6 by Migration Status

	(1) years in school	(2) low-skill occ.	(3) blue-collar occ.	(4) white- collar occ.	(5) ln(occscore)	(6) ln(wages)
Panel A. Stayer						
Exposure	-0.0016 (0.0052)	0.0006 (0.0007)	-0.0002 (0.0005)	-0.0004 (0.0008)	0.0005 (0.0007)	-0.0012 (0.0016)
Obs.	123,464	129,318	129,318	129,318	85,151	70,286
R-squared	0.181	0.107	0.115	0.079	0.110	0.134
Mean pre-Y	9.767	0.250	0.120	0.277	3.252	7.049
Panel B. Mover (Out of State)						
Exposure	0.0171*** (0.0065)	-0.0004 (0.0007)	-0.0019** (0.0008)	0.0017* (0.0010)	0.0013 (0.0014)	0.0062** (0.0028)
Obs.	57,360	58,396	58,396	58,396	41,242	32,644
R-squared	0.229	0.090	0.097	0.151	0.100	0.121
Mean pre-Y	10.581	0.189	0.129	0.370	3.310	7.271
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Year of Birth FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports how antitoxin exposure during childhood affected labor market outcomes as adults by mover status. The sample includes 5 to 15-year-old children who lived in Massachusetts in 1900 linked to 1940 using the crosswalks from the Census Tree Project. The following outcomes in 1940 are used as dependent variables: educational attainment (column 1), a dummy of whether the individual works in a low-skilled (column 2), blue-collar skilled (column 3), or white-collar occupation (column 4), the ln occupational income score (column 5), and ln wages (column 6). The variable of interest, “Exposure”, denotes the average number of antitoxin bottles per 1,000 people that a child during the first nine years was exposed to. All regressions control for municipality and year of birth fixed effects and a set of individual controls (see page 37 for details). Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table A.13: The Long-run Effects of Antitoxin Treatment – Different Periods

	(1) low-skilled	(2) blue-collar skilled	(3) white-collar	(4) ln(occscore)
Panel A: Sample 1900-1920				
Exposure	-0.0023*** (0.0007)	0.0004 (0.0004)	0.0022*** (0.0005)	0.0005 (0.0004)
Obs.	293,980	293,980	293,980	209,851
R-squared	0.101	0.132	0.067	0.107
Mean pre-Y	0.256	0.138	0.311	3.240
Panel B: Sample 1900-1930				
Exposure	0.0003 (0.0004)	-0.0004 (0.0003)	0.0003 (0.0006)	-0.0003 (0.0005)
Obs.	267,371	267,371	267,371	184,227
R-squared	0.097	0.120	0.105	0.119
Mean pre-Y	0.218	0.135	0.325	3.290
Municipality FE	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes
Ind. Controls	Yes	Yes	Yes	Yes

Notes: This table reports how antitoxin exposure during childhood affected labor market outcomes as adults based on a sample 5 to 15-year-old children who lived in Massachusetts in 1900 and can be linked to 1920 (Panel A) and 1930 (Panel B) using the crosswalks from the Census Tree Project. The dependent variable is a dummy of whether the person works in the terminal year in a low-skilled (column 1), blue-collar skilled (column 2), or white-collar occupation (column 3), and the ln occupational income score (column 4). The variable of interest, “Exposure”, denotes the average number of antitoxin bottles per 1,000 people that a child during the first nine years was exposed to. All regressions control for municipality and year of birth fixed effects and a set of individual controls (see page 37 for details). Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table A.14: The Long-run Effects of Antitoxin on Employment Status

	(1) wages	(2) bus. income over \$50	(3) works at all	(4) works in PWA projects	(5) without work	(6) unable to work
Exposure	3.3394** (1.4516)	-0.0002 (0.0007)	-0.0002 (0.0005)	-0.0009*** (0.0002)	-0.0001 (0.0004)	0.0001 (0.0001)
Obs.	135,361	231,721	247,920	247,920	164,261	247,920
R-squared	0.126	0.036	0.350	0.022	0.010	0.007
Mean pre-Y	1654.7	0.253	0.620	0.039	0.064	0.026
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Birth Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind. Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table reports how antitoxin exposure during childhood affected the employment status as adults. The sample includes 5 to 15-year-old children who lived in Massachusetts in 1900 and can be linked to 1940 using the crosswalks from the Census Tree Project. The following outcomes in 1940 are used as dependent variables: wages (column 1), a dummy of whether the individual has a business income over 50 USD (column 2), works at all (column 3), was employed on public emergency work projects (column 4), was unemployed (column 5), or unable to work (column 6). The variable of interest, “Exposure”, denotes the average number of antitoxin bottles per 1,000 people that a child during the first nine years was exposed to. All regressions control for municipality and year of birth fixed effects and a set of individual controls (see page 37 for details). Standard errors (in parentheses) account for arbitrary heteroskedasticity and are clustered at the municipality level. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

A.1 Individual death records and mortality rates by age

More than 17 million individual death records from Massachusetts between 1880 and 1914 have been digitized and published by *FamilySearch.org*, one of the world’s leading genealogy platforms. These records provide detailed geographic and demographic information, enabling the aggregation of individual data into death counts at age-year-municipality level. However, processing these raw death records involves two major technical challenges.

The first challenge arises from a significant number of records with missing age information, caused by incomplete details in the original records or transcription errors. Before 1906, the proportion of records with missing ages was below 2%, but it rose to an average of 26.8% between 1906 and 1914. To address this issue, we redistribute the counts of these records across ages 0 to 100, based on the age distribution of records with non-missing ages. This redistribution is performed within each group defined by municipality, death year, death season (April-September or October-March), decedent’s sex (male, female, or unknown), nativity (born in Massachusetts, other U.S. states, foreign countries, or unknown), and marital status (never married, ever married, or unknown).

Figure A.16 compares the original death counts (excluding records with missing ages) to the death counts incorporating the redistributed missing-age records for the 1906-1914 period. The results demonstrate a strong positive correlation, indicating that records with missing ages do not significantly distort the overall mortality distribution. Our baseline analysis relies on these imputed death counts.

The second challenge involves standardizing geographic information for both the place of death and the decedent’s residence. The original records lacked consistent place names, necessitating the cleaning of 12,229 unique transcribed death locations across 1,765,574 death records. These locations were then assigned to 319 municipalities with standardized boundaries. As a result, we identified a municipality of death for 99.59% of all records (98.95% within Massachusetts and 0.64% outside Massachusetts but registered in the state).

To account for the possibility that many individuals died away from their usual residence, often in hospitals, we also standardize residence information where available. Unfortunately, residence data was only available for deaths before 1906. For this period, we cleaned 9,344 unique non-standardized residence places and assigned them to the same standardized municipalities. Among 1,227,487 death records from 1880 to 1905, we successfully identified residence municipalities for 1,225,708 records (1,205,558 within Massachusetts and 20,150 outside the state). Furthermore, among the 1,201,174 deaths with identified death and residence municipalities within Massachusetts, 95.4% (1,145,497 cases) occurred in the same municipality. This evidence suggests that deviations between death and residence locations were not a significant issue during this period. For our baseline analysis, we rely on death locations, which are more comprehensively recorded after 1906.

Using the death counts and population data, we then calculate age-specific mortality rates, which serve as the basis for constructing life tables and estimating life expectancy. Population data by age are primarily derived from the full-count microdata of the federal census for the years 1880, 1900, and 1910 (Ruggles et al. 2017). For these census years, we aggregate individuals by single-year ages from 0 to 79 and by municipality, then log-linearly interpolate the population for non-census years between 1880 and 1910. For the years 1911 to 1914, we extrapolate population estimates.

In addition to general age-specific mortality rates, we place particular emphasis on infant mortality (age under 1) and child mortality (ages 1-4), given that young children were the most vulnerable to diphtheria infections. To measure mortality rates in these early age groups more accurately, we utilize birth counts by municipality and year from the vital statistics. Specifically, we define the infant mortality rate as follows:

$$IMR_{mt} = \frac{Deaths_{mt}^0}{Births_{mt}}, \quad (\text{A.1})$$

where $Deaths_{mt}^0$ denotes the number of deaths at age 0 in municipality m and year t , and $Births_{mt}$ represents the number of births in the same municipality and year.

Next, we calculate the child mortality rate (ages 1-4) in municipality m and year t as follows:

$$CMR_{mt} = \frac{Deaths_{mt}^{1-4}}{Pop_{mt}^{1-4}}, \quad (\text{A.2})$$

where $Deaths_{mt}^{1-4}$ is the number of deaths among children aged 1-4 in municipality m and year t , and Pop_{mt}^{1-4} represents the population of children aged 1-4 in the same municipality and year. Rather than relying on interpolated population data, we impute the annual population of children aged 1-4 based on cumulative births and deaths for the corresponding cohorts in prior years, assuming that migration among children in this age group is negligible. This approach is also employed by Alsan and Goldin (2019) and Eriksson et al. (2020). Specifically, the population of children aged 1-4 is imputed as follows:

$$Pop_{mt}^{1-4} = \sum_{a=1}^4 \left[Births_{m,t-a} - \sum_{k=1}^a Deaths_{m,t-k}^{a-k} \right]. \quad (\text{A.3})$$

Here, $Births_{m,t-a}$ denotes the number of births in municipality m during year $(t-a)$, while $Deaths_{m,t-k}^{a-k}$ represents the number of deaths at age $(a-k)$ in municipality m during year $(t-k)$. For clarity, consider the cohort of children born two years prior ($a=2$) as an example. In this case, $Births_{m,t-2}$ refers to the total number of children born in municipality m two years earlier (relative to year t), while the sum of $Deaths_{m,t-1}^{2-1}$ and $Deaths_{m,t-2}^{2-2}$ corresponds to the cumulative number of these children who died over the past two years.

Finally, We compare our imputed population of children aged 1 to 4 with the population directly reported by federal or state censuses in years when were available. Figure A.17 shows the imputed population fit well the census-reported population at the municipality level in census years.

A.2 Construction of life tables and life expectancy

This appendix explains the setup of the period life tables used to derive the various life expectancy measures. These tables estimate the number of years a hypothetical person or cohort can expect to live at a given age, based on the prevailing age-specific mortality rates in calendar year t . Our baseline life expectancy measure is derived from the single-year age

life table, which we will explain first. However, we also present life expectancy derived from the abridged life table (i.e., using broader age groups instead of single-year intervals), which will be explained afterward.

A.2.1 The single-year age life table

In the baseline measure of life expectancy, the death counts from the individual-level death certificates are combined with population data by age from the federal censuses (?). We use data for the years 1880, 1900, and 1910, and log-linearly interpolate for the intervening years (and extrapolate from 1911 to 1914). This allows us to compute m-type mortality rates for each municipality and year:

$$m_{xmt} = \frac{\theta_{xmt}}{P_{xmt}}, \quad (\text{A.4})$$

where x denotes age, m indicates municipality, and t represents the calendar year. Since we calculate mortality rates for each municipality and year, the population count for some ages can be zero, particularly in smaller municipalities and at older ages. When this occurs, it leads to missing values in Eq. (A.4), and the summation of life years will stop when computing the life table. To mitigate this issue, we stop the life-table calculations at age 80 (further details below). Additionally, if the population count at age x is zero and the number of deaths is also zero, we assume $m_{xmt} = 0$.¹

In the next step, we calculate the q-type mortality rates as:

$$q_{xmt} = \frac{m_{xmt}}{1 + (1 - a_{xmt})m_{xmt}}, \quad (\text{A.5})$$

where a_{xmt} is the average age of death at each age x . For $x > 1$, this is set to 0.5 (i.e., for those who die, say, at age 5, we assume that they lived for half a year). For the first year of life, this is set to 1/3 (i.e., those who die as infant—before turning 1—lived on average 4 months), as deaths during the first year of life are typically not equally distributed across the calendar year, but rather concentrated around the first months of life. These weights (a_{xmt}) are assumed not to vary across municipalities and time.

In principle, we could “close” the life table by assuming that everyone dies in the age interval 80-100 (i.e., $q_{80-100mt} = 1$) and calculate:

$$a_{81-100mt} = \frac{1}{m_{81-100mt}}. \quad (\text{A.6})$$

This would be a standard way of closing a life table. The problem with this approach, in our context, is that in some municipalities for some years $m_{80-100mt} = 0$ (or it could be a very small number), which would make life-years lived (per person) in the final age group go to infinity (or become a very high number). Thus, as already indicated, we stop the life table at age 79 and if $a_{79mt} = 0.5$ (as we assume), the number of life-years lived (per person) in the final age is always 0.5 and maximum life expectancy at birth is 79.5. We obtain very similar results if we close the life tables in the more standard way, following the approach outlined

¹This is one additional reason as to why we construct abridged life tables, where such occurrences are less likely.

in Eq. (A.6). In addition, when we aggregate to four areas in Massachusetts by quartiles of treatment, we continue the life-table calculation up to age 99 and close by assuming that life-years lived per person at $x = 100$ is equal to 0.5, as here we do not have the “small-area problem”.

Now, we are ready to set up the life table. The number of deaths during age x is:

$$d_{xmt} = q_{xmt}l_{xmt}, \quad (\text{A.7})$$

where l_{xmt} is the number of people alive in the beginning of age x . We set $l_{0mt} = 1$. This means that the size of the hypothetical cohort is normalized to 1 for all years and all municipalities, which is just a normalization and any other number could have been used without influencing the calculation of life expectancy. The number of life/person-years lived between age x and age $x + 1$ is:

$$\begin{aligned} L_{xmt} &= a_{xmt}l_{xmt} + (1 - a_{xmt})l_{x+1mt}, \quad x \leq 78, \\ L_{xmt} &= a_{xmt}l_{xmt}, \quad x = 79, \end{aligned} \quad (\text{A.8})$$

where we see that the maximum age is 79 and all people are assumed to die in that age.

Next, the total number of person-years lived after age x can be computed as:

$$T_{xmt} = \sum_{u=x}^{79} L_{xmt}. \quad (\text{A.9})$$

Finally, life expectancy at different age can be calculated as:

$$e_{xmt} = \frac{T_{xmt}}{l_{xmt}}. \quad (\text{A.10})$$

Life expectancy at birth and age 1 are thus calculated as:

$$\begin{aligned} e_{0mt} &= \frac{T_{0mt}}{l_{0mt}} \\ e_{1mt} &= \frac{T_{1mt}}{l_{1mt}}. \end{aligned} \quad (\text{A.11})$$

Thus, for life expectancy at age 1, we sum all the person-years lived from age 1 to the final age-group and scale with the number of people alive at age 1.

A.2.2 The abridged life table

In order to reduce the problem of missing mortality rates when using one-year age groups, we also compute life expectancies from an abridged life table, which uses wider age groups instead of single-year ages. This has the additional advantage that we can use population counts from ?, which is available for the years (1880, 1885, 1895, 1905, 1915). We only use these population counts up until the census year of 1905 as the tabulated version of the 1915-Census only contains age groups that cannot be harmonized with the previous tabulated state census years. We make log-linear interpolation in-between census years and

extrapolate from 1906 to 1914.

We consider the following age groups 0, 1-4, 5-9, ..., 75-79. We stop the life table at age 75-79. In part due the difficulties mentioned above, but also because we do have population data above age 80 in a consistent manner from the tabulated state censuses. In this case, this means that life expectancy at birth can maximum take the value 77.5, because everyone dies off (by assumption) in the last age interval, and we assume that deaths are evenly distributed within this age-group, so life years lived in this age interval is on average 2.5

The m-type mortality rates are given by:

$$\begin{aligned} {}_1m_{0mt} &= \frac{{}_1\theta_{0mt}}{{}_1P_{0mt}} \\ {}_4m_{1mt} &= \frac{{}_4\theta_{1mt}}{{}_4P_{1mt}} \\ {}_5m_{xmt} &= \frac{{}_5\theta_{xmt}}{{}_5P_{xmt}}, \end{aligned} \quad (\text{A.12})$$

where the first year of life is treated as a single-year age group (${}_1m_{0mt}$), ages 1 to 4 as a 4-year age group (${}_4m_{1mt}$), and the remaining ages as 5-year groups. This is transformed into q-type mortality rates as:

$$\begin{aligned} {}_1q_{0mt} &= \frac{{}_1m_{0mt}}{1 + (1 - a_{0mt}){}_1m_{0mt}} \\ {}_4q_{1mt} &= \frac{{}_4m_{xmt}}{1 + 4(1 - a_{1mt}){}_4m_{1mt}} \\ {}_5q_{xmt} &= \frac{{}_5m_{xmt}}{1 + 5(1 - a_{xmt}){}_5m_{xmt}} \end{aligned} \quad (\text{A.13})$$

where ${}_1a_{0mt} = 1/3$ and the remaining a 's are set to $1/2$.

The number of deaths are now indicated by ${}_1d_{0mt}$, ${}_4d_{1mt}$ and ${}_5d_{xmt}$ and the number of person-years lived between two ages are:

$$\begin{aligned} {}_1L_{0mt} &= a_{0mt}l_{0mt} + (1 - a_{0mt})l_{4mt} \\ {}_4L_{1mt} &= 4(a_{xmt}l_{1mt} + (1 - a_{xmt})l_{4mt}) \\ {}_5L_{xmt} &= 5(a_{xmt}l_{xmt} + (1 - a_{xmt})l_{x+5mt}), \quad x \leq 70 \\ {}_5L_{xmt} &= 5(a_{xmt}l_{xmt}) \cdot x = 75 \end{aligned} \quad (\text{A.14})$$

Life expectancies at birth and age 1 can thus be calculated as:

$$\begin{aligned} e_{0mt} &= \frac{{}_1L_{0mt} + {}_4L_{1mt} + \sum_{i=x}^{75} {}_5L_{imt}}{l_{0mt}} \\ e_{1-4mt} &= \frac{{}_4L_{1mt} + \sum_{i=x}^{75} {}_5L_{imt}}{l_{1mt}} \end{aligned} \quad (\text{A.15})$$

A.2.3 Limitations

There are several limitations in how we calculate life expectancy for all municipalities in Massachusetts each year. First, we rely on two different data sources to estimate age-specific mortality rates: death certificates and federal/state census data. However, the federal/state population is only measured every 10 or 5 years, which introduces considerable uncertainty in the interpolated population estimates by age, especially for smaller areas. When the population for a specific age group in a given area is estimated to be zero but deaths are recorded in that category, the age-specific mortality rate is set to missing. If this occurs at relatively early ages, it can bias life expectancy estimates downward, making them artificially short.

Another problem is related to small populations tend to have more variability in mortality rates, making estimates less stable and more susceptible to random fluctuations. In particular, in some calendar years, there might be very few deaths, while in other years many (relative to population size). The few death years will make life expectancy too high and vice versa.

As discussed in previous subsections, we address these limitations in several ways. First, we truncate the life table at age 79, assuming that individuals who reach this age live an additional half year, capping life expectancy at 79.5 years. Second, we observe similar results when using life expectancy from the abridged life table, which is less susceptible to these issues. Third, we group Massachusetts into four larger areas based on treatment intensity, where the smallest population is approximately 100,000 people. This aggregation shows similar descriptive patterns to those observed using the municipality-level data.

A.3 Main data sources:

Annual report of Birth, Marriages and Deaths, (i)

- Publication title: *“39th to the 73th Report(s) to the Legislature of Massachusetts related to the Registrar and Return of Birth, Marriages, and Deaths in the Commonwealth”*
 - Period: annually from 1880 to 1914 (publications years 1881-1915)
 - Variables (by municipality and year):
 1. Causes/diseases: bronchitis, lung tuberculosis, scarlet fever, pneumonia, whooping cough, measles, typhoid, digestive (diarrhea cholera dysentery), apoplexy, and accidents.
 2. Live births
 3. Population counts, reprinted from State and Federal censuses.
 - Reference in text: Vital Statics Report (year) or source (i)
 - Publisher: BOSTON: Rand, Avery, & Co., Printers to the Commonwealth, 117 Franklin Street.

Individual Death Records, (ii)

- Publication title: “Massachusetts Deaths, 1841-1915, 1921-1924.”
 - Period: annually from 1880 to 1914 (publications years 1841-1915 and 1921-1924)
 - Variables (by municipality, death year, and death age):
 1. aggregated death counts
 2. infant mortality rate
 3. child mortality rate (age 1-4)
 - Reference in text: individual death records (certificates)
 - Publisher: Images. FamilySearch. <http://FamilySearch.org>. Citing Secretary of State. State Archives, Boston.

Annual report of the State Board of Health of Massachusetts, (iv)

- Publication title: “*27th to 46th Annual Report(s) of the State Board of Health of Massachusetts*”
 - Period: annually from 1895 to 1914 (publications years 1896-1915)
 - Variables (by municipality and year):
 1. Supply of antitoxin bottles .
 2. Case counts for diseases: diphtheria, typhoid, measles, smallpox.
 3. Water from public works (year of introduction).
 - Reference in text: SBH (year) or source iii
 - Publisher: Boston: Wright & Potter Printing Co., State Printers, 32 Derne Street.

Full-count Federal Censuses, (i)

- Publication titles: “IPUMS Ancestry Full Count Data: Version 4.0”
 - Period: census years of 1880, 1900, 1910 (publication years are 1850, 1860, 1870, 1880, 1900, 1910 ,1920, 1930, 1940, 1950)
 - Variables: (by municipality and census year)
 1. total population in census years
 2. age-specific population and population for certain age groups (including age 0, 1-4, 5-10, 11-14)
 - Reference in text: full-count censuses
 - Publisher: Steven Ruggles, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren. Minneapolis, MN: IPUMS, 2024. <https://doi.org/10.18128/D014.V4.0>

Massachusetts State Censuses, (vi)

- Publication titles: “Massachusetts State Censuses, 1855-1915”
 - Period: 1885, 1895, 1905, and 1915 (publication years: every ten years between 1855 and 1915)
 - Variables: (by municipality and census year)
 1. total population in census years (also used to interpolate annual total population)
 2. population for certain age groups (age 0, 1-4, 5-10, and 11-14)
 - Reference in text: Massachusetts state censuses
 - Publisher: Haines, Michael R. Massachusetts State Censuses, 1855-1915. Inter-university Consortium for Political and Social Research [distributor], 2022-03-09. <https://doi.org/10.3886/ICPSR38179.v1>
 - Original Reference:
 - 1885: Massachusetts. Bureau of statistics of labor. The census of Massachusetts: 1885. Prepared under the direction of Carroll D. Wright, chief of the Bureau of statistics of labor...Boston, Wright & Potter printing company, state printers, 1887-1888. 3 v. in 4;
 - 1895: Massachusetts. Bureau of statistics of labor. Census of the Commonwealth of Massachusetts, 1895. Prepared under the direction of Horace G. Wadlin, chief of the Bureau of statistics of labor... Boston, Wright & Potter printing co., state printers, 1896-1900. 7 v.
 - 1905: Massachusetts. Bureau of statistics of labor. Census of the Commonwealth of Massachusetts, 1905. Prepared under the direction of the chief of the Bureau of statistics of labor... Boston, Wright & Potter printing co., state printers, 1908-1910. 4 v.
 - 1915: Massachusetts. Bureau of statistics. The decennial census, 1915. Taken under the direction of Charles F. Gettemy, Director of the Bureau of statistics. Boston, Wright & Potter printing co., state printers, 1918. Ix, 749 p.

Linked individuals between Censuses, (vii)

- Publication titles: “The Census Tree, 1900-1940”
 - Period: 1900 and 1940 Census
 - Variables: (at linked individual level)
 1. school absence rate
 2. years of schooling, observed in the 1940 Census
 3. labor market outcomes (occupational categories as white-collar, blue-collar unskilled, or skilled; occupational income score, and wage), observed in the 1940 Census.
 - Reference in text: Census Tree Project/Links

- Publisher: Price, Joseph, Buckles, Kasey, Haws, Adrian, and Wilbert, Haley. The Census Tree, 1900-1940. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2023-08-11. <https://doi.org/10.3886/E193262V1>

A.4 Variable definitions and remarks:

“**Population**” \equiv municipality population size

- The baseline population measure is obtained from Federal and State Census enumerations.
- These counts are available every fifth year from 1880 to 1915 by municipality and reprinted in Source (i) from where we digitized them.
- We apply log-linear interpolation in-between census years.
- Population size in 1895 from this sources is used as weight in the weighted regressions.
- We also derive population-by-age counts from the complete count U.S. census records (1880, 1900, 1910) from Helgertz et al. (2023) and apply log-linear interpolation in-between these Federal census years. This source is used in the single-year life table construction.
- We also use population-by-age-group counts from ?. These are available, for our purpose, 1880, 1885, 1895, and 1905. We apply log-linear interpolation in-between these years. This source is used in the abridged life-table construction and for the age composition variables.
- “**diphtheria all**” \equiv diphtheria mortality rate (for both sexes), defined as diphtheria (and croup) deaths per 1,000 people.
 - Nominator: death counts obtained from Source (i).
 - Denominator: baseline population variable.
 - Variation: municipality-by-year from 1880 to 1914.
 - In the main text used in: Table 2 and Figures 3, 4, and 5.
 - *Note: all cause-specific mortality rates have been derived like this and from these two sources.*
- “**treatment**” \equiv average diphtheria mortality rate from 1888 to 1894.
- “**life exp all**” \equiv life expectancy at age 1 for both sexes.
 - In the baseline, derived from the single-year life tables.
 - This life-table type uses data from Source (ii) and the Federal Census population data (Source iv).
 - Variation: municipality-by-year from 1880 to 1914.
 - In the main text used in: Table 2 and Figures 3, 4, and 5.
 - We also construct this variable by sex (“life exp female” & “life exp male”)
 - We also construct life expectancy at birth from this source. See estimates in Appendix Figure A.14 (“at birth”).

- We also construct life expectancy at age 1 using place of residence instead of place of occurrence, but only up to 1905, due to data availability. See estimates in Appendix Figure A.14 (“residence”).
- We also derived life expectancy at age 1 using the abridged table, which uses population counts by age-group data from the State Censuses (?). See estimates in Appendix Figure A.14 (“abridged”).
- See further details in Appendix A.2.
- “**antitoxin p.c.**” \equiv number of antitoxin bottles per 1,000 people
 - Nominator: bottle counts are obtained from Source (iii).
 - Denominator: baseline population variable.
 - A municipality not listed in a given report is assumed to received zero bottles that year.
 - Municipalities never listed with bottles during the sample period 1895 to 1914 are dropped from the sample.
 - Distribution years do not match perfectly with the calender year. We assign the distribution year to the calender year with the most overlaps in terms of months. See Appendix Table A.8 for the exact distribution years.
 - In the baseline, antitoxin p.c. has been winsorized, using cuts(0 95) with “winsor2” in STATA.
 - Appendix Figure A.2 reports the distribution of antitoxin p.c. with and without winsorizing, and it documents that that winsorizing improves the first-stage fit due to large values of antitoxin to hospitals.
 - Variation: municipality-by-year from 1880 to 1914. (by definition zero before 1895).
- “**infant rate**” \equiv infant mortality rate, defined as death below the age of 1 per 1,000 live births
 - Nominator: death counts by age obtained from Source ii.
 - Denominator: Live-birth counts obtained from vital statistics.
 - We also construct this variable by sex (“infant rate female” & “infant rate male”).
 - Variation: municipality-by-year from 1880 to 1914.
 - In the main text used in: Table 3.
 - See further details in Appendix A.1.
- “**child rate**” \equiv child mortality rate, defined as death between ages 1 and 4 per 1,000 children of this age-group
 - Nominator: death counts by age obtained from Source ii.

- Denominator: we use death and birth from vital statistics to calculate the size of the age group 1-4.
 - We also construct this variable by sex (“child rate female” & “child rate male”).
 - Variation: municipality-by-year from 1880 to 1914.
 - In the main text used in: Table 3.
 - See further details in Appendix A.1.
- **“infec. rate, 88-94”** \equiv average mortality rate (per 1,000 people) from 1889 to 1894 for infectious diseases.
 - Diseases included: bronchitis, lung tuberculosis, scarlet fever, pneumonia, whooping cough, measles, typhoid, and digestive.
 - Variation: municipality.
 - In the main text used in: Table 1.
- **“apoplexy rate, 88-94”** \equiv average mortality rate from apoplexy per 1,000 people from 1889 to 1894.
 - Apoplexy is a historical cause of death and refers to sudden death, which is only the measure of heart/stroke related death during this time period at the municipality level.
 - Variation: municipality.
- **“doctor pr. capita in 95”** \equiv medical doctors per 1,000 people in 1895.
 - Nominator: medical-doctor counts obtained from the Federal Censuses (Source iv) for the census years 1880, 1900, and 1910.
 - Denominator: baseline population variable.
 - Variation: municipality.
- **“dist Boston** is aerial distance to Boston.
 - Calculated using “geodist” in STATA based on longitude and latitudes for each municipality.
- **“person pr 1,000 sqm in 95”** \equiv number of people in 1895 per 1,000 square miles
 - Nominator: baseline population variable, measured in 1895.
 - Denominator: area of municipality obtained in 1915 (?).
 - Variation: municipality.
- **“person pr dwelling in 95”** \equiv number of people per dwelling houses in 1895
 - Nominator: baseline population variable, measured in 1895.

- Denominator: total number of dwelling houses in 1895 (?).
- Variation: municipality.
- **“person pr room in 95”** \equiv number of people per number of room in dwelling houses in 1895
 - Nominator: baseline population variable, measured in 1895.
 - Denominator: total number of rooms in dwelling houses in 1895 ?.
 - Variation: municipality.
- **“fb share in 95”** \equiv number of foreign born people per population in 1895
 - Nominator: baseline population variable, measured in 1895.
 - Denominator: total number of foreign born from State Census in 1895 ?.
 - Variation: municipality.
- **“diph ratio”** \equiv the number of diphtheria (and croup) death per 1,000 deaths
 - Nominator and denominator: both death counts are obtained from Source (i).
- **“case-fatality”** \equiv the number of diphtheria (and croup) death per 1,000 cases of diphtheria
 - Nominator: baseline population variable, measured in 1895.
 - Denominator: case counts are obtained from Source (iii).
 - Due to case data availability, this variable is only available from 1891 onward.
 - Municipalities with values greater than 1,000 is truncated at 1,000.
 - Variation: municipality-by-year from 1891 to 1914.
- **“apoplexy”** \equiv mortality rate from apoplexy per 1,000 people.
 - Apoplexy is a historical cause of death and refers to sudden death, which is only measure of heart related death during this time period.
 - Variation: municipality-by-year from 1880 to 1914.
- **“infec rate”** \equiv mortality rate from infectious diseases per 1,000 people
 - Diseases included: bronchitis, lung tuberculosis, scarlet fever, pneumonia, whooping cough, measles, typhoid, and digestive.
 - Variation: municipality-by-year from 1880 to 1914.
- **“accidents”** \equiv mortality rate from accidents per 1,000 people.
 - Nominator: death counts from Source (i).

- Denominator: baseline population variable.
- Variation: municipality-by-year from 1880 to 1914.
- “**birth rate**” \equiv crude birth rate, defined as the number of live birth per 1,000 people.
 - Nominator: birth counts from Source (i).
 - Denominator: baseline population variable.
 - Variation: municipality-by-year from 1880 to 1914.
- “**Doctors**”, “**Nurses**”, & “**Pharmacists**” \equiv as the number of people in the given occupation per 1,000 people.
 - Nominator: occupation counts from the Federal Census (Source iv).
 - Denominator: baseline population variable.
 - The rate variables have interpolated in-between census years.
 - This is also calculated, where the occupation counts have divided into the age-groups (20-29, 30-39, 40-49, and 50-59).
 - Variation: municipality-by-year from 1880 to 1914.
- “**Water intervention**” \equiv roll-out variable, based on the establishment date of public water works.
 - Records the establishment date of public water works.
 - Source: SBH (1930).
 - Variation: municipality-by-year from 1880 to 1914.
- “**Infectious hospital**” & “**General hospital**” \equiv roll-out variable, based on the establishment date of an infectious disease or a general hospitals.
 - Records the establishment date of the hospital by type (infectious or general hospital).
 - Source: Benevolent Institutions, 1910. Department of Commerce, Bureau of the Census.
 - Variation: municipality-by-year from 1880 to 1914.
- “**Age composition**” \equiv population shares for the age groups 0, 1-4, 5-9, and 10-14.
 - Population share are obtained using data from ?.
 - Variation: municipality-by-year from 1880 to 1914.

A.5 Historically Consistent Municipality Boundaries

During our primary sample period (1880-1914), the county boundaries in Massachusetts underwent several changes due to the separation or annexation of municipalities (towns). To construct a balanced and comparable panel, we adjusted the year-by-year town boundaries as follows:

(1) If municipality *A* was separated from municipality *B* and incorporated between 1876 and 1914, we combined the two municipalities and aggregated them into municipality *B*. We extend the adjustment period back to 1876 to account for lagged births and deaths between 1876 and 1880, which are used to estimate child mortality rates for ages 1-4 in 1880.

(2) If municipality *A* was incorporated before 1876 but annexed into municipality *B* between 1876 and 1914, we also combined the two municipalities and aggregated them into municipality *B*.

Below is a complete, chronologically ordered list of municipality boundary changes that occurred during the sample period:

- 1876: Merrimac was separated from Amesbury and incorporated as an independent town.
- 1878: North Adams was separated from Adams and incorporated as an independent town; Hampden was separated from Wilbraham and incorporated as an independent town.
- 1880: Cottage City was separated from Edgartown and incorporated as an independent town. It was later renamed Oak Bluffs in 1907.
- 1881: Wellesley was separated from Needham and established by the Massachusetts legislature.
- 1884: Bourne was separated from Sandwich and incorporated as an independent town.
- 1885: Millis was separated from Dedham and incorporated as an independent town.
- 1886: Hopedale was separated from Milford and incorporated as an independent town.
- 1888: Avon was separated from Stoughton and incorporated as an independent town.
- 1897: Bradford was annexed to the city of Haverhill; Westwood (then referred to as "West Dedham") was separated from Dedham and incorporated as an independent town.
- 1905: Plainville was separated from Wrentham and incorporated as an independent town.
- 1912: Hyde Park was annexed into the city of Boston.

A.6 Relation to shift-share like instrument

In this appendix, we show that our baseline instrumental variable, reported in Equation (1), is closely related to an alternative instrument, where the aggregate number of bottles is distributed according to municipality specific diphtheria shares (a “shift-share” like instrument). For convenience, we repeat the structure of our baseline instrument here:

$$IV_{mt}^{base} = treatment_m \times I_t \times (t - 1894), \quad (\text{A.16})$$

where treatment intensity is defined as the diphtheria mortality rate averaged across the pre-antitoxin years 1889 to 94:

$$treatment_m = d_m^{pre}. \quad (\text{A.17})$$

A shift-share type of instrumental variable can be defined defined as:

$$IV_{mt}^{alt} = B_t \frac{D_m^{pre}}{D_{MA}^{pre}} \frac{1}{P_m^{94}}, \quad (\text{A.18})$$

where B_t is the total number of antitoxin bottles supplied to the municipalities in our sample, D_i^{pre} is the total number of diphtheria deaths from 1889 to 1894 in municipality m , D^{MA} is the total number of diphtheria deaths in our sample of municipalities during the same pre-antitoxin years (i.e., D_m^{pre}/D_{MA}^{pre} is the share of diphtheria deaths in municipality m), and P_m^{94} is the municipality population size in 1894. Accordingly, IV_{mt}^{base} is the predicted number of antitoxin bottles per capita, where the aggregate number of bottles supplied by the SBH each year is distributed according to the pre-antitoxin mortality share and the scaling population size is fixed to the pre-antitoxin year of 1894. Alternatively, we could have let the population vary by year, but this assumption would be less conservative (as population size itself is influenced by the use of the antitoxin) and the connection to our baseline instrument would be less obvious. Let us provide a simple example of how the prediction works. If Boston had, say, 20% of all diphtheria deaths prior to the antitoxin treatment, the municipality is allocated 20% of all bottles in each year and then the predicted number of total bottles available to Boston is scaled by its pre-antitoxin population size.

In the following, we show how the two instruments relate to each other. Assume that the baseline treatment takes this slight alternative form:

$$treatment_m = \frac{D_m^{pre}}{P_m^{94}}, \quad (\text{A.19})$$

where instead of taking the average mortality rates over multiple years, we sum all pre-antitoxin diphtheria deaths and scale with the population size of 1894. The interpretation of this ratio remains relatively close to a (mortality) rate, and using the formulation in Equation (A.19) as treatment intensity for the baseline instrument provides very similar

results (available upon request). Next, we substitute this into Equation (A.16) and rearrange:

$$\begin{aligned}
 IV_{mt}^{base} &= \frac{D_m^{pre}}{P_m^{94}} \times I_t \times (t - 1894) \Leftrightarrow \\
 D_m^{pre} &= \frac{P_m^{94} IV_{mt}^{base}}{I_t \times (t - 1894)},
 \end{aligned}
 \tag{A.20}$$

which we combine with Equation (A.19) to give:

$$\begin{aligned}
 IV_{mt}^{alt} &= B_t \frac{\frac{P_m^{94} IV_{mt}^{base}}{I_t \times (t - 1894)}}{D^{MA}} \frac{1}{P_m^{94}} \Leftrightarrow \\
 IV_{mt}^{alt} &= \frac{B_t}{D^{MA} \times \tau} IV_{mt}^{base},
 \end{aligned}
 \tag{A.21}$$

where we, in the last line, have omitted the indicator (I_t) for simplicity, since this only reflects the fact that B_t is per definition zero before 1895, and τ is accordingly defined as the linear trend $\tau \equiv (t - 1894)$ for $t > 1894$. From this last expression, we observe that the difference between the instruments is the scaling factor ($B_t / (D^{MA} \tau)$), which is possibly time-varying, but unrelated to municipality specific conditions. Therefore, whether we use one or the other instrument should not be important in terms of obtaining consistent estimates. The 2SLS estimates for diphtheria and life expectancy using the alternative shift-share type of instrument are reported in Appendix Figure A.14 as the specification “shift-share IV”.

A.7 Determinants of the antitoxin diffusion

In this appendix section, we provide a more detailed description of on how the adoption of antitoxin is related to different pre-antitoxin municipality characteristics. This exercise is based on different versions of the estimating equation (2). The estimates are reported in Table A.2. All regressions include municipality and county-by-year fixed effects and are weighted by population size in 1895. The baseline estimate, reported in column 1, implies that 10 years into the use of antitoxin (in 1904), the adoption rate is 1.49 bottles per 1,000 people in a municipality with average treatment intensity ($0.35 * 10 * 0.43$) and the range goes from 0 to 7.01 bottles when moving from zero to maximum treatment. This is our baseline first-stage specification, which is used in Table 2.

In the remaining columns, we add the different pre-antitoxin characteristics, also interacted with the indicator and year trend. In column 6, where all controls are included at the same time, we see that the most robust determinants of the adoption of antitoxin is the instrument, along with the number of doctors per capita, distance to Boston, population density, and the opening of new infectious disease hospitals, in particular.

Using a decomposition exercise following Gelbach (2016), we find that population density, dwelling size, and doctors per capita (all interacted with a linear trend) are the key factors driving the decrease in the first-stage coefficient when robustness controls are added (Appendix Table A.3)

References

- Alsan, Marcella and Claudia Goldin**, “Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920,” *Journal of Political Economy*, 2019, 127 (2), 586–638.
- Anderson, Theodore W and Herman Rubin**, “Estimation of the parameters of a single equation in a complete system of stochastic equations,” *The Annals of mathematical statistics*, 1949, 20 (1), 46–63.
- Cattaneo, Matias D, Richard K Crump, Max H Farrell, and Yingjie Feng**, “On binscatter,” *American Economic Review*, 2024, 114 (5), 1488–1514.
- Eriksson, Katherine, Peter Z. Lin, and Gregory Niemesh**, “Immigration and Infant Mortality in Massachusetts: Evidence from the Age of Mass Migration,” 2020. Working Paper.
- Gelbach, Jonah B.**, “When Do Covariates Matter? And Which Ones, and How Much?,” *Journal of Labor Economics*, 2016, 34 (2), 509–543.
- Helgertz, Jonas, Steven Ruggles, John Robert Warren, Catherine A. Fitch, J. David Hacker, Matt A. Nelson, Joseph P. Price, Evan Roberts, and Matthew Sobek**, “IPUMS Multigenerational Longitudinal Panel: Version 1.1,” 2023.
- Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter**, “Valid t-ratio Inference for IV,” *American Economic Review*, 2022, 112 (10), 3260–3290.
- Rambachan, Ashesh and Jonathan Roth**, “A more credible approach to parallel trends,” *Review of Economic Studies*, 2023, 90 (5), 2555–2591.
- Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek**, “Integrated Public Use Microdata Series: Version 7.0 [dataset],” 2017.