Exploring the Impact of Medical Technology on Life Expectancy: Evidence from the Free Supply of Diphtheria Antitoxin in Massachusetts *

Philipp Ager Casper Worm Hansen Peter Zhixian Lin†

[Latest version: October 2022 - preliminary draft]

Abstract

In this paper, we explore the impact of the first effective medical treatment for an infectious disease—diphtheria antitoxin—on the historical health transition. In 1895, the State Board of Health in Massachusetts began providing free supplies of the antitoxin for medical use throughout the state. This policy has later been recognized as a significant event in the public health history of Massachusetts. We use cross-municipality variation in pre-antitoxin diphtheria mortality rates and the availability of free antitoxin in 1895 to create an instrumental variable for local adoption rates, as measured by the number of antitoxin bottles per capita. By analyzing approximately 1.6 million death certificates from 1880 to 1914, we find that a hypothetical 10-year delay in the development of antitoxin would have reduced life expectancy at birth by one year, primarily due to reductions in child mortality. Our results suggest that medicine played a significant role in the increase of life expectancy in the state of Massachusetts in the early 20th century.

Keywords: Life expectancy; medical technology; antitoxin.
JEL Codes: J11, N32, I15

*Acknowledgments:

†Ager: University of Mannheim and CEPR, philipp.ager@gmail.com; Lin: UCLA, zxlin@ucla.edu; Hansen: University of Copenhagen and CEPR, casper.worm.hansen@econ.ku.dk

1 Introduction

Life expectancy at birth has been steadily increasing since the second half of the 19th century. The increase has been about 3 months per year at the "frontier" (Oeppen and Vaupel 2002). Before the advent of modern medicine (sulfa drugs and penicillin), the increase in life expectancy in the United States was even greater, at about 4.8 months per year from 1880 to 1930 (see, e.g., Hacker 2010; Goldin and Lleras-Muney 2019). Various explanations have been offered in the literature for the significant improvements in population health that have been observed, and the pre-1930s ones include: sanitation (e.g., Cutler and Miller 2005; Ferrie and Troesken 2008; Alsan and Goldin 2019), nutrition (e.g., Fogel 1994, 2004; McKeown 1976), living conditions (e.g., Ager et al. 2020), public-health programs (e.g., Moehling and Thomasson 2014; Egedesø et al. 2020), and environmental factors (e.g., Barreca et al. 2016; Beach and Hanlon 2018).[1]

While the breakthrough of modern medicine is often associated with innovations from the 1930s onward, such as sulfa drugs (e.g, Jayachandran et al. 2010) and penicillin (e.g., Acemoglu and Johnson 2007; Alsan et al. 2021), the first ever (clinical) effective medical treatment against an infectious diseases—diphtheria antitoxin—was invented some 40 years earlier. However, perhaps due to historical data availability on the diffusion of antitoxin and local health measures, there is to our knowledge little empirical evidence on its contribution to the rise in life expectancy.[2] This paper attempts to fill this gap by studying how antitoxin influenced population health in the state of Massachusetts, combining newly-collected municipality data on the diffusion of antitoxin and circa 1.6 million death certificates from 1880 to 1914. Our main finding is that a counterfactual 10-year delay in invention of antitoxin would reduced life expectancy at birth by about one year (or 1.2 month less increase per year).

Using Massachusetts as a laboratory for studying the potential population-health effects of antitoxin has several advantages. Firstly, the state's policy of providing free antitoxin to all residents reduce the possibility of an income gradient in the adoption. The free-supply policy has subsequently been regarded as a milestone in the public-health history of the state. Second is data availability: The State Board of Health (henceforth the SBH) kept track on the number of antitoxin bottles distributed to each municipality from 1895 to 1914, which we digitized and use to measure the local diffusion of the new medical technology.[3] We combine these data with individual death certificates data, which in our settings can be used to calculate life tables for more than 200 municipalities annually from 1880 to 1914. Thirdly, mortality and cause of death data from Massachusetts are well-documented and considered reliable.

The main challenge when estimating the causal effect of antitoxin on population health is reverse

[1]The literature is, however, far from settled and there is an ongoing lively discussion on the relative importance of the proposed different explanations (e.g., Anderson et al. 2019b,a; Clay et al. 2020)

[2]Antitoxin is, for example, not mentioned in literature-review papers such as Cutler et al. (2006) and Costa (2015).

[3]To our knowledge, most previous larger-area studies on the effects of 1930s- and 1940s-medicine are not able to measure the local diffusion directly and rely on indirect measures (e.g., Acemoglu and Johnson 2007; Jayachandran et al. 2010; Alsan et al. 2021; Bhalotra et al. 2022).

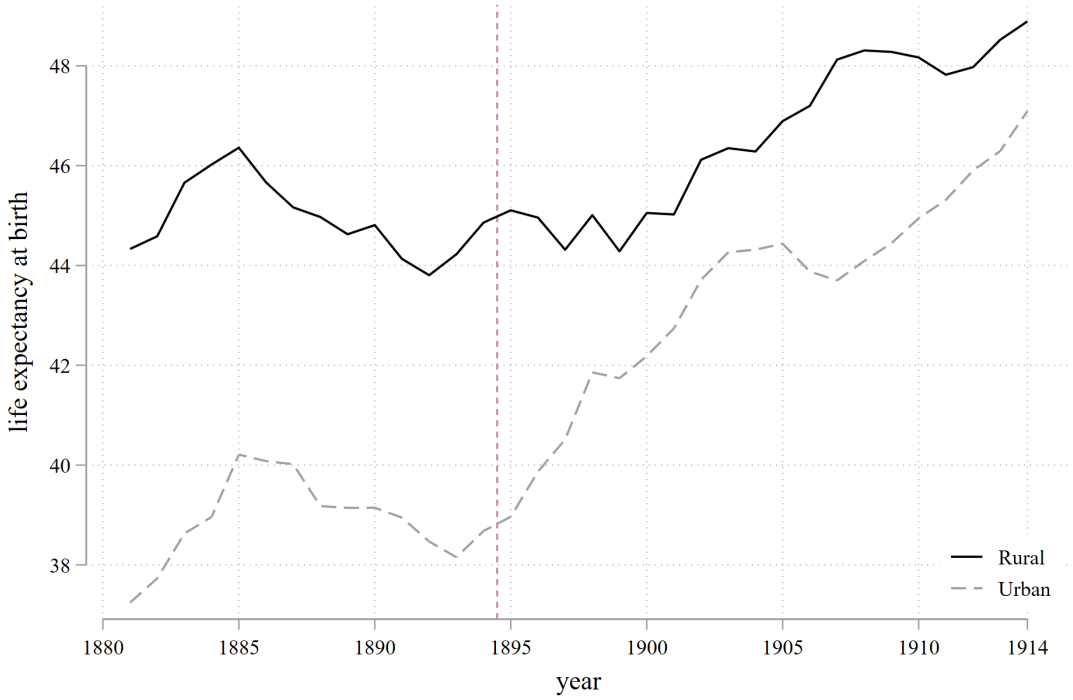
causality since the demand for a given medication will be higher during epidemics. Our empirical strategy exploits that antitoxin became suddenly and freely available for adoption starting in 1895 together with the fact that some municipalities stood to benefit more from the technology, in terms of potential mortality reductions, as they were systematically more plagued by diphtheria prior to antitoxin to construct an instrumental variable for the observed antitoxin adoption rates. In the reduced form, this strategy corresponds to a variant of the double-difference with a continuous measure of treatment intensity, which has been applied previously (e.g., Acemoglu and Johnson 2007; Bleakley 2007). Our 2SLS estimates indicate that the adoption of antitoxin increased life expectancy at birth in particular due to reductions in child mortality (ages 1-4), which is consistent with the pre-antitoxin age-profile of diphtheria mortality. We also find evidence of robust second-order effects on strokes mortality, which in our view indicate that a "general equilibrium" analysis like ours is important when assessing the effects of new medical technologies.

Before the availability of antitoxin, diphtheria was more common in urban and densely populated areas, as it was transmitted from person to person through infectious droplets. Therefore, urban municipalities had a greater potential to benefit from the new technology compared to rural areas. We argue that the decline in the urban health penalty during this period could be partly due to the availability of this new treatment. In Figure 1, we use death certificates data and census population counts to calculate life expectancy for urban and rural residents of Massachusetts. We observe that the urban-rural gap narrows from around 6 years to almost 1 year, and the timing of this narrowing is consistent with the initial diffusion of antitoxin.

In the recent historical demographic literature, there seems to be a perception of limited diffusion of antitoxin in the early 20th century U.S. In Cain et al. (2018, "Medical Technology, Doctors, and Hospitals" by M. Thomasson), it is argued that antitoxin did not readily diffuse, citing a 1907 State Board of Health report from Indiana. Similarly, Preston and Haines (1991) argues that the high diphtheria mortality rates by the turn of the 20th century indicate limited diffusion. A contemporary statistical study by Crum (1917) compared diphtheria mortality rates before and after antitoxin and concluded that it played a positive role in improving health. However, as also noted in Condran (2008), simple before-and-after comparisons could also be capturing other secular changes in the mortality environment. Our newly collected data show that the free-supply regime in Massachusetts led to a relatively fast diffusion of the technology from 1,724 bottles in 1895 to 118,561 bottles in 1914, and our empirical framework allows us to hold constant other time-varying factors that might have influenced health during this time period.[4]

[4]See also our annual bottles per capita numbers from 1895 to 1914 visualized in Figure 5 to get an impression of the temporal diffusion pattern.

Figure 1: Life expectancy at birth in urban and rural Massachusetts



Notes: *The figure shows the development of the average life expectancy at birth calculated in urban and rural Massachusetts, where urban is defined as living in municplaity with more than 10,000 inhabitants in 1890. The vertical line indicates the first year when antitoxin became freely available for adoption.*

2 Background

2.1 Diphtheria

Diphtheria is a contagious bacterial infection that mainly affects the upper respiratory tract, but can also spread to other areas of the body. The bacteria causing diphtheria—Corynebacterium diphtheriae, which was first identified by Edwin Klebs and Friedrich Löffler in the 1883 and later known as the Klebs-Löffler bacillus—produce a toxin that can cause severe damage to the body's tissues and organs. Transmission occurs from person to person via respiratory droplets from coughing or sneezing, as well as via contaminated food products, and symptoms includes general weakness and a swollen neck

The most common complications of diphtheria is the formation of a thick, gray membrane in the throat and nasal passages, which can make it difficult to breathe. This membrane, known as a pseudomembrane, is made up of dead tissue, bacteria, and the diphtheria toxin. If left untreated, the pseudomembrane can obstruct the airways and cause suffocation. Other complications of diphtheria include secondary pneumonia, myocarditis, which is an inflammation of the heart muscle and neuritis. These complications can be life-threatening (strokes and heart attacks) and lead to

long-term health problems, such as paralysis and dysphagia. These complication and sequel diseases can be avoided if the initial infection is treated immediately.[5]

While the vast majority of infections today occur in the developing world, during the second half of the 19th century diphtheria was on the rise, along with industrialization and urbanization, in the US and many other now-developed countries (e.g., Preston and Haines 1991). For example, the cumulated number of diphtheria deaths was 40,877 in the 10 largest US cities (including Boston) during the pre-antitoxin years 1889-93, and the annual mortality rate was close 1,20 deaths per 1,000 people. Boston was close to this average with a mortality rate of 1,18 deaths per 1,000 people, while New York had higher mortality rate of 1.50 deaths per 1,000 (Crum 1917).

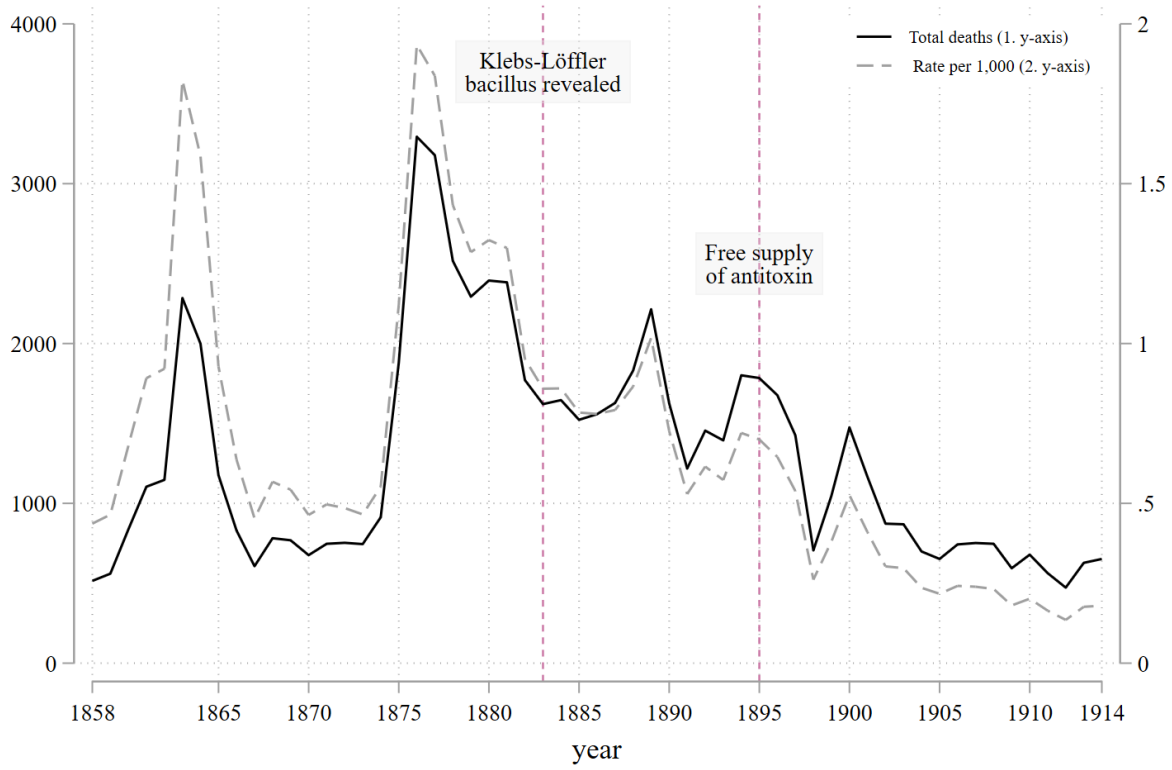
Figure 2 shows the total number of deaths caused by diphtheria (and croup) in Massachusetts from 1858 to 1914.[6] In this period, we see significant annual variations in the number of deaths, with the highest number of deaths (3,294) occurring in the mid-1870s. After that, the number of deaths steadily declined to reach 652 in 1914. This trend in the number of deaths was not caused by changes in the population at risk, as the mortality rate for diphtheria followed a similar trend. This means that simply comparing the mortality rate for diphtheria before and after the introduction of antitoxin in 1895, as has been done in previous studies (e.g., Crum 1917), may not accurately reflect the role of antitoxin in the decline of diphtheria. A similar critique can be found in the study by Condran (2008) on diphtheria mortality in Philadelphia.

The Massachusetts vital records (various years) indicate that diphtheria accounted for up to 10% of all deaths during the peak years in the 1870s. This number declined to less than 1% by the 1910s. The records also reveal that children were disproportionately affected by the disease, with 85% of diphtheria deaths occurring in children under the age of 10 in 1890. The disease was also known as "The strangling angel of children". This skewed age profile suggests that a reduction in diphtheria deaths would have a significant impact on life expectancy at birth, even though the disease was not as deadly as others such as tuberculosis. Additionally, the records show that people died from diphtheria throughout the calendar year, but death rates were generally higher in the autumn and winter months, making the disease similar to others such as pneumonia and influenza in terms of seasonality.

[5]See also the descriptions by the *Centers for Disease Control and Prevention* (CDC)

[6]Diphtheria was classified separately as a cause of death in Massachusetts' vital registration reports starting in 1858.

Figure 2: Total number of diphtheria and croup deaths



Notes: *The figure shows the development of the total number of deaths due to diphtheria and croup (black solid line) and the rate per 100,000 people (gray dashed line) from 1858 to 1914 for the state of Massachusetts. The data have been obtained from the vital statistics of Massachusetts (various years). 1858 marks the first year where diphtheria was recorded as a separate cause of death in the vital statistics.*

2.2 Antitoxin

Less than a decade after identifying the Klebs-Löffler bacillus as the cause of diphtheria, antitoxin was developed by Emil Adolf von Behring and Shibasaburo Kitasato (in 1890 Germany), and Behring went on to win the first ever Nobel Prize in Medicine in 1901 for this achievement. Antitoxin for diphtheria was produced by injecting toxin into a horse. The horse is given many small doses of toxin until a high concentration of the antitoxin build up in the blood of the hose, which then results in the so-called antiserum, which could be used to medically treat patients with the disease.

The use of antitoxin in practice began a few years later in places such Germany, New York City, and Massachusetts. In particular, after having seen the the effectiveness of demonstrated at the International Congress of Hygiene and Demography in Budapest in 1894, the Massachusetts State Board of Health (SBH) began preparing for production and started distributing antitoxin throughout the Commonwealth in March 1895 from its production facilities in Boston. The SBH 1895 annual report explains the motivation behind the free supply of antitoxin as follows (p. 690):

"The State Board of Health, in accordance with its organic purpose as defined in the statutes, having in view the "interests of health and life among the citizens of the Commonwealth," and recognizing the value to the people of any agent which will measurably prevent the ravages of a disease capable of destroying more than a thousand of lives annually in Massachusetts and of causing suffering which cannot be expressed numerically, has prepared a supply of antitoxin, for the benefit, primarily, of such communities in this State as find it difficult or impracticable for any reason to supply themselves with the new agent from reliable sources." .... "The Board does not propose to offer it for sale, but its gratuitous distribution will be under strict conditions designed to prevent abuse and waste and to obtain the most beneficial fruits."

The SBH kept records of the number of antitoxin bottles distributed to municipalities in Massachusetts and published these numbers in their annual reports from 1895 to 1914. These reports have been digitized and will be used in the analysis below

During the first year of antitoxin supply in 1895, each bottle contained 15-20 cubic meters of serum. The strength of the serum gradually increased over time, reaching 100 units per cubic centimeter in 1896 and 200-400 units per cubic centimeter in 1899. This resulted in a decrease in the amount of serum per bottle, with each bottle containing 1,500-2,000 units. The SBH's 1900 annual report provides information on the quantity of treatment given to patients. This varied from less than 1,000 units to more than 20,000 units (equivalent to 0.5-13 bottles per treatment), but 54% of patients received less than 5,000 units. The total number of bottles supplied by the SBH increased from 1,724 in 1895 to 118,561 in 1914. This can also be seen in Figure 5, which shows the supply of antitoxin per 1,000 people for our baseline sample of municipalities.

In 1914, William H. Park developed a diphtheria vaccine by mixing diphtheria antitoxin with diphtheria toxin. This made immunization against diphtheria possible. The SBH also stopped publishing data on the distribution of antitoxin in 1915. This data restriction, along with the development of the vaccine, provides a reason to end our analysis in 1914, as we will only be estimating the effect of antitoxin during a time when medical immunization was not yet available. It is worth noting, however, that contemporary reports indicate that small doses of antitoxin (less than 1,000 units) were given to uninfected individuals to provide short-term immunity against the disease.

2.3 The development in other contemporary areas

Contemporary publications, such as Crum (1917) or George (1923), report various diphtheria mortality statistics for selected areas throughout the world, both before and after the breakthrough of antitoxin in the 1890s, as also mentioned above. According to Crum (1917), the pre-antitoxin (typically 1889-93) diphtheria mortality rates varied from 0.18 (Ireland) to 4.2 (Serbia) deaths per 1,000 people, while post antitoxin (typically 1910-14) rate varied from 0.4 (Serbia) to 0.06 (Chile) deaths

per 1,000 people. The rates of decline were largest in countries such Serbia, US, and Denmark.[7]

Crum (1917) also provides data on comparative mortality rates for 12 larger US cities before and after antitoxin that indicate relative large reductions for the city of Boston (among the top-4 cities, which in addition includes New York, Detroit, and Chicago). He uses these before-and-after comparisons to argue that antitoxin saved 622 lives annually in Boston alone.

The distribution of antitoxin began in New York city around the same time as in the state of Massachusetts and they had a similar policy of providing it free of charge, but only among poor people (see, e.g., the "*Annual report of the Department of Health of the City of New York for the calendar year 1905*"). The state of Pennsylvania also had a policy of free antitoxin for the poor, but its local production of antitoxin was inadequate to meet the general demand (Liebenau 1987). This is also evident if we compare the production of antitoxin to production in Massachusetts. In 1905, for example, the State Board of Health in Pennsylvania produced what corresponds to 15,000 bottles (assuming one bottle contained 1,500 units), while the SBH of Massachusetts produced 47,000 bottles for a population only about half the size. In addition, one can find scattered anecdotal evidence suggesting unequal access to the new medicine at the turn of the 20the century.[8]

3 Data

Our dataset is constructed out of four main data sources: I) "*The Annual Report on Birth, Marriages, and Deaths in Massachusetts*" every year from 1880 to 1914, II) complete-count Federal Censuses (1880, 1900 and 1910) from IPUMS (Ruggles et al. 2021) and Massachusetts State Censuses, 1880-1915 from Haines (2022); III) death registers and certificates from 1880 to 1914, and IV) "*The Annual Report of the State Board of Health of Massachusetts* every year from 1895 to 1914. This section discusses these sources and the variables derived from them.

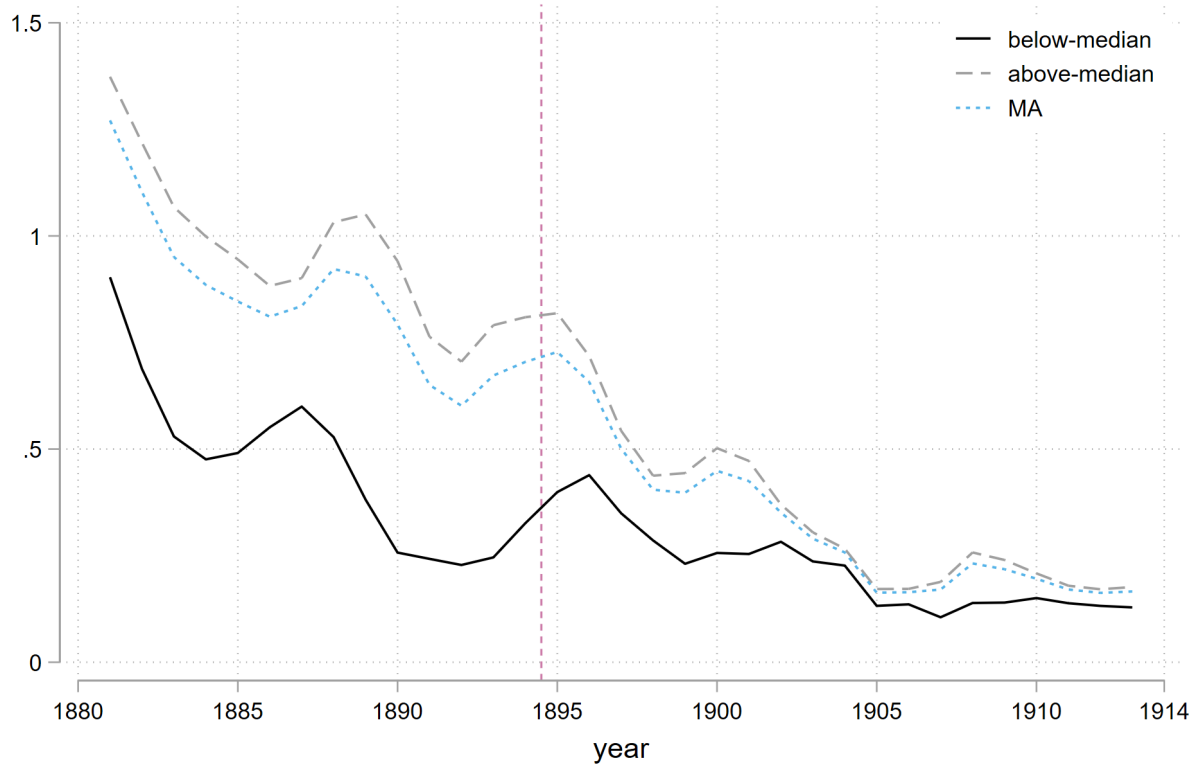
The history of Massachusetts' vital records is well documented and the death registration system is generally considered reliable and of high-quality. By 1900, only around 1% of all deaths were unregistered. The decline in unknown causes of death towards the end of the 19th century further reveals that the returns of the causes of deaths increased substantially in accuracy (Gutman 1956, 1959; Condran and Crimmins 1980). From Massachusetts' vital statistics, we digitized the annual number of total deaths (and by cause-of-death, which we also refer to as "disease"), as well as the number of live birth, for each municipality (i.e., cities and towns) and year from 1880 to 1914. The disease data are from the table titled "Special Causes", which are available at the municipality level. Our main disease variable is the number of deaths from diphtheria. From 1880 to 1901, these deaths are reported in the category "diphtheria and croup", while from 1902 to 1912 they are reported separately as "diphtheria" and "diphtheritic croup" and for 1913 and 1914 only as one category "diphtheria" but this also contains deaths from croup. Therefore, our diphtheria variable includes deaths from diphtheria and (diphtheritic) croup in all years. In addition, we use the

[7]The US statistics are based on the 10 largest cities, including Boston.

[8]Read, for example, the story about how science "conquered" diphtheria here or the description of diphtheria in Canada here.

following diseases (and accidents): bronchitis, digestive diseases (diarrhea, cholera, and typhoid), tuberculosis of the lungs (TB), pneumonia, scarlet fever, whooping cough, strokes (apoplexy and cerebral hemorrhage), and accidents. We use municipality population data from state censuses (Haines 2022) (1880, 1885, 1895, 1905, 1915) and interpolate linearly in-between census year to construct annual mortality rates.[9] Figure 5 shows the development of the diphtheria mortality rate as a three-year moving average for municipalities grouped by "treatment intensity" (defined below) and for all municipalities. The averages are weighted by the municipality's 1895 population size. While the figure shows a steady decline in diphtheria deaths from 0.9-1.4 per 1,000 people in 1880 to around 0.25 death per 1,000 people in 1914, the rates converges across groups mainly during the post-antitoxin era. This pattern suggests that municipalities with higher pre-antitoxin diphtheria mortality rates benefited more from the free supply of antitoxin starting in 1895.

Figure 3: Diphtheria mortality rates



Notes: *The figure shows the development of the average diphtheria mortality rate per 1,000 people by treatment intensity (below/above median) and all municipalities in our Massachusetts (MA) sample. We report three-year moving averages and use 1895 municipality population size as weight. The vertical line indicates the first year when antitoxin became freely available for adoption.*

We use the complete-count federal census records from IPUMS for the decades 1880 and 1900-

[9]From Haines (2022) we also obtain a number of other municipality characteristics such the as the number of dwellings, rooms in dwellings houses, and area size.

1910 (Ruggles et al. 2021) to measure population size by one-year age groups, which we use to construct the life tables. The aggregation of the census data at the municipality level is based on georeferenced crosswalks of individuals from Berkes et al. (2021), which contain the longitude and latitude for every Census designated place. For every individual listed in the census, the crosswalk lists the historical individual-level identifier (HISTID) provided by IPUMS together with the georeferenced location of the individual. The crosswalks are merged with the complete-count census records by HISTID, which allow us to construct different municipality-level characteristics, such as the number of medical doctors, the number of foreign born, etc. The population sizes by age are interpolate linearly in between census years.

We use individual death certificates to calculate annual age-specific mortality rates at the municipality level. The death certificates are digitized and provided by *FamilySearch.org*, and are part of the collection Massachusetts Deaths, 1841-1914. We consider the period 1880 to 1914, which in total includes 1,633,553 deaths. We derive the annual infant mortality rates at municipality level from the infant deaths divided by the birth counts. We also derive the mortality rates for children age 1-4 from the deaths at corresponding ages divided by age 1-4 population, which is imputed based on births and cummulative deaths for each age cohort.[10] Lastly, we calculate the single-year age-specific mortality rates by aggregating the deaths certificates to the annual municplaity level for single-year ages (up to 100) and combining the deaths at particular ages and the interpolated census population age counts from IPUMS. We provide further details in the appendix on how we use and tabulate the death certificates.[11]

Using the age-specific mortality rates, we can construct annual municipality life tables. In the construction of the life tables, we assume deaths to be equally distributed across the calendar years for all ages, except for the first year of life, where it is assumed that an infant death corresponds to 1/3 of a life year lived. We close the table at age 100 by calculating the life years lived as one divided by the mortality rate at age 100. From the life tables, we can compute life expectancy at all ages, but we focus on life expectancy at birth as our main measure of population health. In addition, we use the age-specific mortality rate, which are used as input into the life tables, as outcomes in the regression analysis.[12]

Figure 4 displays life expectancy at birth, following the same structure as in the previous figure. First, the population weighted average life expectancy increased from 40 years in the beginning of the 1880s to 47 years in 1910s (blue dashed line). Second, as with the diphtheria mortality rate, we

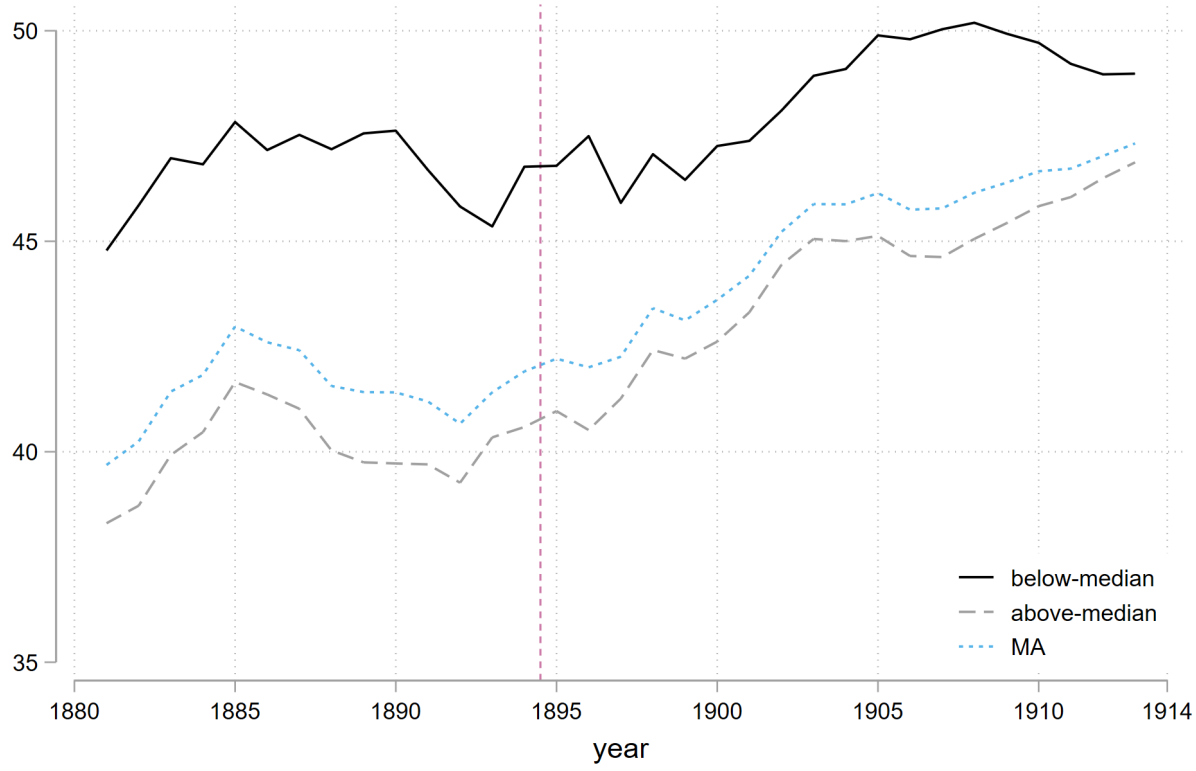
[10]We follow the same formula used by Alsan and Goldin (2019). For each age from 1 to 4, we impute the population stock by subtracting the cumulative deaths for an age cohort from the births for the age cohort, assuming limit emigration of young children. See more details in the appendix.

[11]In the raw digitized death certificates, we have a significant number of death records with missing age after the year 1905. We assign these age-missing deaths into ages in proportion to the age distribution for all death records with non-missing age in the same year, sex, and municipality. In addition, we only assign age-missing deaths with documented spouse to ages above 15, the minimum marriageable age we assume.

[12]Since we calculate annual life tables for all municipalities, including municipalities with smaller populations, it is evident that life expectancy will be measured with error, in particular for smaller areas, which is one reason as to why all our estimation results are derived using the 1890-municipality population size as weight.

only see convergence across the two municipality groups after the introduction of antitoxin in 1895: the gap in life expectancy remained constant at about 6 years from 1880 to 1894 and reduced to circa 2 years by 1914. Lastly, while our empirical analysis never uses the level of life expectancy (but only changes), we nonetheless obtain a reasonable estimates. For example, we find that the population-weighted average life expectancy in Massachusetts in 1890 to be 41.4 years and 44.2 in 1901, while the official contemporary statistics (Glover 1921) reports 42.5 years and 46 years, respectively. Possible explanations for these smaller level differences can be that our numbers are only based on 225 (out of more than 300 possible) municipalities in Massachusetts and individual death records up until the age 100.

Figure 4: Trends in life expectancy at birth



Notes: *The figure shows the development of the average life expectancy at birth by treatment intensity (below/above median) and all municipalities in our Massachusetts (MA) sample. We report three-year moving averages and use 1895 municipality population size as weight. The vertical line indicates the first year when antitoxin became freely available for adoption.*

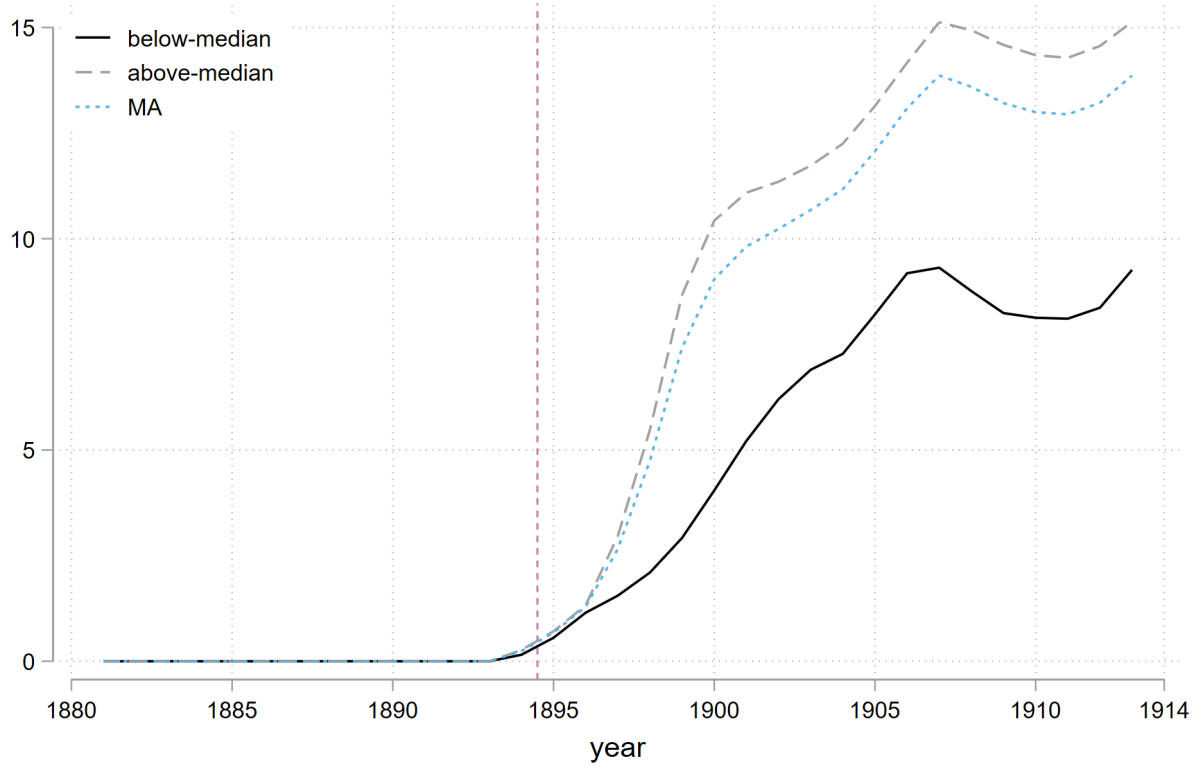
Finally, we collected data on the supply of antitoxin to each municipality from the State Board of Health's (SBH's) annual reports from 1895 to 1914. These reports contain information on the number of bottles supplied to each municipality. In principle, if municipalities were not listed, they did not receive any antitoxin directly from the SBH, however, the possibility of redistribution of antitoxin from listed to non-listed municipalities cannot be excluded according to the SBH.

In the baseline, we assume that non-listed municipalities did indeed not receive any bottles but replace extreme large per capita values by the 95 percentile value ("winsorize"). Our conclusions are generally robust to not making this adjustment, however, it increases the precision of our estimates.

Figure 5 displays the supply of antitoxin per 1,000 people, following the same structure as the previous two figures. The population weighted average supply was close 14 bottles per 1,000 people at the end of our sample, and we observe that municipalities more affected by diphtheria prior to antitoxin (the above-median sample) adopted more antitoxin in per capita terms throughout the antitoxin period.

Therefore, taking the patterns in Figures 3-5 together, we conclude that municipalities with the potential to benefit more from the new technology also adopted more of it and experienced larger declines in diphtheria mortality rates and larger increases in life expectancy at birth. We are going to exploit these features more systemically in our identification strategy below.[13]

Figure 5: Antitoxin adoption rates



Notes: *The figure shows the development of the average number of antitoxin bottles per 1,000 people by treatment intensity (below/above median) and all municipalities in our Massachusetts (MA) sample. We report three-year moving averages and use 1895 municipality population size as weights. The vertical line indicates the first year when antitoxin became freely available for adoption.*

From the SBH reports, we also collected annual data on the number of infections (cases) from

[13]Summary statistics of the key variables are reported in Appendix Table A.1.

1891 to 1914 for the diseases: diphtheria, scarlet fever, typhoid fever, measles, and smallpox. The coverage increase from 68 municipalities in 1891 to around 300 municipalities at end of our sample period. The SHB notes that their case data are very likely to suffer from under-reporting, so these data should be treated with caution.

4 Estimation strategy

The main challenge for estimating the causal effect of antitoxin on measures of population health is reverse causality—the demand for the new medication is higher during disease outbreaks. In this section, we describe our strategy for addressing this difficulty.

Using our data on the local adoption of antitoxin, we start the analysis by estimating the relationship between antitoxin and population health within the framework of the following two-way fixed effect estimation equation:

$$y_{mt} = \beta antitoxin_{mt} + \mu_m + \mu_{ct} + \mathbf{X}'_{mt}\Gamma + \varepsilon_{mt}, \tag{1}$$

where y_{mt} is some measure of population health (e.g., life expectancy, mortality rates by disease or age) in municipality m at year t, where our initial focus is on the diphtheria mortality rate and life expectancy at birth, but later we also report estimates for other diseases, infections, and fertility. The municipality specific adoption of antitoxin is given by $antitoxin_{mt}$, which is antitoxin bottles per 1,000 people supplied to municipality m each year t from the SBH. Since the Board started its supply of antitoxin in 1895, this variable is by construction zero for all municipalities prior to that year during the pre-antitoxin period. Municipality and county-year fixed effects are given by μ_c and μ_{ct}, respectively, and $\mathbf{X}'_{mt}$ is a vector containing various pre-antitoxin municipality characteristics interacted with year fixed effects. The error term is ε_{mt} and standard errors are clustered at the municipality level. The regression is weighted by the municipality population size in 1895, so the estimates reflect changes for the average person in Massachusetts.

While the baseline sample period is annually from 1880 to 1914, some specifications exploit long-difference variation instead. The sample always ends in 1914 because of data availability on the supply of antitoxin, but this year also marks the beginning of the possibility of immunization via the development of the diphtheria vaccine as mentioned in Section 2.

In the two-way fixed effect model, we can control for time invariant differences across municipalities and secular trends, but the LS estimate of β in Equation (1) is likely to biased due to reverse causality: the demand for antitoxin is higher during a diphtheria outbreak, for example. If this effect is sufficiently strong it might even seem as if antitoxin *reduced* population health from estimating β.

In our framework, we address this estimation problem by exploiting the fact that antitoxin became suddenly and freely available in 1895 and that some municipalities stood to benefit more from this technology, in terms of potential mortality reductions, as they were systematically more plagued by diphtheria prior to antitoxin. For example, we observe in our mortality data that

diphtheria was more widespread in urban municipalities, which is to be expected given that the disease is airborne.[14] A similar empirical strategy is applied in studies such as Acemoglu and Johnson (2007), Bleakley (2007), or Ager et al. (2018).[15]

The first-stage equation takes on the following form:

$$antitoxin_{mt} = \alpha treatment_m \times I_t \times (t - 1894) + \mu_m + \mu_{ct} + \mathbf{X}'_{mt}\Omega + \epsilon_{mt}, \tag{2}$$

where $treatment_m$ is the logged average pre-antitoxin (1891-94) diphtheria mortality rate, which is our cross-sectional measure of treatment intensity, I_t is a post-1895 indicator, and $(t - 1894)$ is a linear trend. The remaining variables are defined above. Both $antitoxin_{mt}$ and I_t are zero for all municipalities prior to antitoxin, and $\hat{\alpha}$ quantifies how differences in pre-diphtheria mortality rates translate into differences in the the adoption speed of antitoxin. The linear-trend specification is motivated by the gradual adoption as observed in Figure 5, and if we find that $\hat{\alpha} > 0$ this means that municipalities more affected by diphtheria prior to 1895 have stronger trends in the adoption of antitoxin when the technology became available.[16]

We will exploit the sharp cut-off date in 1895, when the SBH started to supply antitoxin, to conduct falsifications exercises in which we show that treatment cannot explain changes in diphtheria mortality and life expectancy in the time before antitoxin. As mentioned, the baseline specification uses annual data, but we also report estimates based on so-called "long differences", which are more directly comparable with these falsification tests. e.g.

Finally, we are also going to present evidence from a model in which the disease data are stacked, and the main right-hand-side variable in Equations (1) and (2) are interacted with an indicator for diphtheria, which thus is the disease expected to be most influenced by the supply of antitoxin. This model opens the possibility of controlling for interaction fixed effects (municipality-by-year, cause-of-death-by-year, and cause-of-death-by-municipality), and modifies the identifying assumption (Olden and Møen 2020). In our case, the key condition now becomes similar trends in the *difference* in the mortality rate of diphtheria and other diseases across municipalities with different treatment intensities.

5 Results

5.1 Determinants of antitoxin diffusion

We begin the empirical investigation by studying how the adoption of antitoxin is related to different municipality characteristics by estimating versions of Equation (2), which we later use as the first stage in the 2SLS framework. The estimates are reported in Table 1, where all regressions include

[14] In rare cases, milk products can also serve as a source for transmission.

[15] Although, often in these studies, specific data on the medical innovations are not available, and so it is not possible to document directly something similar to Equation 2.

[16] We obtain very similar results using alternative functional forms (e.g., concave instead of linear), but prefer the linear model for simplicity.

municipality and county-year fixed effects and are weighted by population size in 1890. The baseline estimate, reported in column 1, implies that 10 years into the antitoxin era the adoption rate is close to two bottles per 1,000 people in a municipality with average treatment intensity. If treatment intensity is increased by a one standard deviation, this number would be about 2.7 bottles per 1,000 people.

Column 2 of the Table 1 shows that this pattern of adoption is robust to controlling for the mortality rate from other respiratory infectious and waterborne diseases (i.e., bronchitis, tuberculosis, scarlet fever, pneumonia, whooping cough, measles, digestive, and typhoid) and the stroke mortality rate. These are constructed in a similar way as "treatment" by averaging over the pre-antitoxin years 1889 to 1894 and taking logs, in order to deal with the concern that "treatment" could to be correlated with other infectious diseases. The remaining columns control for distance to Boston, population density, dwellings per capita, number of rooms per dwellings and the foreign-born share (all measured in 1895, interacted with the indicator and a linear trend). The magnitude reduces to 1.1 bottle per 1,000 people (after 10 years) in the most conservative specification, reported in column 7. From the table, we also note that the rate of antitoxin adoption was higher in larger municipalities with higher mortality rates from other infectious diseases, located closer to Boston, and with living arrangements more conducive to disease transmission (see e.g., Ager et al. 2020).

5.2 Antitoxin and population health

This subsection reports evidence of how the adoption of antitoxin influenced population health outcomes. We start estimating the relationship by LS, controlling for municipality and county-year fixed effects, for the two main outcomes: diphtheria mortality rates (columns 1 and 2) and life expectancy at birth (columns 3 and 4) in Table 2. We report the results from two different specifications. The first one is a long-difference type, where the pre-period is 1880-1894 and the post-period is 1910-1914 (columns 1 and 3). The second one is the annual model from 1880 to 1914 (columns 2 and 4). We see that the estimated β's for diphtheria mortality is close to zero and statistically highly insignificant, whereas the corresponding estimate for life expectancy goes from positive (but insignificant) to negative. Overall, when estimating the relationship by LS, we cannot reject that the adoption of antitoxin did not improved population health.

Table 1: Antitoxin adoption by municipality characteristics

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)
treatment x I x (t-1894)	0.195***	0.168***	0.169***	0.128***	0.112***	0.115***	0.110***
	(0.030)	(0.029)	(0.029)	(0.026)	(0.028)	(0.026)	(0.027)
ln(infec. rate, 89-94) x I x (t-1894)		0.195***	0.197***	0.311***	0.141	0.007	-0.039
		(0.068)	(0.070)	(0.067)	(0.091)	(0.080)	(0.090)
ln(stroke rate, 89-94) x I x (t-1894)			0.022	0.012	-0.002	0.009	0.030
			(0.042)	(0.040)	(0.035)	(0.036)	(0.040)
dist Boston x I x (t-1894)				-0.008***	-0.006***	-0.005***	-0.005***
				(0.002)	(0.002)	(0.002)	(0.002)
pop density in 95 x I x (t-1894)					0.030***	0.027***	0.026***
					(0.010)	(0.009)	(0.009)
dwellings pr. capita in 95 x I x (t-1894)						0.232	0.289
						(0.229)	(0.226)
rooms pr. dwelling in 95 x I x (t-1894)						0.083***	0.079***
						(0.023)	(0.026)
fb share in 95 x I x (t-1894)							0.325
							(0.283)
Observations	8,505	8,505	8,365	8,365	8,365	8,365	8,365

Notes: *This table report how the number of antitoxin bottles per 1,000 people are related to baseline municipality characteristics: "treatment" is the logged diphtheria mortality rate averaged over the pre-antitoxin years 1891 to 1894; "ln(resp. rate, 89-94)" is the respiratory mortality rate, which is constructed in the same way as "treatment"; "ln(stroke rate, 89-94)" is the stroke mortality rate, which is also constructed the same way as "treatment"; "dist Boston" is the aerial distance to Boston; "pop density in 95" in population density in 1895 as measured by the number of people per 1,000 square miles; "dwellings pr. capita in 95" is the number of dwellings pr. capita; "rooms pr. dwelling in 95" is the number of rooms per dwellings in 1895; "fb share" in the number of foreign born pr. capita in 1895. These municipality characteristics have been interacted with a post-1895 dummy and a linear trend ($I \times (t - 1894)$). All regressions control for municplaity and county-year fixed effects and are weighted by population size in 1890. The sample period is annually from 1880 to 1914.*

Table 2: OLS estimates

VARIABLES	(1) diphteria rate	(2) diphteria rate	(3) life expectancy	(4) life expectancy
antitoxin p.c.	0.000	-0.005	0.050	-0.036
	(0.003)	(0.008)	(0.045)	(0.034)
Sample	1880-88 and 1910-14	annual 1880-14	1880-88 and 1910-14	annual 1880-14
Mean pre-y	0.994	0.888	41.33	41.38
N	3393	8530	2700	7875
N_clust	243	244	225	225

Notes: *This table reports OLS estimates of the relationship between adoption of antitoxin per 1,000 people and the diphtheria mortality rate (columns 1 and 2) and life expectancy at birth (columns 3 and 4). Columns 1 and 3 use a pre-period from 1880 to 1888 and a post-period from 1910 to 1914. In columns 2 and 4, the sample period is annually from 1880 to 1914. All regressions control for municipality fixed effects and county-year fixed effects and are weighted with the municplaity population size in 1895. "Mean pre-y" is the mean of the outcome measured over the relevant pre-antitoxin period. Standard errors clustered at the municipality level in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

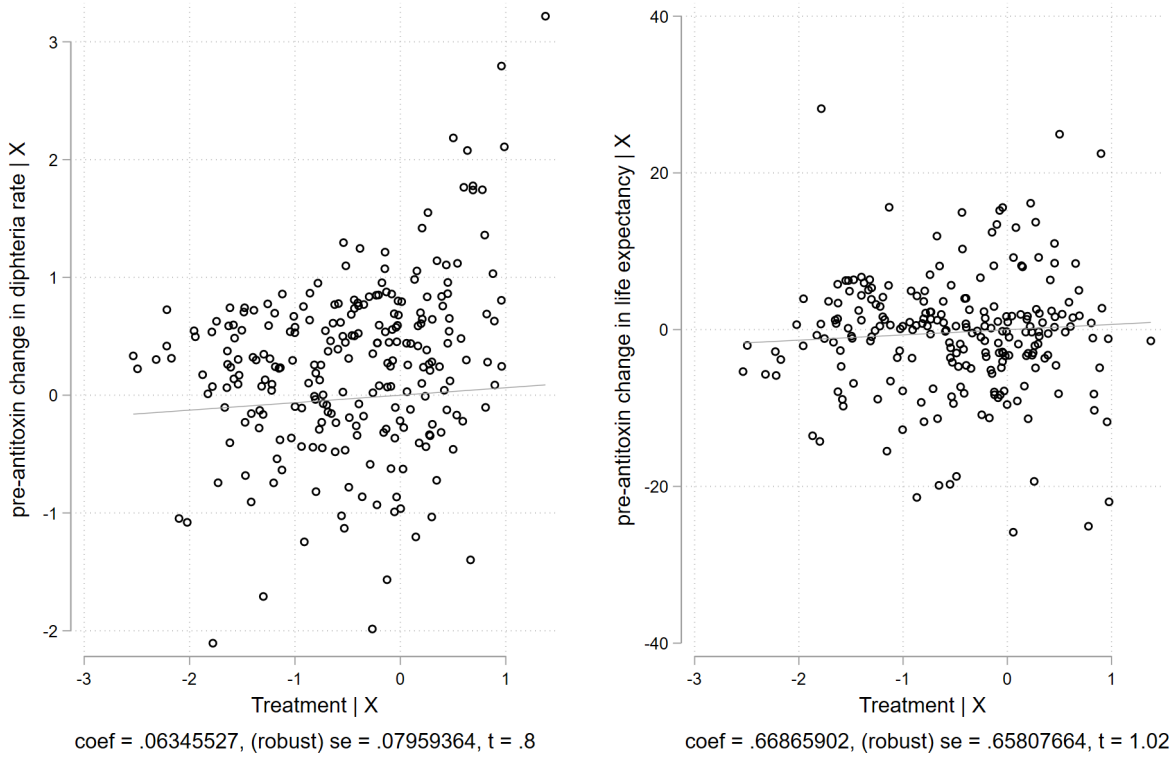
In order to isolate the effect of antitoxin on population health from the demand being higher when population health is lower (i.e., during disease outbreaks), we now exploit that some municipalities had more to benefit from the new technology in terms of potential health improvements as measured by the local severity of diphtheria prior to antitoxin ("*treatment*").

In Figure 6, we exploit that antitoxin only became available for adoption from 1895 onward to conduct a reduced-form falsification exercise, where we test if *treatment* is predictable of change in diphtheria mortality and life expectancy prior to the availability of antitoxin.[17] In the reported specifications, we average mortality (or life expectancy) over the years 1880-1883 and and 1890-93 (to reduce the noise from annual fluctuations) and regress the difference on *treatment*, controlling for county fixed effects As the more descriptive evidence (in Figures 3 and 4) indicates, we do not see any pre-antitoxin changes in the outcomes systematically related to the treatment measure, making it more plausible that any post-antitoxin reductions/increases in outcomes are, in fact, related to the spread of antitoxin.[18]

[17]We can only do this in a reduced-form sense, since adoption of antitoxin per definition is zero before 1895.

[18]This conclusion is robust to (1) choosing alternative pre-years to average over; (2) not controlling for county fixed effects; or (3) not weighting by population size in 1890.

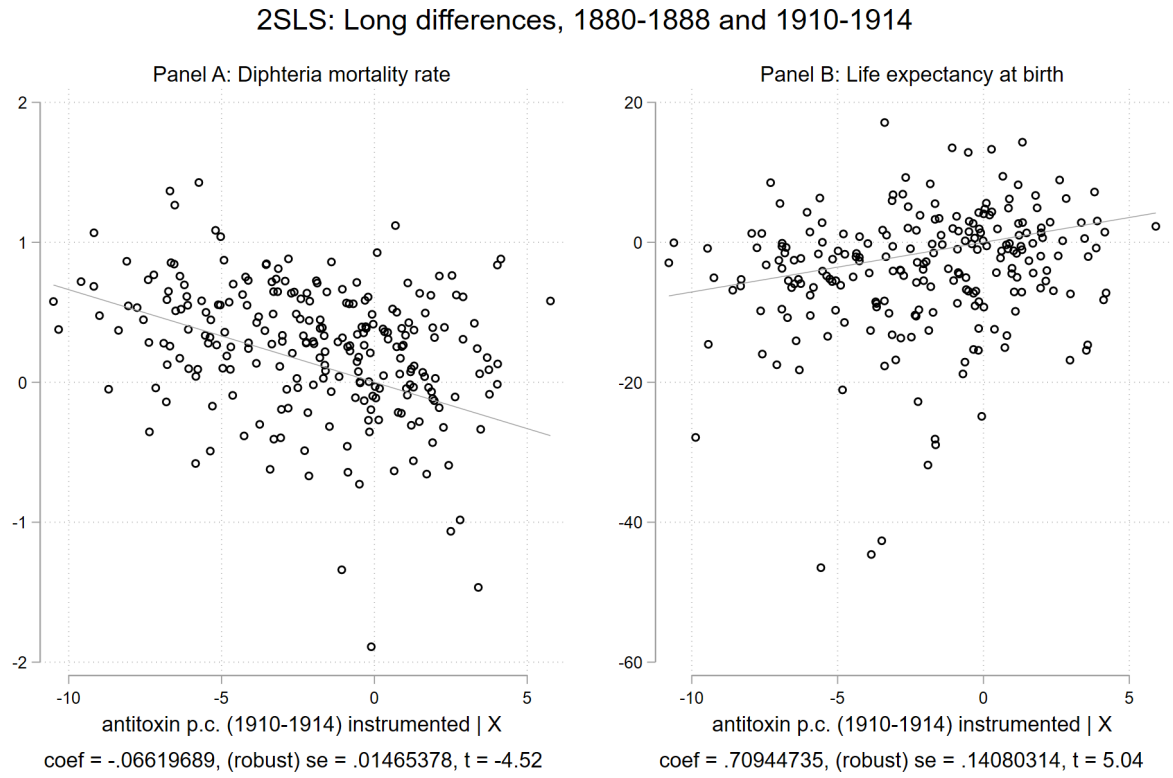
Figure 6: Falsification, 1880-1883 and 1890-1893



Notes: *The figure shows the relationship between the pre-antitoxin change in the diphtheria mortality rate and "treatment" in Panel A, and the pre-antitoxin change in life expectancy at birth and "treatment" in Panel B. The difference in the outcomes are based on four years averages (1880-1883 and 1890-1893) to reduced the noise coming from annual fluctuations in mortality and life expectancy.*

The main insights of the paper are essentially summarized graphically in Figure 7. These are akin to long-difference specifications, but we average the variables over several years in the pre-period (1880 to 1888), and several years at the end of the sample in the post-period (1910 to 1914) to reduce noise coming from annual fluctuations in mortality, which is particular pronounced for smaller municipalities. We instrument the adoption of antitoxin with "treatment" interacted with post-1895 dummy. One advantage of a long-difference model is that we do not to need assume a specific functional form for the diffusion of antitoxin as we, for example, do with the linear-trend break model in Equation (2) when using data annually. The estimates are statistically significant and imply that an increase of one bottle of antitoxin per 1,000 people reduces the diphtheria mortality rate by 0.06 per 1,000 people and increases life expectancy by 0.7 years.

Figure 7: Antitoxin effects on diphtheria and life expectancy



Notes: *This figure shows the relationship between the change in diphtheria mortality rate (Panel A) and life expectancy (Panel B) during the antitoxin period and the adoption of antitoxin, instrumented with "treatment" interacted with a post-1895 indicator. The pre-period is the average of the variables from 1880 to 1888 and the post-period is the average of the variables from 1910 to 1914.*

In columns 1 and 3 of Table 3, we also report estimates from long-difference specifications, using the same pre and post-period, but here we simply restrict the pre-period to 1880-1888 and the post-period to 1910-1914 and do not average the variables "manually".[19] We obtain very similar conclusions from these. In columns 2 and 4, which we consider to be our baseline specification, we use data annually starting in 1880 and model the diffusion of antitoxin as specified in Equation (2) with a linear trend break from 1895 onward. The estimates have the expected negative and positive signs and are both statistically significant with a strong first stage. The magnitudes are relative similar to the long-difference specification. We can use this specification to quantify how a counterfactual delay in the invention (or in the free supply from SBH) of antitoxin would influence population health.[20] The first-stage estimate is 0.2 (in column 4), and so a 10-year delay would reduce the adoption by 2 bottles of antitoxin per 1,000 people in a municipality with the average

[19]The advantage of doing the averaging manually is that we show the results graphically in a transparent way as we do in Figure 7.

[20]The counterfactual where the delay is in the free supply of antitoxin by the SBE assumes that it would not be possible for the municipalities to adopt antitoxin elsewhere.

treatment intensity, which again would reduce life expectancy at birth by one year. Relative to the mean of life expectancy, measured prior to antitoxin, this corresponds to increase of 2.5 percent.[21]

Table 3: Effects on the diphtheria mortality rate and life expectancy at birth: 2SLS estimates

	(1)	(2)	(3)	(4)
VARIABLES	diphteria rate	diphteria rate	life expectancy	life expectancy
antitoxin p.c.	-0.051***	-0.081***	0.590**	0.491**
	(0.015)	(0.016)	(0.229)	(0.209)
Sample	1880-88 and 1910-14	annual 1880-14	1880-88 and 1910-14	annual 1880-14
Mean pre-y	0.994	0.888	41.33	41.38
N	3393	8495	3150	7875
N_clust	243	243	225	225
widstat	44.04	43.93	42.46	40.87

Notes: *This table reports 2SLS estimates of how the adoption of antitoxin per 1,000 people affects the diphtheria mortality rate (columns 1 and 2) and life expectancy at birth (columns 3 and 4). Columns 1 and 3 use treatment intensity interacted with a post-1895 indicator as the instrument for antitoxin, where the pre-period is from 1880 to 1888 and the post-period is from 1910 to 1914. Columns 2 and 4 use treatment intensity interacted with a post-1895 indicator and a linear trend with a sample period annually from 1880 to 1914. "Mean pre-y" is the mean of the outcome measured over the relevant pre-antitoxin period. All regressions control for municipality fixed effects and county-year fixed effects and are weighted by the municplaity population size in 1895. Standard errors clustered at the municipality level in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

Table 4 reports different robustness checks based on the baseline annual specification. First, we find that our previous quantification is robust to controlling for various municipality characteristics possibly associated with urbanity and infectious diseases. These include the pre-antitoxin mortality rates from other infectious respiratory and waterborne diseases and strokes, as well as distance to Boston, the number of doctors per capita, living arrangements, dwelling size, population size, and foreign-born share (all measured in 1880 and interacted with a full set of year fixed effects). Second, we obtain similar insights using alternative functional forms: In columns 2 and 6, the reported 2SLS estimates are based on a level-log model, while in column 7 it is based on a log-log model. The latter functional-form is better suitable for life expectancy at the outcomes, since here we only have very few zeros, while this is not the case with the diphtheria mortality rate, which in some years could be zero in smaller places. Therefore, we report a reduced-form Poisson estimate in column 4. This model takes into account that the outcome variable is highly skewed. Finally, columns 3 and 8 also report reduced-form coefficients estimated by OLS, which is standard to report in a 2SLS framework. In all the reported specifications we reach the conclusion that antitoxin improved population health.

In Table 5, the outcomes are age-specific mortality rates (ages 0 to 10). We find negative

[21]Because of data availability of this death certificates, used to calculate life expectancy, the sample is reduced by 18 municipalities in columns 3 and 4.

Table 4: Robustness to controls and specification

VARIABLES	(1) diph rate	(2) diph rate	(3) diph rate	(4) diph rate	(5) life exp	(6) life exp	(7) ln(life exp)	(8) life exp
antitoxin p.c.	-0.133***				0.720*			
	(0.038)				(0.393)			
ln antitoxin p.c.		-0.844***				4.578**	0.143**	
		(0.187)				(2.263)	(0.064)	
treat x I x (t-1894)			-0.016***	-0.024***				0.100***
			(0.001)	(0.004)				(0.037)
Add controls	yes	no	no	no	yes	no	no	no
Spec.	2SLS level-level	2SLS level-log	r-f OLS	r-f Poisson	2SLS level-level	2SLS level-log	2SLS log-log	r-f OLS
N	8355	7261	8495	8475	7735	6698	6697	7875
N_clust	239	243	243	243	221	225	225	225
widstat	16.74	27.57			16.41	28.48	28.48	

Notes: *This table reports various robustness checks to the baseline annual (1880-1914) linear trend-break model, similar to the ones reported in Columns 2 and 4 of Table 3. The outcomes are the diphtheria mortality rates (columns 1-4) and life expectancy at birth (columns 5-8). Columns 1 and 5 include is pre-antitoxin respiratory, waterborne and stroke mortality rates, distance to Boston, population density, dwellings per capita, rooms per dwelling and the foreign-born share (all measured prior to antitoxin and interacted with a full set of year fixed effects). Columns 2, 6, and 7 use level-log and log-log functional forms. Column 3 and 8 report reduced-form coefficients estimated by OLS, while column 4 report a reduced-form coefficient estimated by Poisson. All regressions control for municipality fixed effects and county-year fixed effects and are weighted by the municplaity population size in 1890. Standard errors clustered at the municipality level in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

coefficients across all ages and most are statistically significant at conventional levels. However, evaluated relative to the relevant pre-antitoxin mean mortality rate, the negative effects are most pronounced at ages 2-5, which is consistent with the historic age profile for diphtheria mortality. For example, the estimated coefficient at age 4 implies that a one-bottle increase in antitoxin leads to 5 percent decrease in the mortality rate.[22]

In Table 6, we report the effects on the infant and child mortality rates (for all and by gender). While these variables are often used in the literature as important markers of (child) population health, this tables also serves as a check of the robustness for denominator data used to construct the rates up until now. For diphtheria mortality (and other diseases, reported below), we use population counts from the state census, which are reported in the vital statistics, and for the life tables we use population counts by single-year age groups from the federal census.[23] Since deaths during the first year of life often occurs within the first couple of months, it is common to use the number of live births to represent the population at risk in the construing the infant mortality *rate*. We follow this approach, but obtain the birth counts from individual birth records. The child mortality rate includes the ages 1 to 4, and we use the individual death and birth records to construct the number of people in that age group (see details in the data appendix). The effects are negative across the board. For the infant mortality rate, this is driven by males, whereas for the child mortality rates the effects are similar across genders. Consistent with the estimates reported in the previous table, the percent declines are largest for the child mortality rates (relative to the pre-antitoxin mean). In particular, the estimate in column 4 suggests that a one-bottle increase in antitoxin per 1,000 people reduces the child mortality rate by about 3 percent.

The next two tables show the antitoxin effects on other causes of death, cases, and fertility. In column 1 of Table 7, we report a negative and significant estimate for the diphtheria mortality ratio (i.e., the number of diphtheria deaths per 1,000 deaths), indicating that the effect of antitoxin diphtheria was of first order and that it thereby reduced the mortality burden of diphtheria. In the second column, we do see some evidence suggesting that antitoxin reduced fatality (diphtheria death per 1,000 cases of diphtheria), but the estimate is not statistically significant at any conventional levels. In columns 3 and 4, we control for the pre-antitoxin mortality rates of the outcomes (as explained above) interacted with the linear trend break in order not to capture changes in other diseases driven by other interventions. The antitoxin effect on strokes is negative and significant, while the effect on other infectious diseases (i.e., bronchitis, tuberculosis, scarlet fever, pneumonia, whooping cough, measles, typhoid, and digestive) is small in magnitude (relative to the pre-antitoxin mean) and insignificant once we controls for the its initial mortality rate. In general, second-order mortality effects are not all that surprising given the complications of diphtheria and secondary

[22]The outcomes are so-called q-type mortality rates, which are being used in the calculation of the life tables. The relationship between the m-type mortality rate and the q-type is given by: $q_x = m_x/(1 + (1 - a_x)m_x)$, where m_x is the (m-type) age-specific mortality rat and a_x is the expected number of months an individual at a given age lives within the calendar year. We use $a_0 = 1/3$ and for all other ages, in the tables, a_x is equal to an half.

[23]We obtain similar results using population counts from the federal censuses to construct the cause-specific mortality rates. Results not reported but available upon request.

Table 5: Effects on age-specific mortality rates

VARIABLES	(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.	-3.367*	-1.504*	-0.967***	-0.620***	-0.695***	-0.496***	-0.307**	-0.082	-0.132	-0.237*	-0.031
	(1.900)	(0.837)	(0.315)	(0.210)	(0.209)	(0.163)	(0.119)	(0.100)	(0.096)	(0.130)	(0.076)
Mean pre-y	194.5	57.23	25.69	17.86	13.73	10.76	8.427	7.222	5.840	4.921	4.115
N	7873	7869	7869	7869	7870	7870	7874	7871	7872	7872	7874
N_clust	225	225	225	225	225	225	225	225	225	225	225
widstat	40.92	40.82	40.78	40.82	40.81	40.74	40.87	40.76	40.89	40.89	40.87

Notes: *This table reports the effects on q-type age-specific mortality rates (ages 0-10), using the baseline annual (1880-1914) linear trend-break model, similar to the ones reported in Columns 2 and 4 of Table 3. The outcomes are expressed in per 1,000 individuals of the relevant age groups. All regressions control for municipality fixed effects and county-year fixed effects and are weighted by the municplaity population size in 1890. Standard errors clustered at the municipality level in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

Table 6: Effects on the infant and child mortality rates

VARIABLES	(1) infant rate	(2) infant rate female	(3) infant rate male	(4) child rate	(5) child rate female	(6) child rate male
antitoxin p.c.	-1.184	-1.879*	-0.349	-0.765***	-0.719***	-0.808***
	(1.049)	(1.060)	(1.156)	(0.211)	(0.212)	(0.234)
Mean pre-y	153.6	140.5	164.9	23.40	22.71	24.08
N	7647	7641	7645	7860	7860	7860
N_clust	225	225	225	225	225	225
widstat	38.81	38.82	38.75	40.81	40.81	40.81

Notes: *This table reports effects on infant and child mortality rates, using the baseline annual (1880-1914) linear trend-break model, similar to the ones reported in Columns 2 and 4 of Table 3. The infant rate uses the number of live birth to measure the population at risk, while the child rates uses the population ages 1-4. The rates are expressed in per 1,000 births and or children. All regressions control for municipality fixed effects and county-year fixed effects and are weighted by the municplaity population size in 1890. Standard errors clustered at the municipality level in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

infections. We find no effects on the deaths from accidents, which is to expected (columns 5), whereas there is some evidence that antitoxin reduced the crude birth, albeit the effect size is small in magnitude.[24]

In Table 8, we "stack" the cause-of-death mortality rates and case rate, so the dimension in the panel becomes municipality-year-disease. We can control for municipality-year FE, disease-year-county FE, and municipality-disease FE, since we interact the main RHS variables in Equations 1 and 2 with an indicator for diphtheria (since our working hypothesis is that antitoxin should have a first order effect on this disease). While the baseline model can be interpreted along the lines of double-difference estimation (with a continuous measure of treatment intensity), the stacked model is more similar to triple-differences estimation. In the first column, all 13 causes-of-deaths, which we have obtained from the tabulated vital statistics are included as controls, while in columns 2-5 we vary the included control causes (e.g.,, column 3 only include childhood diseases as controls). In all five specification, the coefficient is estimated in the range -0.11 to -0.06, which is close the baseline double-difference estimate. This demonstrates that municipalities with higher rates of antitoxin adoption experienced larger declines in their diphtheria mortality rate relative to other diseases. The final column shows the effect on the number of diphtheria infections (or cases) relative to number of scarlet fever infections, typhoid fever infections, measles infections, and smallpox fever infections. The coefficient is negative, but far from being statistically significant

[24]Similar results are obtained for accidents if we control for its initial mortality rate.

Table 7: Other vital outcomes

VARIABLES	(1) diph ratio	(2) fatality	(3) stroke	(4) infec rate	(5) accidents	(6) birth rate
antitoxin p.c.	-3.514***	-9.386	-0.043***	-0.062	0.027	-0.419*
	(0.699)	(13.789)	(0.011)	(0.053)	(0.032)	(0.232)
ln(stroke rate, 89-94) x I x (t-1894)			-0.005			
			(0.003)			
ln(infec. rate, 89-94) x I x (t-1894)				-0.102***		
				(0.026)		
Mean pre-y	43.27	442.7	0.535	7.664	0.729	26.78
N	8495	3485	8355	8505	8495	8495
N_clust	243	236	239	243	243	243
widstat	42.55	36.81	42.08	34.36	42.55	42.55

Notes: *This table reports effects on the diphtheria death ratio (columns 1), fatality (column 2) rate, other causes of deaths (columns 3, 4, and 5), and the crude birth rate (column 6), using the baseline annual (1880-1914) linear trend-break model, similar to the ones reported in Columns 2 and 4 of Table 3. All variables are expressed in per 1,000. All regressions control for municipality fixed effects and county-year fixed effects and are weighted by the municplaity population size in 1890. Standard errors clustered at the municipality level in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

Table 8: Stacked model

VARIABLES	(1) mortality	(2) mortality	(3) mortality	(4) mortality	(5) mortality	(6) cases
antitoxin p.c. x I	-0.080***	-0.109***	-0.087***	-0.071***	-0.058***	-0.043
	(0.019)	(0.032)	(0.020)	(0.017)	(0.016)	(0.065)
Controls	all	exogenous	childhood	declining	waterborne	case data
N	92134	19743	26324	39486	19743	15650
N_clust	243	243	243	243	243	231
rkf	25.25	24.63	24.63	24.63	24.63	28.40

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 all 13 controls diseases (typhoid, tuberculosis, pneumonia, scarlet fever, measles, whooping cough, bronchitis, accidents, suicides, childbirth, meningitis, strokes, and digestive diseases). Column 2 only includes "exogenous" causes as controls (accidents and suicides). Column 3 only include childhood diseases as controls (scarlet fever, whooping cough, measles). Column 4 only includes diseases where we observe secular declines during the pre-antitoxin period as controls (typhoid, tuberculosis, scarlet fever, meningitis, and digestive diseases). Column 5 only include waterborne diseases as controls (typhoid and digestive diseases). In column 6 the outcome in the infection rate (or cases per 1,000 people), where the controls the infection rates from scarlet fever, typhoid, measles, and smallpox. All regressions control municipality-year FE, disease-year-county FE, and municipality-disease FE and are weighted by the municplaity population size in 1890. Standard errors clustered at the municipality level in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

6 Concluding remarks

Based on our analysis of death certificates and municipality data on the adoption of antitoxin, we have found that the provision of free antitoxin increased life expectancy in Massachusetts. Our estimates indicate that a 10-year delay in the adoption of antitoxin would have reduced life expectancy by approximately one year. This effect was primarily driven by reductions in child mortality. While these findings suggest that medicine may have played a greater role in the increase of life expectancy in Massachusetts than previously thought, it is unlikely that the impact of antitoxin would have been similar throughout the United States, as many other states did not implement policies providing free supplies of antitoxin to their citizens. However, this result may help to explain why life expectancy in Massachusetts increased at a similar or faster pace compared to many other U.S. states during this period, despite being one of the most urbanized states at the turn of the 20th century.

Our research design used pre-diphtheria mortality rates to identify the effect of antitoxin on mortality. In the first-stage regression, we observed a significant adoption gradient in physical distance to Boston, for example. We believe that future research could benefit from our detailed local adoption data, which to our knowledge are rarely available in historical settings, to further study such possible adoption gradients in greater detail. Understanding these patterns is important given that antitoxin did save lives.

References

Acemoglu, Daron and Simon Johnson, "Disease and development: the effect of life expectancy on economic growth," *Journal of political Economy*, 2007, *115* (6), 925–985.

Ager, Philipp, Casper Worm Hansen, and Peter Sandholt Jensen, "Fertility and early-life mortality: Evidence from smallpox vaccination in Sweden," *Journal of the European Economic Association*, 2018, *16* (2), 487–521.

_ , James J Feigenbaum, Casper Worm Hansen, and Hui Ren Tan, "How the other half died: Immigration and mortality in US cities," Technical Report, National Bureau of Economic Research 2020.

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.

_ , Vincenzo Atella, Jay Bhattacharya, Valentina Conti, Iván Mejía-Guevara, and Grant Miller, "Technological Progress and Health Convergence: The Case of Penicillin in Postwar Italy," *Demography*, 2021, *58* (4), 1473–1498.

Anderson, D. Mark, Kerwin Kofi Charles, and Daniel I. Rees, "Public Health Efforts and the Decline in Urban Mortality," *NBER Working Paper 25027*, 2019.

_ , _ , Claudio Las Heras Olivares, and Daniel I. Rees, "Was the First Public Health Campaign Successful?," *American Economic Journal: Applied Economics*, 2019, *11* (2), 143–75.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro, "Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century," *Journal of Political Economy*, 2016, *124* (1), 105–159.

Beach, Brian and W Walker Hanlon, "Coal smoke and mortality in an early industrial economy," *The Economic Journal*, 2018, *128* (615), 2652–2675.

Berkes, Enrico, Ezra Karger, and Peter Nencka, "The Census Place Project: A Method for Geolocating Unstructured Place Names," *Unpublished Working Paper*, 2021.

Bhalotra, Sonia R, Atheendar Venkataramani, and Selma Walther, "Fertility and labor market responses to reductions in mortality," Technical Report, National Bureau of Economic Research 2022.

Bleakley, Hoyt, "Disease and development: evidence from hookworm eradication in the American South," *The quarterly journal of economics*, 2007, *122* (1), 73–117.

Cain, Louis P, Price V Fishback, and Paul W Rhode, "The oxford handbook of: American economic history," in "The Oxford Handbook of American Economic History," Oxford University Press, 2018, pp. 1–442.

Clay, Karen, Peter Juul Egedesø, Casper Worm Hansen, Peter Sandholt Jensen, and Avery Calkins, "Controlling tuberculosis? Evidence from the first community-wide health experiment," *Journal of Development Economics*, 2020, *146*, 102510.

Condran, Gretchen A, "The elusive role of scientific medicine in mortality decline: diphtheria in nineteenth-and early twentieth-century Philadelphia," *Journal of the history of medicine and allied sciences*, 2008, *63* (4), 484–522.

Condran, Gretchen A. and Eileen Crimmins, "Mortality differentials between rural and urban areas of states in the northeastern United States 1890?1900," *Journal of Historical Geography*, 1980, *6* (2), 179–202.

Costa, Dora L., "Health and the Economy in the United States from 1750 to the Present," *Journal of Economic Literature*, 2015, *55* (3), 503–70.

Crum, Frederick S, "A statistical study of diphtheria," *American Journal of Public Health*, 1917, *7* (5), 445–477.

Cutler, David and Grant Miller, "The role of public health improvements in health advances: the twentieth-century United States," *Demography*, 2005, *42* (1), 1–22.

_ , Angus Deaton, and Adriana Lleras-Muney, "The determinants of mortality," *Journal of economic perspectives*, 2006, *20* (3), 97–120.

Egedesø, Peter Juul, Casper Worm Hansen, and Peter Sandholt Jensen, "Preventing the white death: tuberculosis dispensaries," *The Economic Journal*, 2020, *130* (629), 1288–1316.

Eriksson, Katherine, Peter Z Lin, and Gregory Niemesh, "Immigration and Infant Mortality in Massachusetts: Evidence from the Age of Mass Migration," 2020.

Ferrie, Joseph P. and Werner Troesken, "Water and Chicago's mortality transition, 1850–1925," *Explorations in Economic History*, 2008, *45* (1), 1–16.

Fogel, Robert W, "Economic growth, population theory, and physiology: the bearing of long-term processes on the making of economic policy," 1994.

Fogel, Robert William, *The escape from hunger and premature death, 1700-2100: Europe, America, and the Third World*, Vol. 38, Cambridge University Press, 2004.

George, Ruggles, "Diphtheria Mortality," *The Public Health Journal*, 1923, *14* (9), 387–401.

Glover, James Waterman, *United States Life Tables, 1890, 1901, 1910, and 1901-1910*, US Government Printing Office, 1921.

Goldin, Claudia and Adriana Lleras-Muney, "XX¿ XY?: The changing female advantage in life expectancy," *Journal of health economics*, 2019, *67*, 102224.

Gutman, Robert, *The accuracy of vital statistics in Massachusetts, 1842-1901*, Columbia University, 1956.

_ , "Birth and Death Registration in Massachusetts. IV. The System Attains Its Basic Goals, 1870-1900," *The Milbank Memorial Fund Quarterly*, 1959, *37* (4), 386–417.

Hacker, J David, "Decennial life tables for the white population of the United States, 1790–1900," *Historical methods*, 2010, *43* (2), 45–79.

Haines, Michael R., "Massachusetts State Censuses, 1855-1915," 2022.

Jayachandran, Seema, Adriana Lleras-Muney, and Kimberly V Smith, "Modern medicine and the twentieth century decline in mortality: evidence on the impact of sulfa drugs," *American Economic Journal: Applied Economics*, 2010, *2* (2), 118–46.

Liebenau, Jonathan M, "Public health and the production and use of diphtheria antitoxin in Philadelphia," *Bulletin of the History of Medicine*, 1987, *61* (2), 216–236.

McKeown, Thomas, "The modern rise of population," 1976.

Moehling, Carolyn M and Melissa A Thomasson, "Saving babies: the impact of public education programs on infant mortality," *Demography*, 2014, *51* (2), 367–386.

Oeppen, Jim and James W Vaupel, "Broken limits to life expectancy," 2002.

Olden, Andreas and Jarle Møen, "The triple difference estimator," *NHH Dept. of Business and Management Science Discussion Paper*, 2020, (2020/1).

Preston, Samuel H and Michael R Haines, "The social and medical context of child mortality in the late nineteenth century," in "Fatal Years: Child Mortality in Late Nineteenth-Century America," Princeton University Press, 1991, pp. 3–48.

Ruggles, S, S Flood, S Foster, R Goeken, J Pacas, M Schouweiler, and M Sobek, "IPUMS USA: Version 11.0 [dataset]," *Minneapolis, MN: IPUMS*, 2021, *10*, D010.

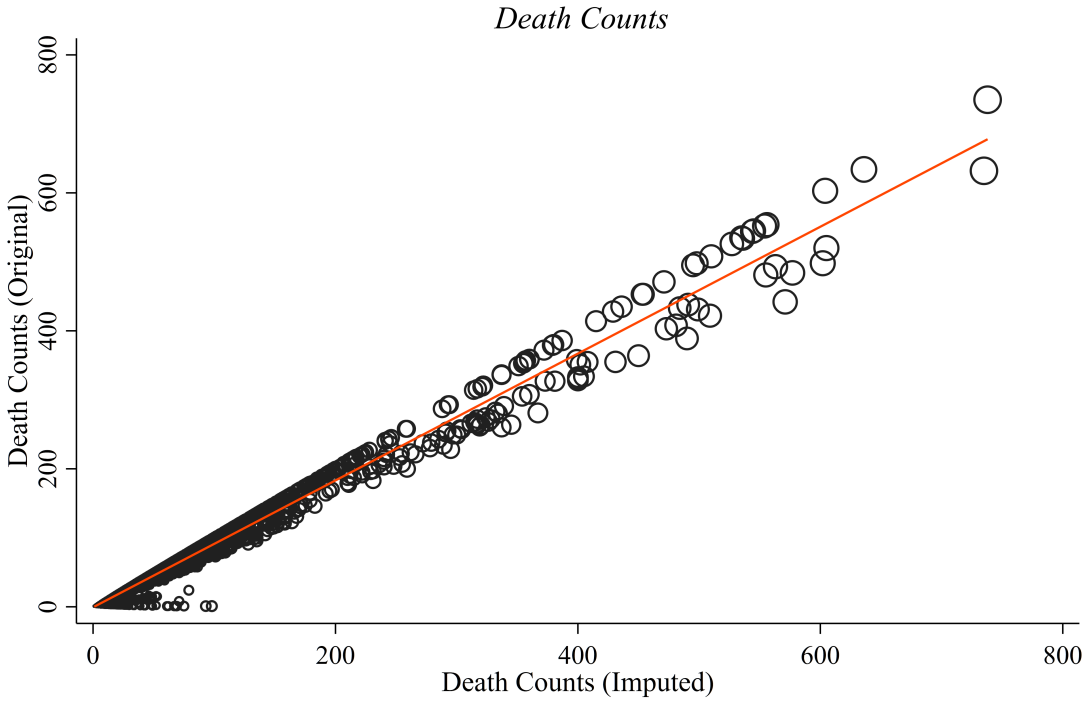
A Online Appendix

A.1 Individual death records and mortality rates by age

We aggregate individual death records at the annual municipality level for a given age to obtain age-specific death counts. We use them to compute the infant mortality rate, the child mortality rate, and life tables, where life expediencies (at various ages) can be extracted, although our regression analysis focuses on life expectancy at birth.

Individual death records have been digitized and are available as part of the data collection Massachusetts Deaths, 1841-1915, through FamilySearch.org. For the death records not reporting decedents' ages or failing to digitize accurate ages, we distribute the number of such records across ages 0 to 100, based on the age distribution of other death records with known ages in the same municipality, year, and other demographic traits. More specifically, we separate all death records into groups, which are defined by a municipality, death year, death season (April-September as warm season, and October-March as cold season), the decedent's sex, the decedent's nativity (born in Massachusetts/other U.S. states/foreign country/unknown), and the decedent's marital status (single/ever-married/unknown). Then, we distribute the number of age-missing death records across ages 0 to 100, based on the age distribution of all death records with known ages in the same group. The rich information on death records allows us to group decedents in such a disaggregated way and make our imputation more accurately by using the age-distribution of similar decedents. We compare the death counts which include or exclude the redistributed age-missing death records in Figure A.1.

Figure A.1: 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.*

The individual micro data, from the federal censuses (IPUMS) for the years 1880, 1900, 1910, allow us to calculate population sizes for single-year age groups by municipality. We linear interpolate in-between census years and combine them the imputed municipal annual death counts by age to calculate age-specific mortality rates for the ages 0, 1, .., 100 for each calendar year between 1880 and 1914, which are used in the construction of the life tables.[25] The construction of the life tables is explained in the main text.

In order to assess the sensitivity of the use of population data from the federal census, we construct other measures of age-specific mortality rates using alternative denominator data sources.[26] In particular, we calculate the infant mortality rate at the annual municipality level by combining the death counts at age 0 and the birth counts aggregated from individual birth records, which are also digitized and available through FamilySearch.org. Specifically, the infant mortality rate is defined as follows:

[25]We extrapolate the population data for the years 1911 to 1914.

[26]In the calculation of the cause-specific mortality rates, we use total population data from the state census, which are available every fifth year, but they are not available by single-year age groups in all the state census years.

$$IMR_{mt} = \frac{Deaths^{0}_{mt}}{Births_{mt}} \tag{3}$$

where $Deaths^{0}_{mt}$ is the death counts for age 0 in municipality m and year t, and $Births_{mt}$ is the birth counts in the same municipality and year. As an additional check of robustness, we also use birth counts from the tabulated vital statistics (same source as cause- of-death data).

We calculate the mortality rate of children aged 1 to 4 at the annual municipality level as follows:

$$CMR_{mt} = \frac{Deaths^{1-4}_{mt}}{Pop^{1-4}_{mt}} \tag{4}$$

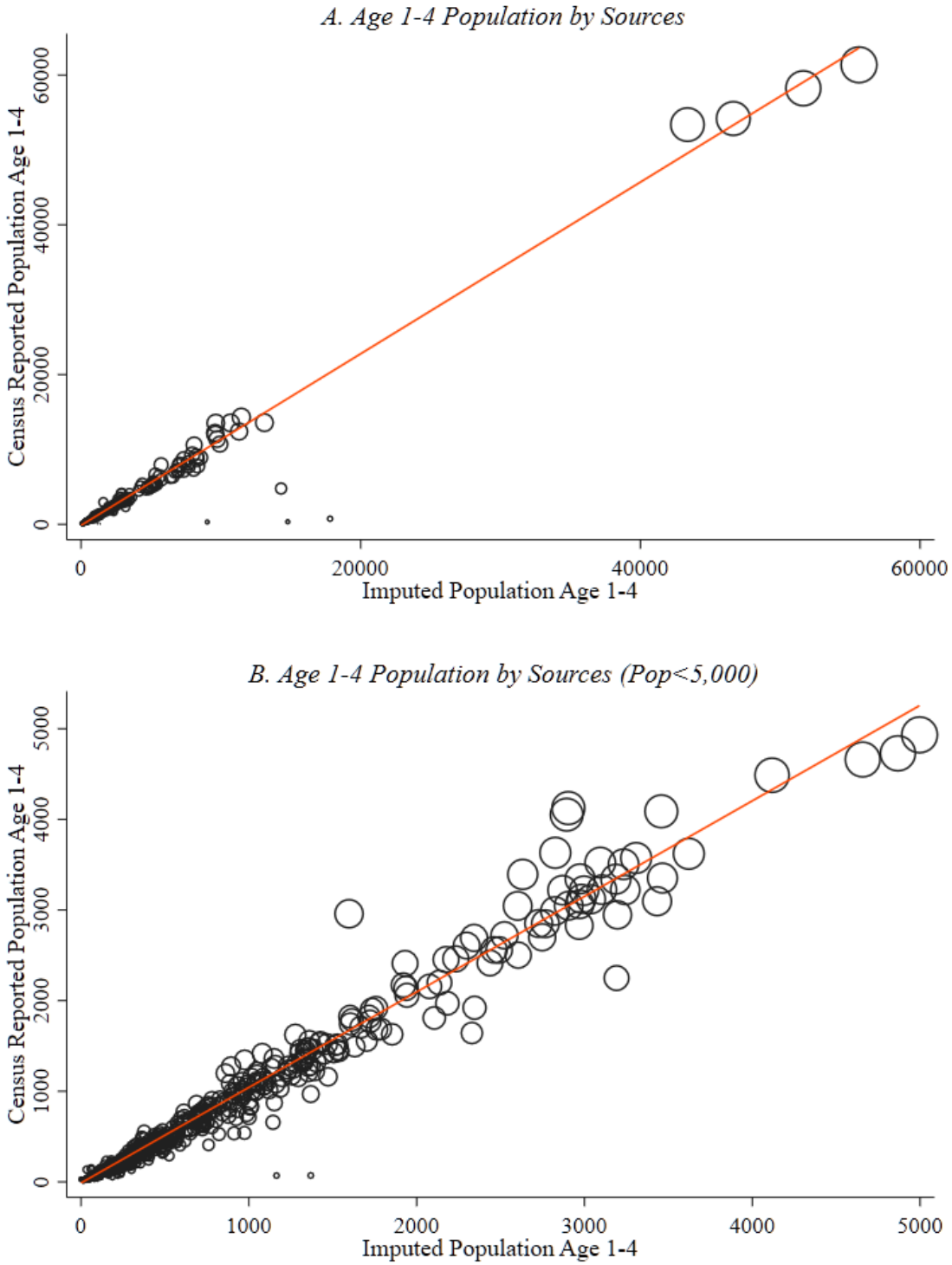
where $Deaths^{1-4}_{mt}$ is the deaths aged 1 to 4 in municipality m and year t, and Pop^{1-4}_{mt} is population aged 1 to 4 in the same municipality and year. Instead of using the federal interpolated population data, we impute the annual population aged 1-4 based on cumulative births and deaths of corresponding cohorts in prior years. This approach is also used by Alsan and Goldin (2019) and Eriksson et al. (2020), Specifically, the population is imputed as follows:

$$Pop^{1-4}_{mt} = \sum_{a=1}^{a=4} \left[Births_{m,t-a} - \textstyle\sum_{k=1}^{k=a} Deaths^{a-k}_{m,t-k} \right] \tag{5}$$

where $Births_{m,t-a}$ is is the number of births in municipality m and year $(t-a)$, and $Deaths^{a-k}_{m,t-k}$ is the number of deaths aged $(a-k)$ in municipality m and year $(t-k)$. In fact, the term $Births_{m,t-a}$ is the total number of children born a years ago (standing in year t), and the term $\sum_{k=1}^{k=a} Deaths^{a-k}_{m,t-k}$ is the cumulative deaths for children aged a between their birth year and the year t. We implicitly assume the migration of children is neglectable and all population changes were accounted by deaths.

We compare our imputed population with the census reported population in years when federal or state censuses were available. Figure A.2 below shows the imputed population fit well the census-reported population at municipality level in census years.

Figure A.2: Imputed and Census Reported Population Age 1-4



Notes: *This figures present scatterplots of imputed population aged 1-4 (x-axis) and census-reported population aged 1-4 (y-axis) in the census years of 1895, 1900, 1905, 1910, and 1915. Census-reported population come from federal censuses 1900 and 1910; and Massachusetts state census in 1895, 1905, and 1915. Each scatterplot 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*

A.2 additional results

Table A.1: Summary statistics by pre- and post-antitoxin periods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	pre-antitoxin					post-antitoxin				
VARIABLES	N	mean	p25	p50	p75	N	mean	p25	p50	p75
diphteria rate	244	0.888	0.646	0.846	1.154	244	0.312	0.227	0.319	0.422
life expectancy	226	41.37	35.46	40.94	46.04	226	45.15	41.45	44.88	49.51
child mortality rate	226	23.40	17.94	24.01	30.80	226	13.06	9.573	12.02	17.74
infant mortality rate	226	153.3	126.8	156.7	168.0	226	133.2	111.9	131.1	137.7
population	244	92,634	5,237	21,782	58,291	244	130,626	7,788	33,165	98,767
antitoxin p.c.	244	0	0	0	0	244	9.953	6.127	9.612	14.70
treatment	244	-0.495	-0.811	-0.416	0.199	244	-0.495	-0.811	-0.416	0.199

Notes: *This table reports summary statistics for selected key variables averaged over the pre-antitoxin period (1880-1894) and the post-antitoxin periods (1895-1914), using 1890 municipality population size as weight.*

Table A.2: Treatment balance

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln(infec. rate, 89-94)	0.953***	0.948***	1.028***	1.059***	0.893***	0.877***	0.857***	0.653**
	(0.256)	(0.257)	(0.229)	(0.225)	(0.235)	(0.238)	(0.249)	(0.277)
ln(stroke rate, 89-94)		-0.223**	-0.222**	-0.239**	-0.218*	-0.219**	-0.220**	-0.144
		(0.108)	(0.109)	(0.115)	(0.111)	(0.110)	(0.111)	(0.115)
dist Boston			-0.009***	-0.009***	-0.010***	-0.010***	-0.010***	-0.009**
			(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
doctors(1880)				73.765	76.181	74.499	68.370	82.284
				(68.529)	(67.977)	(68.322)	(73.140)	(72.698)
living arrangement(1880)					4.328	4.023	3.833	2.263
					(2.685)	(2.749)	(2.843)	(3.039)
dwelling size(1880)						0.002	0.002	0.001
						(0.004)	(0.004)	(0.004)
pop(1880)							0.000	0.000
							(0.000)	(0.000)
foreign(1880)								1.129
								(0.741)
Observations	244	240	240	239	239	239	239	239

Notes: *This table shows how different municipality characteristics are related to "treatment" intensity (i.e., the outcome in this table), which is used in combination with the sudden availability of antitoxin in 1895 as our 2SLS strategy. The explanatory variables are explained in Table 1. All regressions control for county fixed effects and are weighted with the 1890 municipality population size in 1890. Robust standard errors clustered in parenthesis. ***,**,* significant at, respectively, the 10, 5, and 1 percent level.*

Table A.3: Educational outcomes

	(1)	(2)
	attend. rate	attend. rate
VARIABLES	BoE	pub
antitoxin p.c.	-0.039	0.455
	(0.243)	(0.311)
Observations	470	470
Sample	long-diff 1880-1910	long-diff 1880-1910
Mean	90.07	76.81
N	470	470
N_clust	235	235
widstat	18.75	18.75

Notes: *This table reports*