California Center for Population Research
University of California - Los Angeles

Modern Medicine and the 20th Century Decline in Mortality: New Evidence on the Impact of Sulfa Drugs

Seema Jayachandran
Adriana Lleras-Muney
Kimberly V. Smith

CCPR-053-08

December 2008
Latest Revised: December 2008

California Center for Population Research
On-Line Working Paper Series

MODERN MEDICINE AND THE 20TH CENTURY DECLINE IN MORTALITY: NEW EVIDENCE ON THE IMPACT OF SULFA DRUGS

Seema Jayachandran, Stanford University and NBER
Adriana Lleras-Muney, UCLA and NBER
Kimberly V. Smith, Princeton University

December 2008

Abstract

Previous research suggests that medical advances played a negligible role in the large decline in mortality rates during the first half of the twentieth century. This paper, in contrast, presents evidence that sulfa drugs—the first pharmaceuticals effective at treating infectious diseases—were an important cause of U.S. mortality declines after their discovery in the 1930s. Using time-series and difference-in-difference methods (with infectious diseases unaffected by sulfa drugs as a comparison group), we present evidence on the effects of sulfa drugs on mortality. We find that sulfa drugs led to a 25% decline in maternal mortality, a 13% decline in pneumonia and influenza mortality, and a 52% decline in scarlet fever mortality between 1937 and 1943. Sulfa drugs also widened racial disparities in mortality, suggesting that new medical technology diffuses more rapidly among whites than blacks and consistent with the hypothesis that innovation initially increases inequality across population subgroups.

We are grateful to Melissa Thomasson for sharing her maternal mortality data and to James Trussell, Noreen Goldman, and participants at the Office of Population Research seminar at Princeton University, for helpful comments and suggestions.

1. Introduction

The U.S. experienced unprecedented declines in mortality during the 20th century. The age-adjusted death rate declined by 74%, and life expectancy increased by 56%, or 30 years (Bernard Guyer, Mary A. Freedman, Donna M. Strobino and Edward J. Sondik 2000). Seventy percent of the increase in longevity occurred during the first half of the century, driven by improvements in infectious-disease mortality, the leading cause of death at the time. What caused these rapid gains is the subject of great debate and has important policy implications for developing countries today where infectious disease mortality remains high. The prevailing view is that medical advances played a negligible role in the mortality decline between 1900 and 1950. The gains, instead, are attributed to rising living standards, better nutrition, and public health measures that improved water supplies, sanitation systems, and household hygiene (Thomas McKeown 1976, John B. McKinlay and Sonya M. McKinlay 1977, Robert W. Fogel 1994, Samuel Preston 1996, Preston and Michael Haines 1991, David Cutler and Grant Miller 2005).

This view seems uncontroversial for the first thirty years of the 20th century. There were few effective medicines or therapies available to combat infectious diseases before the mid-1930s, so medical care is indeed unlikely to have played an important role during this period. Starting in the mid-1930s, however, there were major medical advances that could have contributed to mortality declines. The effects of these early medical technologies have not received much attention in the literature, due perhaps to data limitations and methodological challenges. Historical epidemiologists, such as McKeown, R.G. Record, and R.D. Turner (1975) and McKinlay and McKinlay (1977), have observed that a large share of 20th-century declines in infectious disease mortality preceded the advent of medical treatments and concluded that the cause of the advances must be something other than modern medicine. This reasoning, however,

does not rule out a role for medical advances. Other researchers have hypothesized about the impact of medical innovation on mortality, but do not empirically test the hypothesis (e.g., Irvine Loudon 1987). The absence of evidence on the effect of medical innovations is very different from evidence of no effect, but it may have led the consensus view to be too dismissive of the role that medical advances played in mortality declines in the 20th century.

In this paper, we attempt to shed new light on this question by examining the introduction of sulfa drugs in the mid-1930s, an important event in the history of modern medicine. Sulfa drugs were the first medicine to effectively treat a range of potentially-fatal bacterial infections, including streptococcal (strep) infections (most notably scarlet fever, puerperal sepsis, and erysipelas), pneumonia, and meningitis. Pneumonia alone accounted for 8% of total deaths in 1936. The production and sales of sulfa drugs increased rapidly after their discovery, with 10 to 15 million people treated with sulfa drugs in 1941 alone (Annabel Hecht 1984). The next major medical advance was not until the mid-1940s when penicillin and other antibiotics became available. Therefore, the period from the mid-1930s to mid-1940s provides a unique opportunity to examine the effect of a new medical technology on mortality.

Despite the historical and medical importance of sulfa drugs, there is little empirical evidence on their contribution to mortality decline. One exception is a study by Melissa A Thomasson and Jaret Treber (2008), who find that hospitalization of childbirth in the U.S. did not have a statistically significant effect on maternal mortality until after sulfa drugs became widely available in 1937. The limited empirical attention given to sulfa drugs may be due in part to the lack of disaggregated data on their diffusion, which makes it more challenging to estimate their effect on mortality. We confront this challenge by using several independent empirical

methods to provide indirect evidence on the effect of sulfa drugs. Collectively, this evidence suggests that sulfa drugs were an important contributor to mortality improvements in the U.S.

The timing of the discovery of sulfa drugs in combination with evidence of their rapid diffusion during a short period—when no other major factors are known to have affected mortality trends—enables us to obtain an estimate of the impact of sulfa drugs on mortality. First, we test for structural trend breaks in the mortality time series. We test both whether the timing of the trend break lines up with the introduction of sulfa drugs and whether the trend break is statistically significant. For diseases treated with sulfa drugs (maternal mortality, pneumonia/influenza, and scarlet fever), we find evidence of trend breaks in the mortality data in 1937, the first year of large-scale sulfa drug production.[1] We find no break during the relevant period for other infectious diseases that were not treatable with sulfa drugs (control diseases).

Second, we estimate the magnitude of the effect of sulfa drugs on mortality using a difference-in-differences regression approach that compares mortality differences between treated and control diseases before and after 1937. Our estimates are based on the identifying assumption that post-1937 mortality declines for treated diseases, beyond those that occurred for the control diseases, are due to sulfa drugs. The results suggest that sulfa drugs were responsible for a 25% decline in maternal mortality, a 13% decline in pneumonia/influenza mortality, and a 52% decline in scarlet fever mortality between the pre-1937 and post-1937 periods. These declines account for between 40% and 75% of total declines for these causes of death during this period.

Third, because sulfa drugs diffused first in cities, we can validate our interpretation of the post-1937 mortality declines as being caused by sulfa drugs by testing whether the declines are

[1] As shorthand, we call maternal mortality a disease, but it is more precisely a cause of death.

more pronounced in urban areas. The declines are indeed stronger in urban areas, providing further evidence that sulfa drugs accelerated improvements in mortality in the post-1937 period.

Next, we compare the effect of sulfa drugs on mortality for blacks and whites. We test the hypothesis that diffusion of new medical technology was lower among blacks during this period, due to the high percentage of southern blacks living in rural areas and the inferior medical care available to blacks under segregation (Douglas Almond, Kenneth Y. Chay, and Michael Greenstone 2006, Loudon 1992). Thus, the mortality reductions from sulfa drugs should be smaller for blacks than for whites. We indeed find greater mortality declines for whites than blacks for each of the treated diseases. This evidence suggests that new medical innovation is initially associated with larger health disparities across population sub-groups.

The rest of the paper is organized as follows. The next section presents a brief history of sulfa drugs. In Section 3, we describe the data. Section 4 presents our tests for structural breaks, and in Section 5 we estimate the magnitude of the effect of sulfa drugs on mortality. In section 6, we investigate racial differences in the effect of sulfa drugs. Finally, Section 7 concludes.

2. History of sulfa drugs

Sulfa drugs were the first effective antibacterial agents to be produced in a pharmaceutical laboratory. Before the first sulfa drug was synthesized in 1932, infectious disease research and treatment was dominated by immunotherapy, which involved the use of either animal serum containing antibodies to treat patients (passive immunization) or vaccines to produce antibodies in individuals to prevent the occurrence of disease (active immunization) (Loudon 2002). While serum was not widely used due to cost and the high risk of serum-related

illness, several important vaccine discoveries were made in the late nineteenth century, including the rabies vaccine (1885) and the diphtheria antitoxin (1891).

In the late 1920s, German researchers began investigating the antibacterial potential of textile dyes (Loudon 1991). In 1932, Gerhard Domagk, a German scientist working at the chemical and dye company I.B. Farben, discovered that a red dye compound, "Prontosil," was successful in treating mice injected with streptococci. Domagk's results, however, were not published until 1935.[2] That same year, the Pasteur Institute in France showed that the active ingredient in the dye compound was sulfonamide. The structure of sulfonamide had been documented in the doctoral dissertation of an Austrian chemist, Paul Gelmo, in 1908 and the patent on it had long expired. Therefore, the technology for making sulfonamide was available, and anyone could produce it. It was also relatively inexpensive to produce. As a result, production and clinical testing of sulfonamide began on a large-scale soon after its discovery.

The first major clinical trial of sulfa drugs occurred in 1936 at Queen Charlotte's Hospital in England, when Prontosil was given to 38 women with serious cases of puerperal fever, a complication from childbirth caused by streptococcal infection that was the leading cause of maternal mortality at the time. The results, published in the June 1936 issue of the *Lancet*, reported a mortality rate of 8% among treated patients versus 24% among the previous 38 untreated patients (Leonard Colebrook and Maeve Kenny 1936a). The success of sulfa drugs in treating puerperal fever was replicated, and even surpassed, in subsequent clinical trials in London, as well as in other parts of Europe (Colebrook and Kenny 1936b, M.A. Foulis and John B. Barr 1937; G. F. Gibberd 1937). Prontosil was first used in the U.S. in 1935 to treat a child with meningitis at Babies Hospital in New York (Andrew J. Schuman 2003). Further U.S. testing

[2] As a result of his discovery, Domagk was awarded the Nobel Prize in Physiology and Medicine in 1939, but due to the Nazi government's opposition to such a prize, was not able to accept the award until 1947 (Kiefer 2001).

in 1936 at Johns Hopkins University and Western Pennsylvania Hospital showed that sulfa drugs were effective against scarlet fever and pneumonia (e.g., Perrin H. Long and Eleanor A. Bliss 1937). Over the next several years, clinical tests provided compelling evidence that various sulfa compounds were effective in treating puerperal fever, pneumonia, scarlet fever, meningitis, gonorrhea, and erysipelas (S.W. Sappington and G.O. Favorite 1939; Paul H. Hamilton 1938).[3]

The lay press created immediate and intense enthusiasm over the drugs. Sulfa drugs became widely-known in the U.S. following a December 1936 *New York Times* article reporting that Prontosil cured President Franklin D. Roosevelt's young son of a deadly streptococcal infection. By 1939, sulfa drugs were commonly referred to as 'magic bullets' and a 'growing miracle' and credited for declines in mortality from childbirth, pneumonia, and other diseases.[4]

Pharmaceutical companies took advantage of the positive press on sulfa drugs. By as early as 1937, sulfa drugs were widely available and dispensed in the U.S. That year, output of sulfa drugs totaled 350,000 pounds; by 1940, it had doubled; and by 1942, production exceeded 14 million pounds (U.S. Tariff Commission, 1936-1950). More than 5000 sulfa compounds were developed by the early 1940s, not all of which proved to be effective (David M. Keifer 2001). Until 1938, sulfa drugs were available without a prescription, though physicians routinely prescribed them, and beginning in 1938, a prescription was required (Barron H. Lerner 1991).[5]

---

[3] It should be noted that for many clinical trials, particularly those involving diseases with high fatality rates, concurrent controls were not used for ethical reasons (Lerner 1991; Loudon 2002). In other cases where disease fatality rates were relatively low, definitive conclusions about the effect of sulfa drugs on mortality could not always be reached. Nonetheless, the successful results of repeated clinical testing were still viewed as compelling evidence of sulfa drugs' effectiveness.

[4] *New York Times*, April 9, 1939; *New York Times*, August 8, 1941, *New York Times*, July 19, 1942; *New York Times*, April 4, 1940.

[5] In 1937, 100 people died from consuming an "Elixir of Sulfanilamide", an untested liquid sulfa preparation. This incident led to the passage of the 1938 Federal Food, Drug, and Cosmetic (FDC) Act that mandated safety testing of drugs before marketing and prohibited the sale of some non-narcotic drugs without a prescription, including sulfa drugs. Sulfapyridine, a sulfa compound, was the first important new drug to be reviewed by the FDC Act (Lesch 2007).

The rapid diffusion of sulfa drugs was aided by the fact that their price was low, about $35 to $100 (2008 dollars) per patient for a full course of the life-saving medicine.[6]

The fact that sulfa drugs were a sensation and became widely available quickly is very useful for analyzing their effects. These characteristics of sulfa drugs' diffusion allow us to look for sudden changes in outcomes (i.e., mortality) right after sulfa drugs were discovered and entered into mass production. This approach would be inappropriate for a medical technology with more gradual adoption, since one could not demarcate a before and after period as sharply.

The subsequent development of penicillin and other antibiotics have largely overshadowed the revolutionary effect of sulfa drugs on medical practice and research. However, the brief 10-year period covering the discovery and widespread use of sulfa drugs was considered 'epochal' at the time, and sulfa drugs are credited today with ushering in a 'therapeutic revolution' (Long and Bliss 1939:1; Richard M. Weinshilboum 1987:1). Sulfa compounds are still used today, but have been largely replaced by antibiotics, which proved to be more effective and less toxic in the treatment of most infectious diseases.[7]

3. Data and disease selection

For our analyses, we use U.S. vital statistics data collected at the national-, state-, and city-level for the period from 1920 to 1950. Data prior to 1930 by state and disease were collected by Grant Miller and made publicly available through the National Bureau of Economic Research.[8] We collected additional national and state-level mortality data from 1930 to 1950 on

[6] The calculation is based on revenue, production volume, and number of patients treated in 1943, as reported by Lesch (2007). As a cross-validation, another price estimate is $4.30 per patient per day (converted to 2008 dollars), suggesting the drugs were taken an average of 8 to 23 days.
[7] Today, sulfa drugs are used to treat some antibiotic-resistant infectious diseases, but are more commonly used in the treatment of urinary tract infections, leprosy, and fungal diseases.
[8] See http://www.nber.org/data/vital-statistics-deaths-historical/.

four causes of death that were shown to be highly responsive to sulfa drugs in clinical trials: maternal mortality, pneumonia, scarlet fever, and meningitis. In our analysis, we refer to these four causes of death as "treated" diseases. The most common cause of maternal mortality, responsible for roughly 40% of maternal deaths in the mid-1930s, was puerperal fever (puerperal sepsis), caused by streptococcus bacteria, which was responsive to sulfonamide; therefore, we use maternal mortality as a proxy for death from puerperal fever (Loudon 1988). In 1939, influenza, which was not affected by sulfa drugs, was combined with pneumonia into a single category in the vital statistics volumes. In order to achieve a consistent data series, we use the combined pneumonia/flu category for all years. For the years when disaggregated data are available, pneumonia accounts for about 75% of total deaths in the combined category.

We selected two diseases to serve as a comparison group, or what we loosely refer to as "control" diseases. An ideal control disease would be an infectious disease that is similar to the infectious diseases treated with sulfa drugs, with the exception that it does not respond to any sulfa compound. Another criterion is that there is a consistent mortality data series available. The two diseases that met these criteria are tuberculosis and diphtheria. We also collected data on three chronic diseases: diabetes, heart disease, and cancer. While chronic diseases are clearly very different from the infectious diseases treated with sulfa drugs, they can provide some insight into whether there were other factors that affected mortality rates for all diseases during the period when sulfa drugs were introduced.

In addition to the state-level series, we look at national series. Age-standardized national mortality series (using the 1940 population as the standard) are available in vital statistics volumes, except for scarlet fever mortality and maternal mortality. For scarlet fever, we aggregated age-specific deaths at the state level and then calculated national age-adjusted

mortality rates using the same direct method of standardization to the 1940 population. The national MMR series is constructed by aggregating total deaths at the state-level and, following convention, is not standardized.

The U.S. death registration area did not include all states until 1933. Therefore, pre-1933 national mortality estimates are approximations to complete national rates and, for our state-level analysis, the number of data points available for each state varies.[9] The results presented are based on unbalanced state panel data, but our results are very similar when we restrict our analysis to states with complete data series. In addition, for the regression sample, we use 1925 as the beginning of our sample period in order to minimize the imbalance.[10] Lastly, data on population size, which were used to calculate state-level death rates, were obtained from the Census Bureau's decennial censuses from 1920 to 1950. Population estimates for years between census years were estimated by linear interpolation by state.

The city-level mortality data were collected by Thomasson and Treber (2008). These data have two limitations for our purposes: (1) they cover only 1928 to 1940 and (2) they contain information on maternal mortality but not other causes of death. During the relevant period, maternal deaths were reported at the city level for each city with a population of at least 10,000 (in states that had entered the death registration system). We focus our analysis on larger cities with populations of at least 25,000 (as of 1930), of which there are 267.

Table 1 reports age-adjusted national mortality rates for 1920 to 1936 (before the introduction of sulfa drugs) and for 1937 to 1950 (after sulfa drugs).[11] In 1920, the maternal

[9] The annual collection of mortality statistics, known as the death registration system, began in 1900. The death registration area expanded from 10 states in 1900 to 36 states in 1920 and the entire United States by 1933. Alaska, Hawaii and the District of Columbia are not included, so there are at most 48 states in a given year.

[10] We begin the state-level panel in 1920 for the trend break analysis since the time-series methods are more reliable with a longer panel.

[11] We use the maternal mortality ratio (maternal deaths per 100,000 live births) rather than maternal mortality rate (maternal deaths per 100,000 women of reproductive age). The maternal mortality ratio is the most commonly used

mortality ratio (MMR) in the U.S. was about 809 deaths per 100,000 live births, much higher than in Western Europe, where the MMR ranged from 235 in Denmark to 665 in France (Loudon 1988). The 1920 MMR level in the U.S. is similar to levels found in developing countries such as Chad and Ethiopia today (Loudon 1988, WHO et al. 2004). The MMR declined by roughly 60% between the pre-sulfa and post-sulfa periods; similarly large declines also occurred in Western Europe. The pneumonia/influenza and tuberculosis disease categories were responsible for the largest share of infectious disease deaths between 1920 and 1950, and accounted for approximately 22% of overall mortality. Deaths due to scarlet fever, meningitis, and diphtheria were much less common during this period. However, mortality for all of these infectious diseases dropped by 50% to 85% between the pre- and post-sulfa period. It should be noted that the 30-year period from 1920 to 1950 covers the discovery of sulfa drugs, as well as the introduction of penicillin and other antibiotics starting in the mid-1940s. For this reason, when estimating the size of the effect of sulfa drugs on mortality, we end our data series in 1943 to avoid possible confounding by the introduction of other medical advances in the mid-1940s. Table 1 also shows summary statistics across states for the 1925-36 pre-sulfa period and 1937-43 post-sulfa period used in our regression estimates. Note that the state means are not an exact match to the national data because the national data series are age-adjusted and the state averages are not population-weighted.

---

measure of maternal mortality and is a more precise measure of risk in that it represents the risk of death once a woman has become pregnant. As noted earlier, the national maternal mortality series are not age-standardized.

4. The effect of sulfa drugs on mortality in the U.S.: Times series evidence

We use indirect methods to measure the effect of sulfa drugs on mortality because disaggregated data on the diffusion of sulfa drugs are not available. Ideally, to estimate the impact of sulfa drugs on mortality, we could estimate the following equation:

$$(4.1) \qquad Log(M)_{it} = \beta_0 + \beta_1 (Sulfa\ drug\ availability)_{it} + e_{it}$$

The dependent variable is the natural log of the mortality rate in state *i* and year *t*. The reason for modeling the log of the mortality rate is that medical advances are hypothesized to have a proportional change on mortality rates: if the initial rate of mortality is high, there is more room for mortality decline, so a medical advance should lead to a larger drop in the level of mortality than if the initial rate is low. Using proportional changes also facilitates comparisons of mortality change across diseases that have different levels of mortality initially.

The ideal independent variable in model (4.1) would be a measure of sulfa drug availability (supply) that is uncorrelated with the demand for sulfa drugs. The hypothesis is that when sulfa drugs become more available, mortality falls ($\beta_1 < 0$). In practice, researchers might use state-year level data on sales volume or number of prescriptions issued as a proxy for sulfa drug availability, but it is worth keeping in mind that such a regression would suffer from endogeneity bias. For example, sales of sulfa drugs are likely to be higher in places with higher mortality risk from diseases treatable with the drugs. There is also a data limitation in our case: there are no data on the availability of sulfa drugs disaggregated by state or other subgroup for the period examined.

Instead, we use two facts about sulfa drugs to arrive at empirical tests of their effects. First, we use the fact that the timing of the sulfa drug discovery can be taken as exogenous and

that, because sulfa drugs could not be patented, diffusion was rapid: the first year of large-scale production and sales of sulfa drugs in the U.S. was 1937 (U.S. Tariff Commission, 1936-1950). This allows us to use time series techniques to test whether the timing of mortality declines corresponds with sulfa drugs' introduction in 1937. Second, there is strong evidence from clinical trials that sulfa drugs were effective against some infectious diseases but not others. Therefore, we are able to use infectious diseases that were impervious to sulfa compounds as a comparison group in our analysis.

Graphical evidence on trend breaks in the mortality time series

Figures 1 to 4 depict trends in overall mortality and in mortality for the treated infectious diseases, control infectious diseases, and chronic diseases.[12] (Treated diseases are those which sulfa drugs were effective against, and control diseases are those which sulfa drugs were ineffective against.) All graphs show mortality on a log scale. The vertical line at 1937 indicates the first year of large-scale production of sulfa drugs in the U.S.

Figure 1 depicts total mortality, and Figure 2 depicts mortality for the four treated diseases: maternal mortality, pneumonia, scarlet fever, and meningitis. With the exception of meningitis, the mortality curves become notably steeper after 1936-7 compared to their pre-1936 trends. Meningitis has a very different pattern than the other diseases, with wide fluctuations that suggest outbreaks. Modeling this type of time-series process requires very different modeling techniques than those we use for the rest of the diseases in this study and, for this reason, we do not pursue meningitis further in our analysis. Also worth noting is that the mortality rate from

[12] Graphs are based on national-level age-adjusted death rates, using 1940 as the standard, except in the case of MMR, which is based on aggregated state-level data.

pneumonia/influenza is very noisy, particularly in the pre-sulfa-drug period, which will most likely result in noisier estimates in our subsequent analysis.

In contrast to the treated diseases, mortality trends for the control and chronic diseases (Figures 3 and 4) do not show any noteworthy changes around the time that sulfa drugs were introduced. Death rates for the control diseases are trending downward and chronic diseases trending upward, with no obvious breaks in the trend lines. When penicillin was introduced around 1945, there is a steepening of the tuberculosis and diphtheria curves; penicillin (but not sulfa drugs) was an effective treatment against these diseases.

The graphs also show that between 1935 and 1937, there is a slight increase in overall mortality as well as mortality from several specific causes of death considered here. The cause of this uptick is a puzzle in the demographic literature and one that we do not attempt to solve in this paper. Its timing is somewhat unfortunate for our purposes since we are interested in structural breaks in our mortality series around 1937; however, this uptick exists in the mortality series for both treated and control diseases (and, to a lesser extent, for chronic diseases). To ensure that our results are not driven by this anomaly, we also check the robustness of our results to the omission of these years in our estimation.[13]

Formal tests of trend breaks in mortality time series

The graphs provide suggestive evidence of an increase in mortality decline associated with the introduction of sulfa drugs in 1937. We next conduct a more formal test for structural breaks in the trend lines. The macroeconomics literature contains several methods for identifying structural breaks in time series data. We use a method that does not require a priori assumptions

[13] For all diseases except pneumonia/influenza, the data suggest that the mortality uptick is limited to 1935 to 1936. For pneumonia/influenza, the uptick is between 1935 and 1937. To be conservative, we present results from models that exclude 1935 to 1937 as our main robustness test. Appendix Table A2 presents results from an alternative test that excludes only 1935 and 1936. In both cases, dropping the years of the uptick has little impact on our results.

about the precise timing of the break.[14] As noted earlier, national data on sulfa drug production indicate that 1937 was the first year of large-scale production and sales, so our hypothesis is that if sulfa drugs affected mortality, the strongest trend break should be detected around 1937.

The method that we use to test for structural breaks is the Quandt likelihood ratio (QLR) test (Richard E. Quandt 1960). The QLR test has been shown to be a reliable test for structural breaks in the case of unknown break points and has been applied recently to program evaluation (Donald W.K. Andrews 1993, Anne M. Piehl, Suzanne J. Cooper, Anthony A. Braga, and David M. Kennedy 2003, James H. Stock and Mark W. Watson 2003). The approach can be described using the following basic model:

$$(4.2) \qquad Log(M)_t = \beta_0 + \beta_1 Year_t + \delta_0 D_t(\tau) + \delta_1 D_t(\tau) * Year_t + e_t$$

where the dependent variable, $Log(M)_t$, is the log of mortality in year t, $Year_t$ is a continuous year variable, and D_t is a break indicator variable. The possible break date is denoted τ; D_t is equal to zero for all years before τ and equal to one for all subsequent years. The equation also includes an interaction between the break indicator, D_t, and $Year_t$, allowing for not just a change in the intercept (mean) at time τ, but a change in the slope.

To conduct the QLR test, one first needs to specify an interval over which to test for possible break dates. The interval must allow for a sufficient number of data points on each end in order to estimate the regression line before and after the break point. Equation (4.1) is then estimated for each possible break date, τ, in the specified test window. After each estimation, the hypothesis of a possible break in the mean and trend is tested by using the F-statistic that tests

[14] Interrupted time series models are commonly used in the program evaluation literature. However, the conventional practice is to assume a particular date that the program would have an impact and to test for a discrete change in the intercept or slope of the regression line at that point using conventional *F*-tests.

the null hypothesis that $\delta_0 = \delta_1 = 0$.[15] The largest of the resulting F-statistics is used to identify the best possible break point and to determine the significance of the break using Andrews critical values (Andrews 2003).[16]

The QLR test, however, is only valid in the case of stationary data series. Our mortality series are not stationary.[17] To achieve stationarity, we use the first difference of log mortality, $\Delta Log(M)_t$, or $Log(M)_t - Log(M)_{t-1}$, as the dependent variable and estimate the following equation:

$$(4.2') \qquad \Delta\, Log(M)_t = \beta_0 + \beta_1 Year_t + \delta_0 D_t(\tau) + \delta_1 D_t(\tau) * Year_t + e_t.$$

This model allows for a break in both the level and slope of the first-differenced mortality series. Therefore, it tests for structural change in the slope of the mortality series and allows for a change in the curvature of the mortality series. As a result of using first differences, however, we cannot detect a change in the level of mortality. In other words, if sulfa drugs caused a drop in the level of mortality in a particular year but no change in the slope, this test would not detect the change. Therefore, we may be missing one element of the effect of sulfa drugs on mortality, and our test of the effects of sulfa drugs is conservative. Also note that to account for serial correlation, we compute Newey-West standard errors and allow the error structure to be correlated up to two lags.

We conduct the QLR test using national- and state-level data. The range of dates over which we test for possible breaks is based on two considerations: (1) having an adequate number

[15] This is often referred to as the Chow Test. The QLR test is, in essence, a sequential Chow Test, where the largest F-test from these sequential tests is used to identify a possible break point.

[16] Andrews (2003) critical values account for the fact that many possible break points are being tested and, therefore, there are many chances to reject the null hypothesis of no structural change. The critical values depend on the number of parameters that are being allowed to break and the fraction of the total date range that is being examined for possible breaks points (Stock & Watson, 2003).

[17] Stationary series fluctuate around a fixed mean level over time, which makes it easier to detect a true break in the series. The graphs strongly suggest the data are not stationary, and if we test for stationarity using an Augmented Dickey-Fuller test, we find that the mortality series are not stationary but that the first-differenced series are.

of data points before (after) the earliest (latest) break point in the range and (2) allowing for the maximum number of states to be included in the analysis (given varying dates of entry into the vital statistics registration system across states). The window that we use to identify possible break dates is 1933 to 1943.[18,19] We interpret finding a trend break between 1936 and 1938 as evidence that the introduction of sulfa drugs had an impact on mortality.

Table 2 presents the results of the QLR test based on national-level mortality data. The results confirm our visual inspection. For all-cause mortality, as well as mortality from treated diseases, the test identifies a structural break between 1936 and 1938.[20] The four treated diseases account for about 12% of total mortality in the pre-period and were not the only diseases treated with sulfa drugs, so it is not surprising that we detect a possible break in total mortality when sulfa drugs became widely available. In spite of our relatively low statistical power to detect a break point due to our short data series, the break is statistically significant at the 5% level in the case of scarlet fever. The only comparison disease that has a possible trend break around 1937 is heart disease, though the break is not statistically significant. This slowdown in the increase in heart-disease mortality, which can be seen in Figure 4, could be due to the fact that sulfa drugs reduced pneumonia, and one of the complications from pneumonia is heart-disease mortality.[21]

As a specification test, we conduct the QLR test using different test windows to see if our results are sensitive to the range of years selected for possible break points. We also test for

[18] There is little guidance in the literature on how much of the sample to "trim" on each side to create an appropriate test window. A common approach is to trim 15% from both ends. However, in the case of small samples such as ours, that would leave only 3 data points on either side of the test window. In addition, there is no evidence that the QLR test requires symmetrical trimming. Andrews (1993) provides a calculation and critical values that enable one to obtain critical values for all possible test intervals (updated critical values are found in Andrews (2003)).

[19] In our model specification, we allow for only one break date. The QLR test can be easily adapted to allow for multiple breaks. However, the identification of multiple breaks requires a longer time series than is used here. Therefore, our test identifies the strongest possible break, but does not rule out the existence of other break points.

[20] The later break date for pneumonia/influenza mortality likely reflects the availability of a new sulfa compound more effective in treating pneumonia, sulfapyridine, beginning in 1938.

[21] This suggests that perhaps some of the diseases we study are not independent. We nevertheless treat them as such in all of the analyses in this paper, since it is extremely difficult to account for dependence without making strong functional form assumptions.

trends breaks in the level (rather than the log) of mortality. These results are presented in appendix Table A1. We find that the results are robust to the window starting in 1925 through 1930, to end dates through 1945, and to using levels rather than logs.[22]

We next conduct the same structural break test using state-level mortality data. Graphical analysis of state mortality trends (not presented) suggests wide variation in the timing and magnitude of mortality change when sulfa drugs were introduced. This is unsurprising, as we would expect certain states to adopt new medical technology earlier and more widely than other states. The results of the QLR test for structural breaks in the state-level data series, presented in Table 2, confirm this variability. In approximately 30% of states, the test identifies possible breaks between 1936 and 1938 for maternal mortality and scarlet fever. For pneumonia/influenza mortality, the test identifies possible breaks between 1936 and 1938 in 67% of states.[23] In the case of our control diseases, tuberculosis and diphtheria, in only 13% of states were possible break dates detected between 1936 and 1938. Note that since the trend break tests require continuous time series, if a disease has zero mortality (undefined log mortality) in any year for a state (which affects diphtheria and scarlet fever), we drop the state-disease for this analysis.[24] In short, the state-level results support our previous findings. They also suggest important differences across states in the adoption and diffusion of sulfa drugs. In unreported results, we

[22] The results are virtually the same when using mortality levels rather than log mortality, with two exceptions (that strengthen our findings): the break in the MMR achieves statistical significance when using levels, and the break date for pneumonia/flu changes from 1938 to 1937. In addition, when using levels, the estimated break date changes from 1937 to 1945 for pneumonia/flu when the test window is expanded to include 1945, the year that penicillin was introduced.

[23] The test identifies a slightly larger percentage of states with possible breaks when using mortality levels rather than logs. More specifically, the test identifies possible breaks in 41% of states for maternal and scarlet fever mortality and 78% of states for pneumonia/flu mortality.

[24] South Dakota and Texas entered the death registration area in 1930 and 1933 and, therefore, are excluded from this analysis too.

were unable to identify any state level factors that predicted the differential patterns across states, however.[25]

5. Difference-in-differences estimates of the effect of sulfa drugs on mortality

Next, we use regression analysis to estimate the impact of sulfa drugs on mortality. Regression analysis allows us, first, to more formally test whether the trend breaks are larger for the treatment diseases than for the control diseases and, second, to measure the magnitude of the effect of sulfa drugs on mortality.

Based on the trend-break results above, our starting point is the assumption that the 1937 trend break for treated diseases, relative to control diseases, represents the effect of the introduction of sulfa drugs. To account for the possibility that other factors may have affected mortality around 1937, we use a difference-in-difference approach, where we compare pre- and post-1937 levels and trends in the treated diseases to pre- and post-1937 levels and trends in a control disease, tuberculosis. Diphtheria is excluded because a substantial percentage of state-year observations have zero mortality and therefore undefined log mortality.) We also exclude chronic diseases in further analyses since they are less comparable to the treated diseases.

The identifying assumption is that the trends in control and treated diseases are similar, and that sulfa drugs are the only factor differentially affecting the treated diseases in 1937. We use state-level mortality data to estimate the following equations:

[25] We investigated whether any state-level factors can explain these observed differences in the adoption of sulfa drugs across states, where adoption was proxied by the level of mortality decline around 1937. Using our difference-in-differences setup, we tested for an association between state characteristics—such as income per capita, high school graduation rates, number of doctors and hospitals per capita, and percent of the population in urban areas—and mortality declines due to sulfa drugs. We found some suggestive evidence that adoption was more extensive in states with higher levels of education and urbanization, but our results were generally inconclusive (in most cases, not statistically significant and sensitive to specification choice). Results are available upon request.

$$(5.1)\ Log(M)_{idt} = \beta_0 + \beta_1\, Treated_d{*}Post\text{-}1937_t + \beta_2\, Treated_d{*}Year_t + \beta_3\, Treated_d + \beta_4\, Year_t + \gamma_{it} + \varepsilon_{idt}$$

and

$$(5.2)\ Log(M)_{idt} = \beta_0 + \beta_1\, Treated_d{*}Year_t{*}Post\text{-}1937_t + \beta_2\, Treated_d{*}Post\text{-}1937_t + \beta_3\, Treated_d{*}Year_t + \beta_4\, Treated_d + \beta_5 Year_t + \gamma_{it} + \mu_{it}{*}Year_t + \varepsilon_{idt}.$$

The dependent variable is the log of the mortality rate in state *i* for disease *d* and year *t*. *Treated* is an indicator variable for whether the disease is a treated disease, *Post-1937* is a dummy variable equal to zero in the period from 1925 to 1936 and equal to one in the period from 1937 to 1943, and $Year_t$ is a continuous year variable (centered on 1937). Standard errors are clustered by state to correct for possible serial correlation (Marianne Bertrand, Esther Duflo, and Sendhil Mullainathan 2004).

The first model (equation 5.1) examines changes in the level of the dependent variable after 1937. The coefficient of interest, β_1, measures whether the post-1937 reduction in mortality was larger for treated diseases than for the control disease. (Whereas in the trend break analysis we were required to use the first difference of mortality and therefore could not examine whether there was a discrete drop in mortality in 1937, in the regression analysis – where stationarity is not a required condition – we can.) The model also allows for a different linear time trend for treated and control diseases. (The results are similar excluding the differential time trend.) Finally, the regression includes *State*Post-1937* fixed effects, denoted by γ_{it}, which, control for the main effect of *Post-1937* and absorb state variation in mortality declines (which is orthogonal to the regressor of interest, but adds noise to the estimates).

In the second model (5.2), we allow for a change in both the intercept and the slope after 1937. In this model, the statistical question of interest is whether β_1 and β_3 are jointly

significantly different from zero. The equation includes *State*Post-1937* fixed effects and their interaction with $Year_t$ (the terms γ_{it} and μ_{it}**Year*$_t$). These variables control for the main effect of *Post-1937* and *Post-1937*Year*, plus they absorb residual variation from state-specific mortality trends. The model estimated in the first equation (5.1) allows for a level decline only (and not a change in slope), but it is more parsimonious, which might be preferable given the short time series, which ends in 1943.

We estimate the equations separately for each treated disease using state-level mortality data from 1925 to 1943.[26] The results are presented in Table 3. Panel A presents the results of our main specification. The coefficients of interest are negative and significant in both models for all three treated diseases, suggesting that the introduction of sulfa drugs led to significant mortality declines. Panel B shows that these results are robust to the exclusion of 1935 to 1937, the years of the uptick in mortality. While the estimated coefficients in the MMR and scarlet fever models are virtually unchanged in this second specification, the coefficients in the pneumonia/flu model are reduced by roughly half. This is not surprising given that the data for this disease are noisier and we have a short panel, and thus point estimates are sensitive to excluding a few points. But in all cases the estimates are negative and statistically significant.

To assess the size of the effect of sulfa drugs on mortality, we use the coefficients from Panel A, column 1, corresponding to model (5.1), which allows for only a mean change in mortality, but is less demanding of our data. In the MMR model, the coefficient of -0.28 implies that sulfa drugs resulted in a 25% decrease in MMR in the post period, or 163 (.25*653) fewer maternal deaths per 100,000 births between 1925-36 and 1937-43. This accounts for 56% of the

[26] As noted previously, years after 1943 are excluded to avoid possible confounding of our estimates from the introduction of penicillin in 1945-1946. Note that one could estimate related models using the national-level rather than the state-level data. Using state data has the advantage that there are over forty separate "experiments" on the effect of sulfa drugs, so the statistical tests have more power.

total decline in MMR during this time period, which fell by about 292 deaths per 100,000 births between the two periods (see Table 1 for summary statistics). The estimates also suggest that sulfa drugs were responsible for a 13% decline in pneumonia/flu mortality (a decrease of 15 deaths per 100,000 people) and 39% of total mortality decline in this disease category. (If one assumes the averted deaths in the combined pneumonia/influenza category were from pneumonia, then sulfa drugs caused a 17% decline in pneumonia mortality.) Finally, the scarlet fever results imply that sulfa drugs led to a 52% drop in scarlet fever mortality, which accounts for 76% of total mortality decline for this disease between the pre-1937 and post-1937 periods. These estimates are quite consistent with clinical trial estimates from the time (G.M. Evans and Wilfrid F. Gaisford 1938, Colebrook and Kenny 1936a).[27]

Using these estimated effect sizes, in combination with the summary statistics presented in Table 1, we can calculate the contribution of sulfa drugs to the overall mortality decline from 1937 to 1943. We first convert the decline in MMR due to sulfa drugs to a decline in the number of maternal deaths per 100,000 individuals, which we estimate to be 24%. We know that mortality from maternal causes, pneumonia/influenza, and scarlet fever accounted for 1.1 %, 10.5%, and 0.2% (respectively) of total deaths in the pre-sulfa period. Therefore, we estimate that sulfa drugs', through their effect on the three treated diseases, were responsible for 2 % of overall mortality decline between 1925-36 and 1937-43 [(.24*.011)+(.13*.105)+(.52*.002)]. This contribution is relatively small, which is not too surprising given the small share of total deaths accounted for by maternal complications and scarlet fever; the three treated diseases in

[27] In clinical trials, sulfa drugs reduced mortality from puerperal fever (responsible for roughly 40% of maternal deaths in the pre-sulfa period) by about 81%, which implies a 32% reduction in maternal mortality, and reduced mortality from pneumonia by 50 to 70%, which implies a 37 to 52% reduction in pneumonia/influenza given that pneumonia comprised about 74% of the pneumonia/influenza category in the pre-sulfa period. As expected, we find that the decline in the population, as measured by our regression coefficients, is smaller than in the trials. Only a portion of the population with the disease took sulfa drugs, and drugs are generally more efficacious in a controlled, clinical setting. We do not have estimates from scarlet fever clinical trials.

our analysis account for 11.7% of mortality in the pre-period. Our estimates, however, do not attempt to capture the total impact of sulfa drugs on mortality because we cannot account for the effects of sulfa drugs on mortality declines after 1943 and because, due to data limitations, we cannot study the effect on other diseases (such as meningitis, erysipelas, and cerebrospinal fever) that sulfa drugs also treated.[28] It is also worth noting that sulfa drugs made a larger contribution to life expectancy gains during the period, since the diseases they treated struck people at a relatively young age.[29]

Urban-Rural Evidence

While the above analysis provides strong suggestive evidence that the introduction of sulfa drugs caused a reduction in mortality starting in 1937, we conduct an additional test to validate our interpretation of the decline as resulting from sulfa drugs. Studies of diffusion of medical technology in the U.S. find that large, urban-based research hospitals are usually the first to adopt new medical technologies, followed, after the importance of the new technology is proven, by urban medical facilities and, lastly, by medical centers in rural areas, often after a considerable delay (Duke University, Medical Technology Assessment Working Group, 2006). In addition, in the case of sulfa drugs, historical accounts relate that clinical testing was first carried out at several urban-based research facilities in the U.S. such as Johns Hopkins University (Baltimore), Babies Hospital (New York City), and Western Pennsylvania Hospital (Pittsburgh), among others. Therefore, one validation test is to compare post-1937 mortality

[28] This paper focuses on the effect of sulfa drugs on mortality, but sulfa drugs were also shown to be effective in the treatment of various infectious diseases and conditions with low fatality rates, including bacillary dysentery and gonorrhea. There is also evidence from clinical trials that they reduced morbidity from pneumonia and our other treated diseases. Therefore, the contribution of sulfa drugs to health is greater than what is captured by focusing on mortality alone.

[29] In 1936, people under the age of 25 comprised 55.2% of maternal deaths, 29.6% of pneumonia deaths, and 82.0% of scarlet fever deaths. In comparison, 17.5% of all deaths were among those under age 25.

reductions in urban and rural areas, with the expectation that urban areas would benefit from sulfa drugs more immediately than rural areas.

We have urban-only mortality data but not rural-only data. Therefore, we compare urban mortality to state-level mortality, where state-level data represent the sum of mortality in urban and rural areas. If the trend breaks are larger in urban areas than in rural areas, then the urban data should show larger trend breaks than the state-level data that aggregate across urban and rural populations.

We restrict the urban data to cities with populations of at least 25,000 for two reasons. First, our hypothesis of a higher rate of diffusion in urban areas is more likely to hold in larger cities. Second, vital statistics data were not collected in cities smaller than 25,000 people in 1931 and 1932. Recall that our city-level data series includes maternal mortality only, and only covers the period from 1928 to 1940. Figure 5 shows that the MMR was higher in cities than at the state level until 1938, after which point an urban advantage emerges. While the mortality curves for both cities and states become noticeably steeper around 1936-7, the graph suggests that this slope change occurs earlier in cities. Because the city-level series end in 1940, however, we cannot estimate the timing of possible trend breaks using the time series methods used above, which require a longer panel.

We can, however, estimate the regression models. We impose a break in 1937 and test whether this break is larger in cities, using pooled city- and state-level maternal mortality data. We estimate the following two regressions:

$$(5.3)\quad Log(MMR)_{ict} = \beta_0 + \beta_1\, Urban_c * Post\text{-}1937_t + \beta_2\, Urban_c + \beta_3\, Urban_c * Year_t + \beta_4 Year_t + \gamma_{it} + \varepsilon_{ict}$$

and

$$(5.4)\ \ Log(MMR)_{it} = \beta_0 + \beta_1\, Urban_c * Post\text{-}1937_t + \beta_2\, Urban_c * Year_t + \beta_3\, Urban_c * Year_t * Post\text{-}1937_t + \beta_4\, Urban_c + \beta_5 Year_t + \gamma_{it} + \mu_{it} * Year_t + \varepsilon_{ict},$$

where the dependent variable is the natural log of the maternal mortality ratio for category *c* (city- or state-level observation) in state *i* in year *t*, *Urban* is an indicator variable for whether the observation is a city, *Post-1937* is a dummy variable equal to zero in the period from 1928 to 1936 and equal to one in the period from 1937 to 1940, and $Year_t$ is a continuous year variable. Equation (5.3) includes *State*Post-1937* fixed effects (denoted γ_{it}), and allows for differential time trends in urban areas. Equation (5.4) also includes *State* Post-1937* fixed effects interacted with $Year_t$. These equations are analogous to those used before, with the treated-disease dummy replaced with an urban dummy. The hypothesis is that the interactions with *Urban* and *Post-1937* (the coefficients β_1 in Equation (5.3) and both β_1 and β_3 in Equation (5.4)) should be negative. Since the error terms will be correlated between state and city data (given that state is the aggregate of urban and rural), we cluster the standard errors by state.

Table 4 presents the regression results. For both equations, β_1 is negative and significant, suggesting that cities experienced statistically significantly larger declines in maternal mortality after 1937 compared to rural areas. In addition, in the second model, the slope of the MMR curve after 1937 is more negative in urban areas (β_3 is negative and significant), suggesting that mortality started to decline faster in cities compared to rural areas after sulfa drugs were introduced. Panel B again shows that, although the magnitude of these effects changes when the 1935-1937 years are excluded, the substantive conclusions of the analysis do not.

6. Racial differences in the effect of sulfa drugs

We next compare the effect of sulfa drugs for whites and blacks. A growing literature posits that new medical technologies may increase inequalities in health because they tend to favor, at least initially, the better off, those who are more able to access and implement health-related innovations (Sherry Glied and Adriana Lleras-Muney 2008, Dana P. Goldman and Darius N. Lakdawalla 2005). More broadly speaking, this idea is the basis of the "fundamental causes" hypothesis, which posits that socioeconomic (SES) gradients in health arise because those with higher SES use their greater resources to improve their health (Bruce G. Link, Mary E. Northridge, Jo C. Phelan and Michael L. Ganz 1998).

We test this hypothesis for the case of sulfa drugs. There are several reasons to believe that sulfa drugs may have diffused more rapidly within the white population, leading to steeper declines in mortality for whites than blacks after sulfa drugs' introduction in 1937. During this period, the vast majority of blacks lived in lower-income Southern states and in predominantly rural areas, often at far distances from hospitals and physicians. Hospital segregation policies and greater resources in white facilities further contributed to lower health care utilization among blacks (Almond et al., 2006).

In the case of maternal mortality, the rate of home deliveries by untrained midwives was still high among black women (and very low among white women) living in the South during this period, which may have reduced black mothers' access to sulfa drugs during and after delivery (Loudon 1992, Almond et al. 2006). Thomasson and Treber (2008) find that in U.S. cities, greater rates of hospital births were associated with significant and large declines in MMR for both races after the introduction of sulfa drugs in 1937, but that the effect of hospitals on

MMR in the post-sulfa period was greater for blacks than whites. This may reflect the relatively poorer hospital care received by black mothers in the pre-sulfa drug period.[30]

Figure 6 shows that there were clear racial disparities in maternal and pneumonia/flu mortality during the 1920 to 1950 time period.[31] The gap between blacks and whites in MMR was stable from 1925 to 1935, but it widened after 1936 as a result of steeper declines in MMR among whites. Pneumonia/flu appears to follow the same pattern, though the evidence of greater gaps after 1936 is not as clear (although, again, these data are noisy). Somewhat surprisingly, mortality from scarlet fever was greater among whites than blacks, which the previous literature has attributed to lower susceptibility to scarlet fever among blacks (William M. Welch and Jay F. Schamberg 1905). Nevertheless after 1936, mortality from scarlet fever declines more rapidly for whites, and the black advantage is greatly diminished.

We test whether these patterns are statistically significant by testing for trend breaks in the time series for blacks and whites, using the same methodology employed in Section 4. As shown in Table 5, the QLR test using national data identifies a statistically significant break in 1937 in maternal and scarlet fever mortality for whites only. The test also identifies a break in pneumonia/influenza mortality in 1937-8 for both races, but the break is not statistically significant. For the state analysis, only the 18 states with black populations greater than 5% of the total population are included; most of these states are in the South.[32] Recall that the time series methods require us to drop state-diseases with zero mortality in any year, which results in a reduced sample size for scarlet fever and diphtheria. Thus, the state-level trend break tests

[30] In their balanced city panel, Thomasson and Treber (2008) find that greater hospitalization of births in the pre-1937 period lowered MMR for whites only, but the effect was larger in the post-1937 period. However, in their unbalanced panel, the mortality effects of hospital-based deliveries in the pre-sulfa drug period are insignificant for both blacks and whites.

[31] The graphs in Figure 6 are based on national-level age- and race-adjusted death rates using 1940 as the standard, except in the case of MMR, which is based on aggregated state-level data.

[32] These states are Alabama, Arkansas, Delaware, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, Missouri, New Jersey, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

should be interpreted with some caution. That said, we find that the trend break in the treated diseases, except scarlet fever, occurs when sulfa drugs were introduced more often for whites than blacks, suggestive of more widespread diffusion of sulfa drugs among whites.[33]

Next, we examine whether post-1937 declines in mortality from our treated diseases were larger in magnitude for whites compared to blacks. Using the same difference-in-differences regression approach outlined in Section 5, we estimate equations 5.1 and 5.2 separately for blacks and whites. Table 6 reports the results. These results suggest that sulfa drugs led to a 26% decline in MMR for whites versus a 12% decline for blacks. In the case of pneumonia/flu, our estimates suggest that sulfa drugs led to a 21% decline in mortality for whites and an 11% decline for blacks. For scarlet fever, we find that sulfa drugs had no statistically significant effect on scarlet fever mortality for blacks, but resulted in a 55% decline in scarlet fever mortality for whites.

To determine whether the differences in the coefficients for blacks and whites are statistically significantly, we pool the race data and estimate fully interacted versions of the models. The results are shown in Panel C of Table 6 and confirm that whites experienced significantly greater mortality reductions from sulfa drugs than blacks for all of the treated diseases (a positive interaction coefficient implies a smaller decline for blacks than whites given that the effects of *Treated*Post-1937* and *Treated*Year*Post-1937* are negative). Appendix Table A3 shows that these results are robust to omitting 1935 to 1937 in our estimations. Thus, whites appear to have benefited more (in the short term) from the introduction of sulfa drugs, in line with previous results in the literature that suggest that innovation increases health gradients, at least initially. Sulfa drugs are a particularly interesting case because, unlike most life-saving

[33] The percent of states with trend breaks in 1936 to 1938 are not comparable to the all-race results in Table 2 because the samples are different. For closer comparability to Table 2, Table 5 also shows the whites-only data for all states (again dropping cases with zero mortality).

innovations, they were inexpensive. (The full course of medicine cost less than $100 (2008 dollars), as discussed in section 2.) This suggests that factors other than income were the major barrier to accessing medical breakthroughs for low-SES groups like blacks.

7. Conclusion

Medicine is regarded as playing a relatively minor role in mortality decline through much of the 20th century. While several scholars have made compelling arguments in support of this view (e.g., McKeown 1976, Fogel 1994, McKinlay and McKinlay 1977), there have been few attempts to statistically test the effect of specific medical measures on mortality decline. This paper addresses this shortcoming in the literature by using several statistical methods to produce an estimate of the impact of sulfa drugs—the first major pharmaceutical innovation to treat disease in the 20th century—on mortality in the U.S.

We use time series and difference-in-difference estimation methods—whereby we compare mortality patterns before and after sulfa drugs were introduced between treated and control diseases—to arrive at an estimate of the impact of sulfa drugs. Our results suggest that sulfa drugs caused significant mortality declines for the three causes of death included in our estimations: maternal mortality, influenza and pneumonia, and scarlet fever. The introduction of sulfa drugs resulted in a 25% drop in maternal mortality, a 13% drop in pneumonia and influenza mortality, and a 52% drop in scarlet fever mortality between 1925-36 and 1937-43. In addition, sulfa drugs significantly increased the pace of mortality decline for these diseases. We also find that these reductions were larger in cities than in rural areas.

Our findings also suggest that sulfa drugs initially increased racial inequalities in maternal and pneumonia/flu mortality by causing a larger decline in mortality among whites than

blacks between the pre- and post-sulfa periods. These results suggest that sulfa drugs diffused more rapidly among whites, and more generally support the hypothesis that medical innovation results in greater inequality across groups, at least initially. This finding is especially striking given that sulfa drugs were an inexpensive life-saving drug.

8. References

Almond, Douglas, KennethY. Chay, and Michael Greenstone. 2006. "Civil Rights, the War on Poverty, and Black-White Convergence in Infant Mortality in the Rural South and Mississippi." MIT Department of Economics Working Paper 07-04.

Andrews, Donald W.K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61(4): 821-856.

____. 2003. "Tests for Parameter Instability and Structural Change with Unknown Change Point: A Corrigendum." *Econometrica* 71(1): 395-397.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119(1): 249-275.

Colebrook Leonard and Maeve Kenny. 1936a. "Treatment of Human Puerperal Infections, and of Experimental Infections in Mice with Prontosil." *Lancet* I:1279-1286.

____.1936b. "Treatment with Prontosil of Puerperal Infections due to Haemolytic Streptococci." Lancet 2:1319-1322

Cutler, David and Grant Miller. 2005. "The Role of Public Health Improvements in Health Advances: The Twentieth-Century United States." *Demography* 42.1:1-22.

Duke University, Medical Technology Assessment Working Group. 2006. *Assessing the Impact of Medical Technology Innovations on Human Capital. Phase I Final Report (Part A): State-of-the-Science Literature Reviews*. Durham: Center for Demographic Studies, Duke University.

Evans, G.M., and Wilfrid.F. Gaisford. 1938. "Treatment of Pneumonia with 2-(p aminobenzenesulphonamide) pyridine." *Lancet* 2: 14-18

Fogel, Robert W. 1994. "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy." *American Economic Review* 84(3): 369-395.

Foulis M.A. and John B. Barr. 1937. "Prontosil Album in Puerperal Sepsis." *British Medical Journal* 1:445-6.

Gibberd G.F. 1937. "Prontosil and Similar Compound in the Treatment of Puerperal Haemolytic Streptococcus Infections." *British Medical Journal* 2:695-698.

Glied, Sherry and Adriana Lleras-Muney. 2008. "Technological Innovation and Inequality in Health." *Demography* 45(3): 741-761.

Goldman, Dana P. and Darius Noshir Lakdawalla. 2005. "A Theory of Health Disparities and Medical Technology." *Contributions to Economic Analysis & Policy* 4(1)8: 1-30.

Guyer, Bernard, Mary Anne Freedman, Donna M. Strobino, and Edward J. Sondik. 2000. "Annual Summary of Vital Statistics; Trends in the Health of Americans During the 20th Century." *Pediatrics* 106:1307-1317.

Hamilton, Paul M. 1938. "Round Table Discussion of Sulfanilamide". *Journal of Pediatrics* 13(4): 605-617.

Hecht. Annabel. 1984. "Sulfa; Yesterday's Hero is Still Taking Bows." *FDA Consumer* 18 (Oct.).

Kiefer, David M. 2001. "Miracle Medicines. The Advent of the Sulfa Drugs in the Mid-1930s Gave Physicians a Powerful Weapon." *Today's Chemist at Work* 10(6): 59-60.

Lerner, Barron H. 1991. "Scientific Evidence Versus Therapeutic Demand: The Introduction of Sulfonamides Revisited". *Annals of Internal Medicine* 115(4), 315-320.

Lesch, John E. 2007. *The First Miracle Drugs: How the Sulfa Drugs Transformed Medicine*. New York, NY: Oxford University Press.

Link, Bruce G., Mary E. Northridge, Jo C. Phelan and Michael L. Ganz. 1998. "Social Epidemiology and the Fundamental Cause Concept: On the Structuring of the Effective Cancer Screens by Socioeconomic Status,: *Milbank Quarterly* 76 (3): 375-402.

Long, Perrin H. and Eleanor A. Bliss. 1937. "Observations upon the Experimental and Clinical Use of Sulphanilamode in the Treatment of Certain Infections". *Journal of the American Medical Association* 108: 32-27.

____. 1939. *The Clinical and Experimental Use of Sulfanilamide, Sulfapyridine and Allied Compounds*. New York: Macmillan.

Loudon, Irvine 1987. "Puerperal Fever, the Streptococcus, and the Sulphonamides, 1911-1945". *British Medical Journal* 22: 485-490.

____. 1988. "Maternal Mortality: 1800-1950. Some Regional and International Comparisons." *Social History of Medicine* 1: 183-228.

_____. 1991. "On Maternal and Infant Mortality 1900-1960." *Social History of Medicine* 4: 29-73.

_____. 1992. *Death in Childbirth*. New York, NY: Oxford University Press.

_____. 2002. "The Use of Historical Controls and Concurrent Controls to Assess the Effects of Sulphonamides, 1936-1945". The James Lind Library (www.jameslindlibrary.org).

McKeown, Thomas. 1976. *The Role of Medicine: Dream, Mirage, Nemesis.* London: Nuffield Provincial Hospitals Trust.

McKinlay, John B. and Sonja M. McKinlay. 1977. "The Questionable Contribution of Medical Measures to the Decline of Mortality in the United States in the 20th Century". *The Milbank Memorial Fund Quarterly. Health and Society* 55(3): 405-428.

McKeown, T., R.G. Record and R.D. Turner. 1975. "An Interpretation of the Decline of Mortality in England and Wales During the Twentieth Century." *Population Studies* 29: 391–421.

Peihl, Anne Morrison, Suzanne J. Cooper, Anthony A. Braga, and David M. Kennedy. 2003. "Testing for Structural Breaks in the Evaluation of Programs." *The Review of Economics and Statistics* 85(3): 550-558.

Preston, Samuel. 1996. "American Longevity: Past, Present, and Future" Maxwell School of Citizenship and Public Affairs, Center for Policy Research, Syracuse University, Working Paper No. 7/1996.

Preston, Samuel H. and Michael R. Haines. 1991. *Fatal Years: Child Mortality in Late Nineteenth-Century America*. Princeton, NJ: Princeton University Press.

Quandt, Richard E. "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes." *Journal of the American Statistical Association*, Vol. 55 (290): 324-330

Sappington, S.W. and G.O. Favorite. "Sulfanilamide and Meningitis". *Annals of Internal Medicine* 13: S76-97.

Schuman, Andrew J. 2003. "A Concise History of Antimicrobial Therapy (Serendipity and All)". *Contemporary Pediatrics.* 20(10): 66-85

Stock, James H. and Mark W. Watson. 2003. *Introduction to Econometrics*. Boston, MA: Pearson Education, Inc.

Thomasson, Melissa A., and Jaret Treber 2008. "From Home to Hospital: The Evolution of Childbirth in the United States, 1927-1940." *Explorations in Economic History* 45(1): 76-99.

Weinshilboum, Richard M. 1987. "The Therapeutic Revolution". *Clinical Pharamcology & Therapeutics* 42(5), 481-4.

Welsch, William Miller and Jay Frank Schamberg. 1905. *Acute Contagious Diseases*. Philadelphia: Lea Brothers & Co.

WHO, UNICEF and UNFPA. 2004. *Maternal Mortality in 2000: Estimates Developed by WHO, UNICEF and UNFPA*. Geneva: World Health Organization.

Figure 1: Total mortality rate (in logs), 1920 – 1950

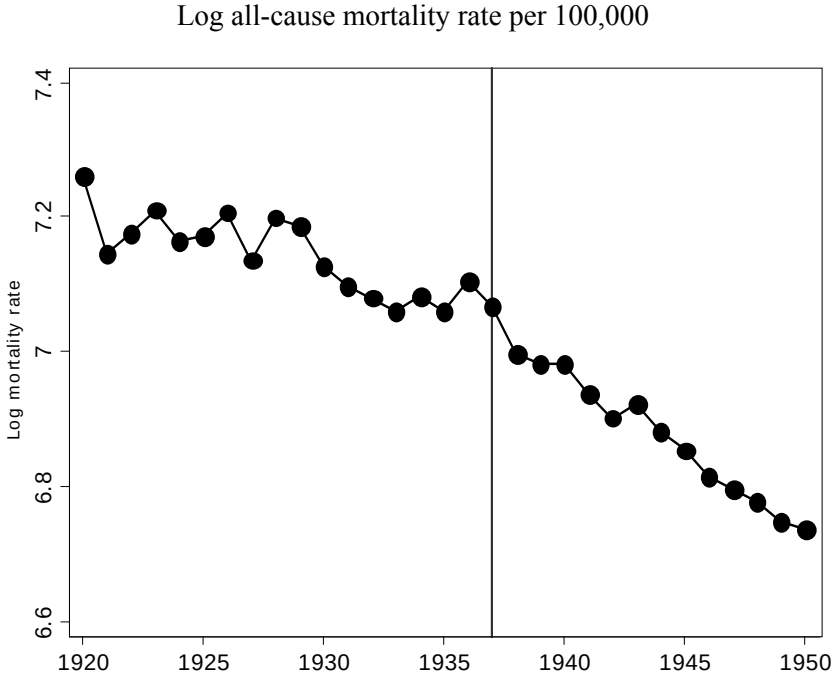
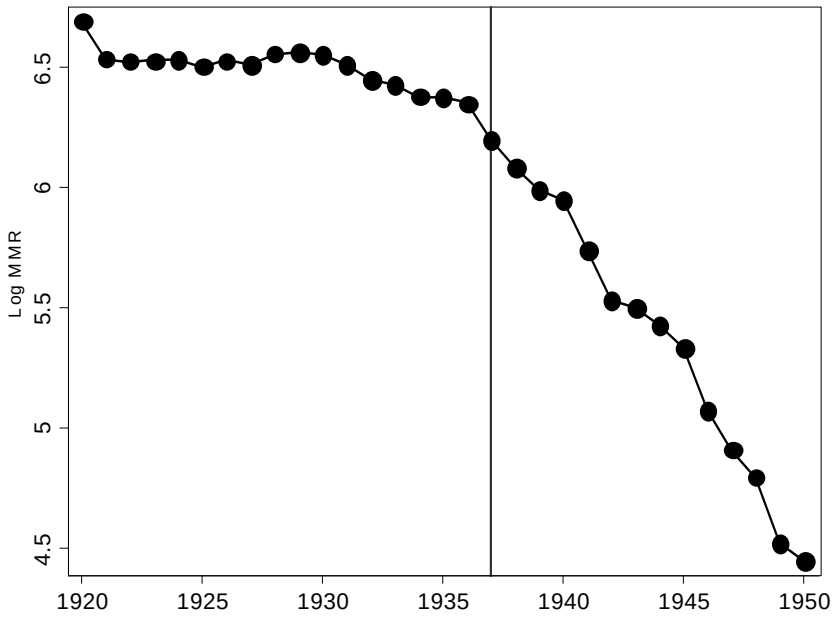


Figure 2: Mortality trends (in logs) for treated diseases, 1920 – 1950

a. Log maternal mortality ratio (deaths per 100,000 live births)



b. Log influenza and pneumonia mortality rate per 100,000

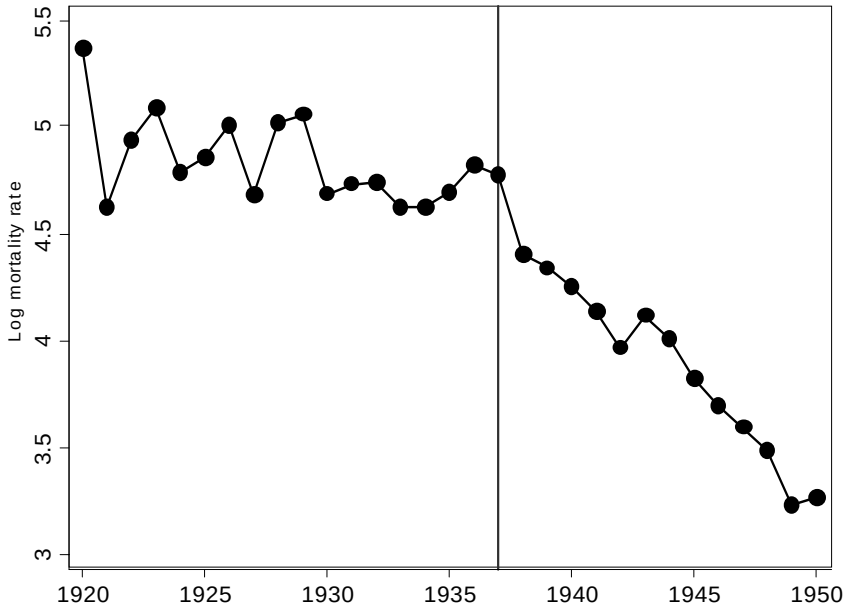
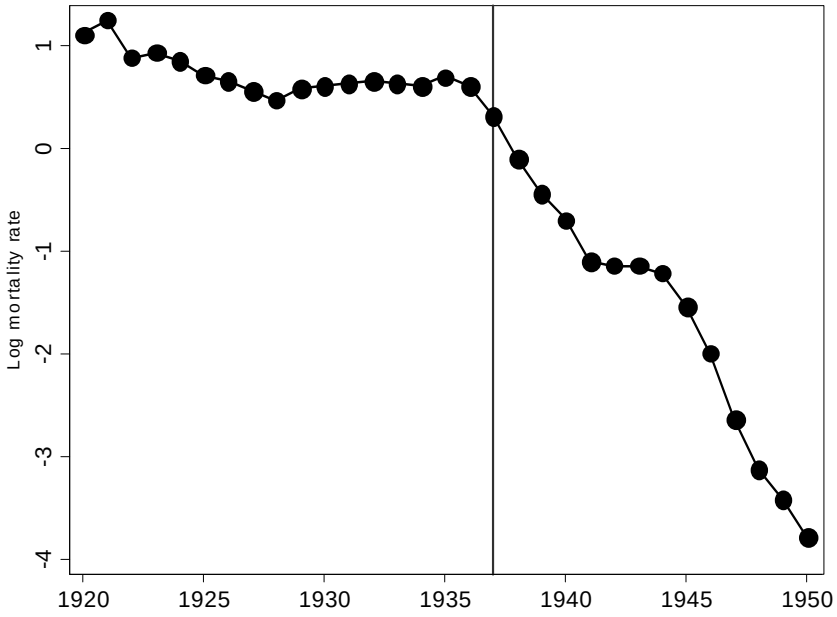


Figure 2, continued: Mortality trends (in logs) for treated diseases, 1920 – 1950

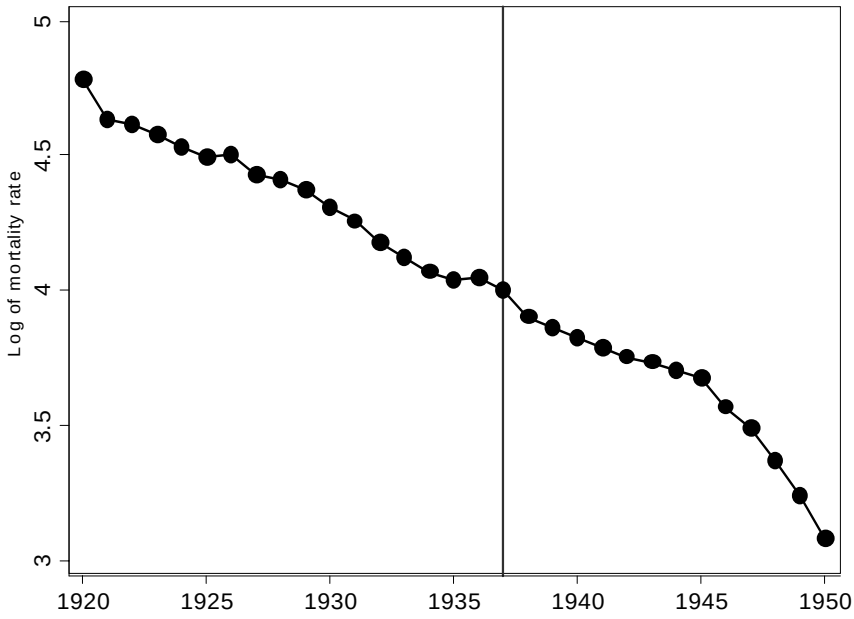
c. Log scarlet fever mortality rate per 100,000



d. Log meningitis mortality rate per 100,000

Figure 3: Mortality trends (in logs) for control diseases

a. Log tuberculosis mortality rate per 100,000



b. Log diphtheria mortality rate per 100,000

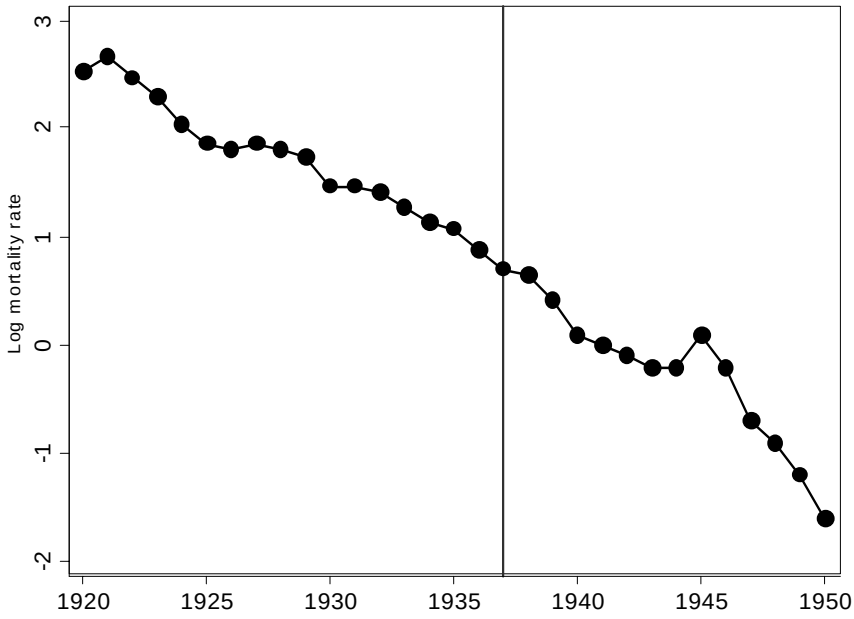


Figure 4: Mortality trends (in logs) for chronic diseases

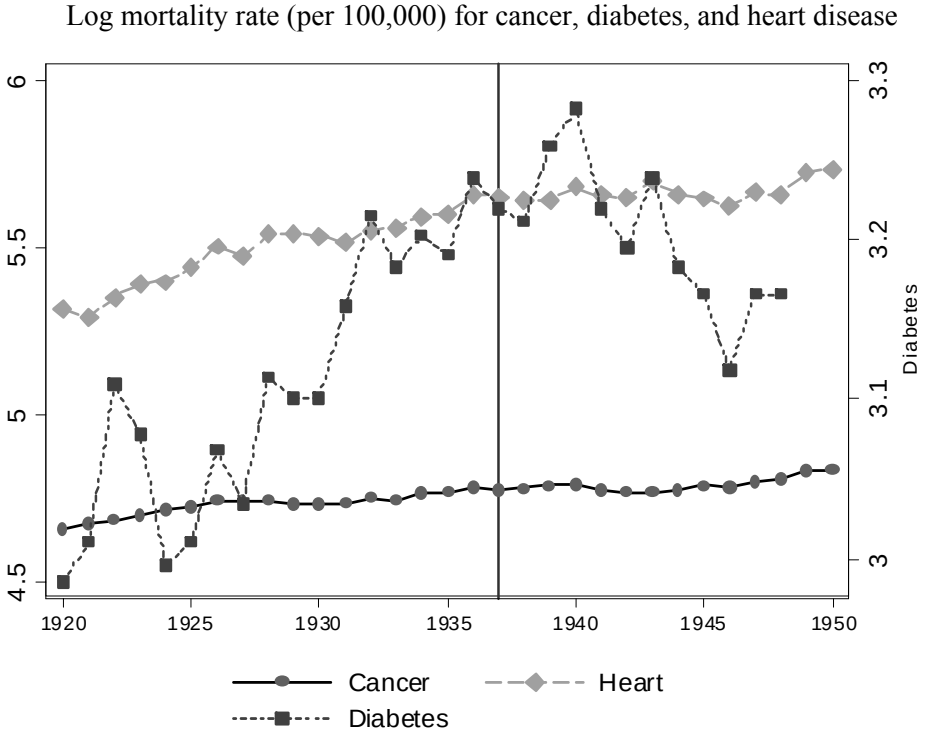


Figure 5: City and State trends in MMR (in logs), 1928 – 1940

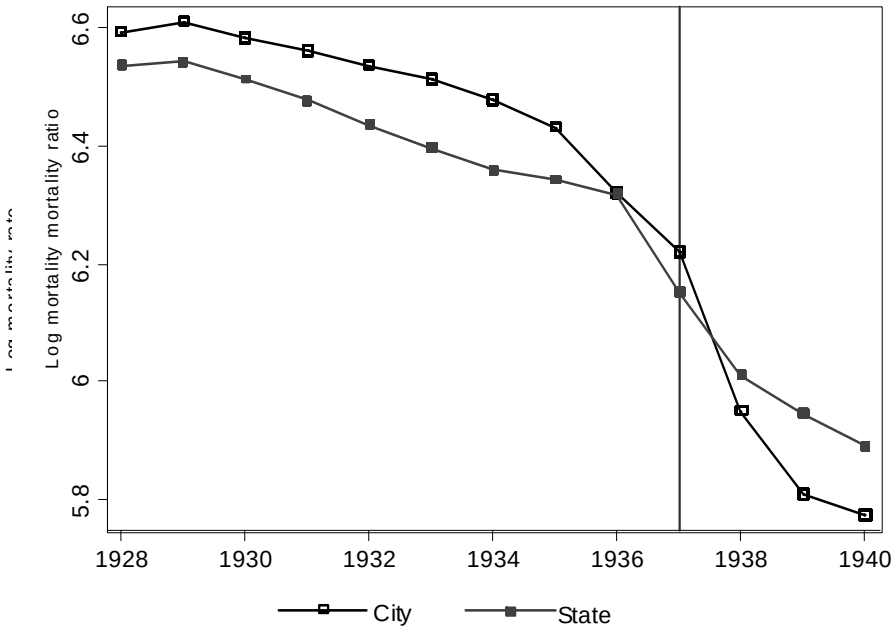
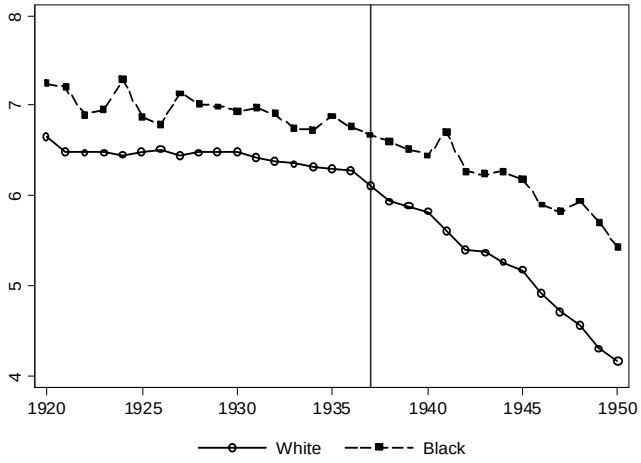
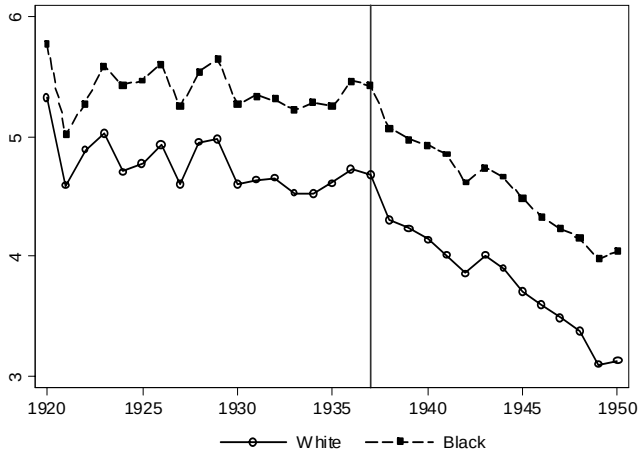


Figure 6: Mortality trends (in logs) by race, 1920 – 1950

a: Log maternal mortality ratio, by race



b: Log influenza and pneumonia mortality rate, by race



c: Log scarlet fever mortality rate, by race

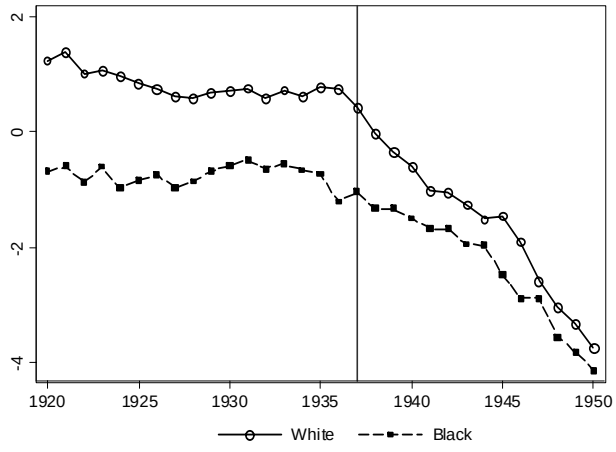


Table 1: National and state-level mortality statistics (deaths per 100,000)

	Panel A: National mortality rates				Panel B: Avg. state mortality rate			
	1920 to 1936		1937 to 1950		1925 to 1936		1937 to 1943	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
All-cause mortality	1270	74	998	114				
Treated diseases								
MMR	672	57	275	152	653	153	361	144
Pneumonia/flu	130	30	61	30	118	30	79	26
Scarlet Fever	2.1	0.5	0.5	0.5	2.2	1.6	0.7	0.7
Control diseases								
TB	81	18	41	10	70	43	46	24
Diphtheria	7	3.6	1.1	0.7	5	3.3	1.5	1.3
Chronic diseases								
Diabetes	22	2	23	4	19	7	23	8
Heart disease	242	25	288	9	205	64	264	78
Cancer	113	4	120	2	95	30	110	33
No. of states	*na*	*na*		*na*	*39 -48*		*48*	
By race								
MMR - White	630	58	242	143	644	149	358	112
MMR - Black	1070	190	546	202	1139	322	739	214
Pneumonia/flu - White	120	29	55	27	123	19	75	22
Pneumonia/flu - Black	223	43	117	56	241	69	141	40
Scarlet Fever - White	2.3	0.6	0.5	0.6	1.8	1.1	0.5	0.5
Scarlet Fever - Black	0.5	0.8	0.2	0.11	0.7	0.8	0.3	0.4
TB - White	68	18	32	8	72	19	38	12
TB - Black	209	32	115	27	219	57	125	40
No. of states	*na*	*na*		*na*	*14-18*		*18*	

Notes: In Panel A, national mortality rates are age-adjusted (using the U.S. population in 1940 as the standard), except for MMR (maternal deaths per 100,000 births). The summary statistics in Panel B are for states and years included in the regression analysis presented in Section 5 and 6. The analysis in Section 5 includes all states; the race analysis in Section 6 includes 18 states with black populations greater than 5% of the total state population. The state-level MMR data by race is only available for 10 of the 18 states in 1925; mortality rates for other causes by race are available for 14 of the 18 states in 1925.

Table 2: Estimated break dates in national and state-level mortality series

	National level		State-level
	Break year	Test statistic	% of states with break years between 1936 and 1938
All-cause mortality	1937	2.43	
Panel A: Diseases treated with sulfa drugs			
MMR	1937	6.02	29%
Pneumonia/flu	1938	3.02	67%
Scarlet Fever	1936**	10.48	32%
Panel B: Control diseases			
TB	1943***	22.94	13%
Diphtheria	1943	6.16	13%
Panel C: Chronic diseases			
Diabetes	1941	2.48	33%
Heart disease	1937	3.83	31%
Cancer	1941	6.49	11%

Notes: Trend breaks are estimated using the Quandt Likelihood Ratio test for a single trend break between 1933 and 1943. The national estimates are based on age-adjusted mortality data from 1920 to 1950 for all diseases except MMR. The state-level analysis is based on the 46 states that entered the registration area before 1930. For scarlet fever and diphtheria, states with zero mortality for any years between 1920 and 1948 are dropped from the analysis; the resulting sample sizes are 28 states for scarlet fever and 39 for diphtheria. The dependent variable is the first difference of the log mortality rate. Newey-West standard errors are computed, allowing for autocorrelation in the error term for up to two lags. The test statistic is the maximal F-statistic from a sequence of Chow tests for each possible break date in the test window. Critical values were obtained from Andrews (2003).
* $p < .10$; ** $p < .05$; *** $p < .01$

Table 3: Effect of sulfa drugs on mortality for treated diseases, 1937 - 1943

Dependent variable = ln (mortality)	MMR		Pneumonia/flu		Scarlet fever	
	(1)	(2)	(1)	(2)	(1)	(2)
Panel A: All years, 1925-1943						
Treated*Post-1937	-0.281***	-0.144***	-0.143***	-0.041**	-0.733***	-0.488***
	(0.028)	(0.027)	(0.020)	(0.018)	(0.078)	(0.079)
Treated*Year*Post-1937		-0.103***		-0.077***		-0.184***
		(0.0086)		(0.0056)		(0.021)
Observations	1736	1736	1736	1736	1720	1720
R-sq.	1.00	1.00	0.82	0.84	0.93	0.94
Panel B: Excluding 1935 to 1937						
Treated*Post-1937	-0.288***	-0.125***	-0.072**	-0.026	-0.713***	-0.510***
	(0.043)	(0.040)	(0.035)	(0.033)	(0.12)	(0.12)
Treated*Year*Post-1937		-0.117***		-0.033***		-0.146***
		(0.011)		(0.0072)		(0.021)
Observations	1448	1448	1448	1448	1432	1432
R-sq.	1.00	1.00	0.84	0.85	0.95	0.95

Notes: Treated diseases refer to maternal mortality, pneumonia/influenza, and scarlet fever. Tuberculosis serves as the control disease. Estimates are based on 1925 to 1943 state-level mortality data. Post-1937 is equal to one for years from 1937 to 1943. Model (1) includes the main effect of Treated, state*post fixed effects, a continuous year variable, and its interaction with Treated. Model (2) additionally includes state*post fixed effects and their interaction with the continuous year variable. Robust standard errors, clustered by state, are shown in parentheses. 16 state/year observations are dropped from the scarlet fever model due to zero mortality from scarlet fever.

* p < .10; ** p < .05; *** p < .01

Table 4: Urban-state differences in the effect of sulfa drugs on MMR

Dependent variable=ln(MMR)	(1)	(2)
Panel A: All years, 1928-1940		
Urban*Post-1937	-0.158***	-0.093**
	(0.042)	(0.036)
Urban*Year*Post-1937		-0.075***
		(0.026)
Observations	3966	3966
R-squared	0.94	0.94
Panel B: Excluding 1935 to 1937		
Urban*Post-1937	-0.329***	-0.298***
	(0.074)	(0.087)
Urban*Year*Post-1937		-0.024
		(0.031)
Observations	3038	3038
R-squared	0.94	0.94

Notes: Estimates are based on 1928 to 1940 mortality data for cities with populations of at least 25,000 (as of 1930) and for states. Post-1937 equals one for years from 1937 to 1940. Model (1) includes the main effect of Urban, state*post fixed effects, and a continuous year variable and its interaction with Urban. Model (2) additionally includes state*post fixed effects and their interaction with the continuous year variable. Robust standard errors, clustered by state, are shown in parentheses.

* p < .10; ** p < .05; *** p < .01

Table 5: Racial differences in national and state-level break years

	National-level		State-level		
	Break year		% of states with break years between 1936 and 1938		
			States with black populations >= 5%		All States
	White	Black	White	Black	White
All-cause mortality	1937		na	na	
Panel A: Diseases treated with sulfa drugs					
MMR	1937*	1941	29%	12%	37%
Pneumonia/flu	1938	1937	94%	82%	78%
Scarlet fever	1937**	1934	15%	15%	29%
Panel B: Control diseases					
TB	1943***	1940**	18%	18%	7%
Diphtheria	1943	1943	30%	10%	15%

Notes: Trend breaks are estimated using the Quandt Likelihood Ratio test for a single trend break between 1933 and 1943 and mortality data from 1920 to 1950. The national estimates are based on age and race-adjusted mortality data, except for MMR. The state-level race analysis is based on 18 states with black populations equal to at least 5% of the total population, except for scarlet fever and diphtheria, where states with zero mortality for any years before 1948 are dropped from the analysis; the resulting sample sizes are 13 for scarlet fever and 10 for diphtheria. The state-level analysis for whites-only is based on 46 states that entered the registration area before 1930, except for scarlet fever and diphtheria, which is based on 28 states and 39 states, respectively. The dependent variable is the log of mortality rate. Newey-West standard errors are computed, allowing for autocorrelation in the error term for up to two lags. The test statistic is the maximal F-statistic from a sequence of Chow tests for each possible break date in the test window. Critical values were obtained from Andrews (2003).

* p < .10; ** p < .05; *** p < .01

Table 6: Racial differences in the effect of sulfa drugs on mortality, 1937-1943

	MMR		Pneumonia/flu		Scarlet fever	
	(1)	(2)	(1)	(2)	(1)	(2)
Panel A: Whites						
Treated*Post-1937	-0.301***	-0.169***	-0.230***	-0.104***	-0.804***	-0.582***
	(0.044)	(0.036)	(0.022)	(0.020)	(0.12)	(0.15)
Treated*Year*Post-1937		-0.109***		-0.094***		-0.155***
		(0.011)		(0.008)		(0.039)
Obs.	644	644	652	652	539	539
R-sq.	1.00	1.00	0.86	0.88	0.98	0.98
Panel B: Blacks						
Treated*Post-1937	-0.133**	-0.029	-0.115**	-0.013	-0.134	-0.124
	(0.060)	(0.045)	(0.040)	(0.042)	(0.15)	(0.17)
Treated*Year*Post-1937		-0.081***		-0.076***		-0.032
		(0.019)		(0.011)		(0.029)
Obs.	644	644	652	652	500	500
R-sq.	1.00	1.00	0.79	0.83	0.98	0.98
Panel C: Fully interacted model						
Treated*Post-1937*Black	0.168***	0.140***	0.115***	0.091**	0.671***	0.458**
	(0.044)	(0.041)	(0.034)	(0.035)	(0.17)	(0.19)
Treated*Year*Post-1937*Black		0.028*		0.0179**		0.123***
		(0.016)		(0.007)		(0.020)
Obs.	1288	1288	1304	1304	1039	1039
R-sq.	1.00	1.00	0.92	0.93	0.98	0.98

Notes: Treated diseases refer to maternal mortality, pneumonia/flu, and scarlet fever. Tuberculosis serves as the control disease. Estimates are based on 1925 to 1943 state-level mortality data. Only states with black populations greater than or equal to 5% of the total population in 1936 are included (18 states for MMR and pneumonia/flu models). In three states, MMR data by race are available four years later than pneumonia/flu and scarlet fever data by race, which results in a total of 8 fewer state/year/race observations in the MMR models for each race. In the scarlet fever models, six states are dropped from the analysis due to a large proportion of year/state observations with zero mortality for blacks. In Panels A and B, Model (1) includes the main effect of Treated, state*post fixed effects, a continuous year variable, and its interaction with Treated; Model (2) additionally includes state*post fixed effects and their interaction with the continuous year variable. In Panel C, Models (1) and (2) include each of the same variables as in the previous panels plus their interaction with Black. Robust standard errors, clustered by state, are in parentheses.

* p < .10, ** p < .05, *** p < .01

Table A1: Robustness check: National break dates, by trimmed year range

	Log Mortality		Mortality Level	
	1925-1944	**1925-45**	**1925-1944**	**1925-45**
All-cause mortality	1944	1944	1944	1944
MMR	1937	1937	1937**	1937**
Pneumonia/flu	1938	1938	1937	1945
Scarlet fever	1936**	1936**	1936	1936
TB	1943***	1943***	1943***	1943***
Diphtheria	1925	1925	1941	1941
Diabetes	1925	1925	1925	1925
Heart disease	1944	1944	1944	1944
Cancer	1931	1931	1931	1931
By race				
MMR-White	1937	1937	1937**	1937**
MMR-Black	1925	1925	1925	1925
Pneumonia/flu-White	1938	1938	1937	1937
Pneumonia/flu-Black	1937	1945	1937	1945
Scarlet Fever-White	1937**	1937**	1937	1937
Scarlet Fever-Black	1932	1932	1932	1945
TB-White	1943	1943	1943	1943
TB-Black	1925	1925	1925	1925

Notes: Trend breaks are estimated using the Quandt Likelihood Ratio test for a single trend break between 1933 and 1943. The estimates are based on national age-adjusted mortality data from 1920 to 1950 for all diseases except MMR. The estimates by race are based on age- and race-adjusted mortality rates. The dependent variable is the first difference of the log mortality rate. Newey-West standard errors are computed, allowing for autocorrelation in the error term for up to two lags. The test statistic is the maximal F-statistic from a sequence of Chow tests for each possible break date in the test window. Critical values were obtained from Andrews (2003).
* $p < .10$; ** $p < .05$; and *** $p < .01$

Table A2: Robustness check: difference-in-difference regressions, dropping 1935-1936

Dependent variable = ln (mortality)	MMR		Scarlet fever	
	(1)	(2)	(1)	(2)
Panel A: Effect of sulfa drugs on mortality: dropping 1935-1936 (Table 3)				
Treated*Post-1937	-0.183***	-0.170***	-0.416***	-0.391***
	(0.037)	(0.034)	(0.11)	(0.12)
Treated*Year*Post-1937		-0.106***		-0.173***
		(0.0097)		(0.020)
Obs.	1544	1544	1528	1528
R-sq.	0.99	0.99	0.94	0.95
Panel B: Urban-state differences in the effect of sulfa drugs on the MMR: dropping 1935-36 (Table 4)				
Urban*Post-1937	-0.201***	-0.157***		
	(0.055)	(0.049)		
Urban*Year*Post-1937		-0.084***		
		(0.026)		
Observations	3348	3348		
R-squared	0.94	0.94		

Notes: Panel A estimates are based on 1925 to 1943 state-level mortality data. Model (1) includes the main effect of Treated, state*post fixed effects, a continuous year variable, and its interaction with Treated; Model (2) additionally includes state*post fixed effects and their interaction with the continuous year variable . Panel B estimates are based on 1928 to 1940 city-level data and state data. Model (1) includes the main effect of Urban, state*post fixed effects, a continuous year variable, and its interaction with Urban; Model (2) additionally includes state*post fixed effects and their interaction with the continuous year variable. For all regressions, robust standard errors, clustered by state, are shown in parentheses.

* p < .10, ** p < .05, *** p < .01

Table A3: Robustness check: Racial differences, dropping 1935-1937

	MMR		Pneumonia/flu		Scarlet fever	
	(1)	(2)	(1)	(2)	(1)	(2)
Panel A: Whites						
Treated*Post-1937	-0.329***	-0.176**	-0.214***	-0.127***	-1.281***	-0.971***
	(0.072)	(0.068)	(0.038)	(0.039)	(0.18)	(0.21)
Treated*Year*Post-1937		-0.125***		-0.062***		-0.211***
		(0.015)		(0.009)		(0.038)
Obs.	536	536	544	544	449	449
R-sq.	1.00	1.00	0.88	0.89	0.98	0.98
Panel B: Blacks						
Treated*Post-1937	-0.098	0.025	-0.053	-0.011	-0.470*	-0.377
	(0.10)	(0.083)	(0.064)	(0.070)	(0.23)	(0.25)
Treated*Year*Post-1937		-0.092***		-0.030***		-0.060
		(0.024)		(0.009)		(0.050)
Obs.	536	536	544	544	417	417
R-sq.	1.00	1.00	0.85	0.86	0.98	0.98
Panel C: Fully interacted model						
Treated*Post-1937*Black	0.231**	0.201*	0.161***	0.116**	0.812***	0.594*
	(0.11)	(0.097)	(0.050)	(0.053)	(0.27)	(0.28)
Treated*Year*Post-1937*Black		0.033		0.032***		0.151***
		(0.021)		(0.007)		(0.039)
Obs.	1072	1072	1088	1088	866	866
R-sq.	1.00	1.00	0.94	0.94	0.98	0.98

Notes: Treated diseases refer to maternal mortality, pneumonia/flu, and scarlet fever. Tuberculosis serves as the control disease. Estimates are based on 1925 to 1943 state-level mortality data. Only states with black populations greater than or equal to 5% of the total population in 1936 are included (18 states for MMR and pneumonia/flu models). In three states, MMR data by race are available four years later than pneumonia/flu and scarlet fever data by race, which results in a total of 8 fewer state/year/race observations in the MMR models for each race. In the scarlet fever models, six states are dropped from the analysis due to a large proportion of year/state observations with zero mortality for blacks. In Panels A and B, Model (1) includes the main effect of Treated, state*post fixed effects, a continuous year variable, and its interaction with Treated; Model (2) additionally includes state*post fixed effects and their interaction with the continuous year variable. In Panel C, Models (1) and (2) include the same variables as in the previous panels plus their interaction with Black. Robust standard errors, clustered by state, are in parentheses.

* p < .10, ** p < .05, *** p < .01