Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions

Sonia Bhalotra
Department of Economics, University of Bristol, UK
s.bhalotra@bristol.ac.uk

Atheendar Venkataramani
Department of Medicine, Massachusetts General Hospital
avenkataramani@partners.org

This version: November 2012

Abstract: We leverage the introduction of the first antibiotic therapies to examine the long-run impacts of pneumonia in infancy on education, employment, income and disability and find strong impacts on each of these outcomes, with the estimates for whites larger and more robust than those for blacks. We show that impacts for blacks are declining in measures of the severity of institutionalized segregation, which strongly inhibited human capital investments among blacks in the US South. Our results demonstrate the importance of responsive investments in translating improved early-life health into adult socioeconomic status as well as the persistent impacts of racial segregation.

Keywords: early childhood, human capital formation, complementarity, race, institutions, segregation, infectious disease, pneumonia, medical innovation, antibiotics, education, income

JEL codes: I18, H41

Acknowledgements – We would like to thank Douglas Almond, Tania Barham, Alan Barreca, Erlend Berg, Prashant Bharadwaj, Alfredo Burlando, Janet Currie, James Fenske, Jason Fletcher, Winnie Fung, Caroline Hoxby, Jane Humphries, Stephan Klasen, Eliana La Ferrara, Adriana Lleras-Muney, Anandi Mani, Bhashkar Mazumder, Alice Mesnard, and Michael Rothschild for helpful comments on this and an earlier version of this paper ("The Captain of the Men of Death and His Shadow: Long-Run Impacts of Early Life Pneumonia Exposure"). We have benefitted from presenting this work at the 2010 American Society of Health Economics conference and at the following events in 2011 and 2012: the American Economics Association Meeting, the NBER Children's Meeting, the UK launch meeting of Academics Stand Against Poverty, the University of Florence, Bocconi University, conferences held at DIAL Paris, DIW Berlin, the CAGE Research Centre at Warwick, European Association of Labour Economics, and the Royal Economic Society, the Health and Human Capital workshop in Mannheim, and seminars at Oxford University (Wellcome Trust Unit for the Study of the History of Medicine), the CSAE (Economics), the Health Economics Research Centre, and the University of Cape Town. We are grateful to Anne Case, Adriana Lleras-Muney, Bhash Mazumdar, Grant Miller, and Nathan Nunn for making available their historical data, Sarah Karlsberg-Schaffer for assistance in extending these historical databases, and Damian Clarke and Eddy Tam for their excellent assistance at stages of the data analysis. All errors are our own.

Abstract (Long Form): While there is mounting evidence of the long-run impacts of early life shocks, less is known about the extent to which investments made later in the life course matter in modulating these impacts. We leverage the introduction of the first antibiotic therapies to examine the long-run impacts of pneumonia in infancy on education, employment, income and disability and find strong impacts on each of these outcomes, with the estimates for whites larger and more robust than those for blacks. We then demonstrate that impacts for blacks are declining by measures of the severity of institutionalized segregation, which strongly inhibited human capital investments among blacks in the US South. Our results highlight the importance of responsive investments in translating improved early-life health into adult socioeconomic status as well as the persistent impacts of racial segregation.

I. Introduction

The setting for our study is America in the 1930s and early 1940s, when pneumonia accounted for one of every ten deaths. Barring mortality from premature birth, pneumonia was the leading cause of infant mortality (Linder and Grove, 1947; Wegman, 2001), as it is in developing countries today (Black et al., 2010). The ubiquity and ruthlessness of the disease during this time period led the iconic physician Sir William Osler to coin pneumonia as the "Captain of the Men of Death". The foundation for the present work is an investigation on the extent to which contemporary American adults carry the scars of exposure to pneumonia in their childhood. Such long-run impacts are theoretically possible given insights from the biomedical literature, which suggests that infections such as pneumonia may lead to permanent physiological changes by provoking inflammatory responses that divert nutritional resources away from physical and mental development. As a result, severe or repeated infections may lead to long-term "scarring" in the form of poorer health, reduced cognitive development, and ultimately reduced lifetime income and utility in adulthood (Crimmins and Finch, 2006; Eppig et al., 2010).

Consistent with this, recent work in economics demonstrates the impacts of the fetal and childhood health environment on a range of adult outcomes, including employment, income, and disability[1]. However, the reduced form impacts that dominate this body of work bundle the effects of initial changes in the biological endowment with those from subsequent investments. That is, the extent to which investments made later in life matter in driving the link between early life endowments and socioeconomic realizations in adulthood is unknown. Clearly, building an understanding of these relationship is relevant for policy. For instance, if the technology of human

[1] For instance, see Case, et al (2005), Almond (2006), Currie and Moretti (2007), Case and Paxson (2009), Bleakley (2007 and 2010), Cutler et al (2010), Lucas (2010), and Venkataramani, (2012). Almond and Currie (2011a, b) survey this literature.

capital formation is such that returns to post-infant investments are increasing in the infant endowment, then subsequent investments would be necessary to fully realize potential of early life interventions[2]. This idea of dynamic complementarity has been recently formalized and incorporated into extensions of the Becker and Tomes (1986) model (Cunha and Heckman 2007, Heckman 2007, Cunha et al., 2010).

However, advances in theory aside, empirical evidence regarding complementarity remains scarce (Almond and Currie 2011b, ostensibly because exogenous variation both in early life endowments and in subsequent investments is required to establish credible identification, a situation which is understandably difficult to achieve. For example, Aizer and Cunha (2012) demonstrate how the impacts of plausibly exogenous exposure to the Head Start program among American children on cognitive test scores increases in the quality of their infant endowment. However, interpreting this as evidence of complementarity requires strong assumptions, as it is not clear whether the infant endowments they use are exogenously determined. Moreover, Aizer and Cunha (2012) do not examine outcomes beyond the age of 7, so the importance of complementarity in driving lifetime income, which is of primary policy interest, cannot be ascertained from these data. We contribute to this literature by using exogenous variation across cohorts in the infant health endowment (here, driven by changes in pneumonia morbidity) and combining this with historically determined and plausibly exogenous variation in returns to human capital investments across racial groups. We then assess the quantitative importance of subsequent investments (i.e., proxied by their effective price) in driving the impacts of early childhood endowments on socioeconomic outcomes in adulthood. The remainder of this section first summarizes our approach and our main findings in greater detail.

---

[2] Subsequent investments, such as those made by parents or households, may respond to endowment shocks in a reinforcing manner if there are complementarities between early and late childhood investments. Adhvaryu and Nyshadham (2012) and Bharadwaj et al. (2012) exploit exogenous variation in endowments to examine whether parents invest differently in children on the basis of endowment and quality. Importantly, these studies do not model adult outcomes and therefore do not shed light on the importance of these responsive investments in the long run.

Whether parental responses to early life endowment shocks are consequential for long-run income realizations or utility has recently been debated in the literature. Bleakley (2010) argues that such investments are unlikely to contribute to the long-run impacts of early life shocks as the marginal returns to investments for families that optimize on the basis of endowments is necessarily zero. However, as Adhvaryu and Nyshadham point out, parental investments may be consequential if parents are unable to invest near the full information/unconstrained optimum or if there are spillovers from such investments that are not properly internalized. As we discuss in more detail below, the identifying variation we use for investments involves institutional constraints to access and the returns to human capital by race. Thus, marginal investments made by parents in this setting are likely not at the optimum.

In the first part of the paper, we establish the impact of pneumonia in infancy on adult outcomes. We focus on infancy because the bulk of childhood morbidity and mortality from pneumonia occurs during this period. Moreover, infancy is a period of high velocity physical and mental growth characterized by high nutritional requirements: this developmental plasticity creates the possibility that nutritional stressors, such as infections, can lead to irreversible damage (Almond and Currie, 2011a, b; Barker and Osmond, 1986; Eppig, et al, 2010).

Recognizing that the identification of causal effects of early life infections on later life attainments is challenged by selectivity into infection, we turn to the sharp and plausibly exogenous cohort variation in pneumonia exposure created by the arrival of sulfonamide antibiotics (sulfa drugs) in 1937. These agents were the first clinically available antibiotics and were quickly utilized to treat a variety of potentially fatal bacterial infections. Their arrival was responsible for sharp decreases in morbidity and mortality from pneumonia (Greengard, et al, 1943; Lesch, 2007; Jayachandran, et al, 2010), especially among infants and young children. We exploit the fact that states most burdened by pneumonia in the pre-sulfa era experienced the largest declines upon the introduction of sulfa drugs, an approach similar to that in Acemoglu and Johnson (2007) and Bleakley (2007). In particular, we investigate whether the post-sulfa convergence in pneumonia mortality rates across the states after 1937 is mirrored in longer run socioeconomic outcomes for cohorts born in the sulfa era for whom adult outcomes are recorded in United States Census data from 1980-2000.

We find that cohorts exposed to sulfa-led declines in pneumonia in their birth year achieved large improvements in schooling, employment and income and lower risks of cognitive and work-related disability and poverty in adulthood. The long run benefits are large, and larger in states with higher pre-intervention levels of pneumonia. For instance, the estimates for men (we do not find evidence of impacts for women) suggest that for post-1937 birth cohorts, a change in pneumonia exposure corresponding to a move from the 75th to the 25th percentile of the pre-sulfa pneumonia death rate (roughly equivalent to the nationwide 30% drop in pneumonia mortality rates after 1937) resulted in a 2.05% [1.02%] point increase in the probability of completing high school [college] and a 2.11% increase in family income. Using a more flexible specification, we confirm that the only consistently significant long run impacts of pneumonia flow from birth year exposure, consistent with the birth year being a critical period for physical and cognitive development. The results, particularly those for schooling, employment, and income, are robust to the addition of birth state by birth cohort level controls for income, health and educational infrastructure, a wide range of

communicable and non-communicable diseases (including placebo diseases), and birth state specific time trends. We show that they hold up in number of other robustness checks, as well, and cannot be explained by selective migration in response to improvements in the disease environment. We establish that the estimated returns are comparable to those from other recently studied public health and educational interventions.

In the second part of the paper, we provide evidence regarding the extent to which the aforementioned gains require rely upon subsequent investments to materialize. To do this, we utilize measures of cross-birth state variation in the intensity of institutionalized racial segregation, which has been shown in numerous other studies to have discouraged human capital accumulation by blacks through reduced access to quality schools and depressed labor market returns to education (Card and Krueger, 1992; Donohue and Heckman, 1991; Smith, 1984). We adopt these measures, which effectively proxy for the (opportunity) cost of human capital investment, in lieu of information on actual investments, which are difficult to come by in historical data and, more importantly for our purposes, likely to be endogenous. Put somewhat differently, by utilizing variation in (actual) *returns* to human capital generated by plausibly exogenous institutions of racial segregation, we are able to examine how gains in income and employment in adulthood flowing from an infant health intervention look with what is effectively restricted *versus* unrestricted intermediate investments.

Interacting our sulfa drug-treatment exposure term with a binary indicator for birth in the US South, we find that blacks born outside of the South, where segregation was not institutionalized, enjoyed sizeable gains in education, employment, and income from reduced early life pneumonia. Indeed, these gains were similar in magnitude, if not larger, than those experienced by whites. However, for blacks born in the South where segregation was *de jure*, sulfa-induced health improvements contributed little to their economic progress. In order to more definitively isolate the role of race-specific institutions, we also interact the sulfa treatment term with the birth state share of slaves in the population in 1860, which has been shown in other work to be predictive of institutions such as the quality of public schools and suffrage seven to eight decades later (Engerman and Sokoloff, 2005; Mariscal and Sokoloff, 2000; Nunn, 2008), as well as the ratio of educational attainment between young black and white men, which provides a continuous measure of actual human capital investments made under institutionalized segregation. In both cases, we find that the impacts of reduced pneumonia in infancy decrease with the intensity of institutionalized segregation.

Notably, we do not find any similar gradients for white men, which suggests that the results are not being driven a more general economic process. In addition, we investigate and undermine competing explanations for these black-white differences such as race*region variation in access to sulfa drugs, measurement error in age reporting and mortality, and selective endogenous migration from the South to the North.

Collectively, these results suggest that responsive investments are critical to the realization of long-run socioeconomic gains from improved infant health, consistent with dynamic complementarity across inputs in the production of human capital. In this way, we illustrate how (good) biology is not necessarily destiny. Understanding the production function for human capital is clearly relevant to understanding the origins of socioeconomic inequalities and, accordingly, for the design of social policy.

Along these lines, our results also inform the literature on race inequality in the United States. Several studies document the long-run effects of segregation in access to health care (Almond and Chay, 2006; Chay et al., 2009), schools and labor markets (Aaronson and Mazumdar, 2011; Donohue and Heckman, 1991; Johnson, 2011; Neal, 2006; Smith, 1984). Our results illustrate another, more insidious, legacy of segregation. In particular, we show how institutionalized segregation in schools and labor markets prevented (Southern) blacks from consolidating the potential long run returns to exogenous improvements in their health and cognitive endowments. As such, the potential of a generation of black children born in the post-sulfa era went underutilized at a time where America was experiencing rapid, inclusive growth as a result of the expansion of state-financed education alongside skill-biased technological change (Goldin and Katz, 2008).[3]

Another contribution of this paper is that we provide the first estimates of the long-run health, cognitive and socioeconomic impacts of pneumonia.[4] This is of immense importance given that pneumonia remains the leading cause of child death, killing 1.6 million children every year, which is more than AIDS, malaria, and tuberculosis combined (World Health Organization, 2011). Despite this, only 20% of children who have pneumonia are able to access antibiotics and pneumonia vaccination rates, too, remain dismal. While pneumonia is concentrated in developing

[3] We show how race discrimination in access to schools and labour markets cramped long run returns to pan-race health interventions. A similar "cross-effect" is studied by Chay et al (2009) who show how discrimination in access to health care at birth contributed to race gaps in test scores.

[4] While the biomedical literature reports associations of childhood pneumonia and adult lung function (Johnston et al, 2012), there is no evidence of its impact on cognitive ability, human capital, earnings or employment.

countries (Bhutta, 2007; Black et al, 2010) it remains a significant concern in developed countries (Farha and Thomson, 2005). Our use of the introduction of sulfa drugs as a source of identifying variation in exposure to pneumonia infection yields what would appear to be the first estimate of the benefits of medical therapeutics. Our findings indicate that the common neglect in the medical and global health literature of long run and socioeconomic benefits of investments in R&D (for vaccines and drugs) will have led to persistent under-investment. Thus, this paper stands to inform investments in medical research and contemporary debates concerning changes to patent laws and marketing structures that may help promote the innovation and distribution of antibiotics and vaccines.

The remainder of the paper is as follows. Section II begins by profiling pneumonia mortality and morbidity in the 1930s United States with particular emphasis on changes in disease rates as a result of the sulfa drug revolution. After describing the data and core research strategy, we then demonstrate the impact of reduced early childhood pneumonia on long-run socioeconomic outcomes. Section III examines gradients in the long-run impacts of reduced early life pneumonia among blacks by indicators for institutional segregation, presenting estimates consistent with complementarity between early life endowments and subsequent human capital investments. Section IV discusses the policy implications of our results and concludes.

II. The Long-Run Impacts of Early Childhood Pneumonia

A. Pneumonia and the Sulfa Drug Revolution

Pneumonia is an acute inflammatory disease of the lung characterized by fevers, shortness of breath, and cough, and is typically caused by bacteria and viruses. Bacterial pneumonia is more severe and more likely to be fatal than its viral counterpart. In this subsection, we describe the prevalence and distribution of pneumonia in the 1930s and 1940s, underlining its contribution to overall mortality rates and its particularly high incidence in infancy. We then discuss evidence of the sharp drop in pneumonia mortality and corresponding declines in morbidity that flowed from the plausibly exogenous introduction of sulfa drugs in 1937. We show that this stimulated convergence in pneumonia rates across the US states.

In the 1930s, pneumonia accounted for 10% of all-age deaths in the United States and about 44% of neonatal deaths. In terms of morbidity, estimates from the US National Health Survey of 1934-1936 show a case rate of 3 per 100 infants, with rates twice as high among the poor and those

living in crowded conditions (Britten, 1942). Rates for infants were twice as high as for 1-4 year olds and nearly 10 times larger than for 10-15 year olds (*Figure 1*). It is estimated that these figures represent a significant (at least two-fold) underestimate of the true pneumonia burden during this era (Klugman and Feldman, 2009). Indeed, morbidity rates among infants from poorer families may have been similar to that in today's developing countries, where it is estimated that there are between 15 and 28 pneumonia cases per 100 children under the age of 4 each year (Lopez et al., 2006; Rudan et al., 2004). The duration of morbidity was significant: in the pre-antibiotic era, pneumonia was a long and trying illness, resulting in an average of 39 days of disability per patient in 1936 (Britten, 1942). In terms of geography, pneumonia was more prevalent in the American South and some parts of the West (*Figure 2*), consistent with its known risk factors (poverty, overcrowding, and poor nutrition) being more prevalent in these regions in the 1930s (Klugman and Feldman, 2009; van der Poll and Opal, 2009).[5]

Prior to the arrival of sulfa drugs, pneumonia was primarily treated with supportive care, particularly among infants and young children[6]. The seeds for pharmaceutical therapy were initially sown in 1932, when German chemists conducting experiments on textile dyes discovered the antibiotic properties of sulfonamides. The first scientific evidence of their potential was published in 1935 and this was confirmed in clinical trials conducted in the following two years (Gibberd, 1937; Kiefer, 2001; Lesch, 2007; Long and Bliss, 1937). A *New York Times* article in the winter of 1936 lauded the potential benefits of sulfa and, by early 1937, the drugs became widely available in the United States. They were relatively inexpensive and heavily promoted and quickly adopted to treat a range of conditions, leading to a "sulfa craze" that lasted until the mass-availability of the first penicillins in the mid 1940s (Jayachandran et al., 2010). The first sulfa agents, such as Prontosil, were only somewhat effective against *Streptococcus pneumonaie*, the agent responsible for the majority of bacterial pneumonias. However, in 1938, sulfapyridine (also known as M&B 693) became available for clinical use. Clinical trials conducted soon after sulfapyridine became available showed striking

[5] The estimated equation includes state*year varying variables such as income the impacts of which are allowed to break in 1937, and also state fixed effects and trends.

[6] Intravenous serum therapy, where antibodies to the bacteria infecting a patient were harvested in animals and thereafter introduced into the patient intravenously, was introduced among hospitalized patients in the early 1930s (Lesch, 2007). While this was successful in certain contexts (Finland, 1960), it was not utilized all that widely and appears to have had no impact on pneumonia mortality rates at the population level (*Figure 3*). Moreover, serum therapy was *less likely to be used in infants and young children* given greater difficulty in administration and more pronounced side effects (Connolly, et al 2012).

reductions of 50-70% in pneumonia case fatality rates among inpatients (Evans and Gaisford, 1938; Gaisford, 1939; Lesch, 2007).

Consistent with findings from small clinical trials, sulfa drugs had large impacts on mortality from pneumonia at the population level. Jayachandran et al. (2010) demonstrate a structural break in the time series data for all-age mortality from sulfa treatable diseases in 1937, which is evident in *Figure 3, Panel A*. They estimate that sulfa drugs led to a 17% decline in pneumonia mortality. *Panel B* of *Figure 3* illustrates that the largest (absolute) decline accrued to infants, which is consistent with their higher pre-antibiotic era infection rates. Moreover, as illustrated in *Figure 4*, larger absolute reductions in pneumonia mortality were seen in states with higher pre-sulfa drug era disease burdens. This pattern was evident for all age groups (*Panel A*) but, again, it was infants (*Panel B*) in high pneumonia states who experienced the most dramatic reductions in disease mortality. As we discuss below, we exploit the implied convergence across states after 1937 in our identification strategy.

In addition to reducing mortality, there is strong evidence that sulfa drugs led to reductions in the severity of pneumonia episodes. Connolly, et al (2012) note how physicians at a Baltimore hospital were struck by "how favorably (and rapidly) children with pneumonia responded to sulfapyridine…[they] described one severely ill thirteen-month old girl who 'within 24 hours after sulfapyridine' could be found 'sitting up in bed entirely well.'" Clinical trials on infants and children from the era back up these anecdotal claims, citing rapid improvements in fever, mental status and other physical examination findings, and demonstrating that the average inpatient case of pneumonia was shorter and followed a less severe course as a result of sulfa drug therapy (Greengard et al., 1943; Hodes et al., 1939; Moody and Knouf, 1940; Smith and Nemir, 1939)[7].

Sulfa drugs likely also led to reductions in pneumonia morbidity in the community, where roughly 70% of cases were treated in the mid 1930s, in addition to their profound impact on hospitalized patients (Britten 1942). This is because sulfa drugs were widely available to, and utilized by, laypersons and community physicians soon after their arrival in the US (Lesch, 2007; Lerner, 1991). Indeed, they were available without prescription until 1939. Furthermore, sulfa antibiotics cost $4.3 per patient per day or $28-$100 (in 2008 USD) for a complete course, which is inexpensive for a life saving drug (Jayachandran et al. 2010).

[7] Among adults, comparison of US Army experiences between the first and second World Wars suggest that sulfa drugs were instrumental in drastically reducing the severity of and infirmity time from pneumonia among soldiers. Data on industrial workers illustrate a 20-30% reduction in the number of illness days after the arrival of sulfa drugs (Ungerleider et al., 1943).

B. Data

Data for the outcome variables – years of schooling, high school and college completion, age of first marriage, logged family income, poverty status, employment status, work limiting or preventing disability, and physical and cognitive disability - were taken from 5% samples of the United States Census for the years 1980, 1990, and 2000 (Ruggles et al., 2010). Cohorts born in 1937, the year sulfa drugs became available, were 43, 53, and 63 years old during these enumerations, respectively. We calculated means for each outcome by birth state, birth year, census year, race and gender and used these cell means in our analysis. To avoid measurement error, we dropped cells in the bottom 1% of the cohort size distribution (those with less than 50 persons), although our results are not sensitive to this restriction.

We chose to examine high school and college completion in addition to years of schooling given that sample cohorts were entering high school and college towards the end and the beginning, respectively, of booms in enrollment at these levels of education (Goldin and Katz, 2008). Examining disability measures other than self-reported work related disability allows us to address potential "justification bias" in welfare regimes where there is an incentive to exaggerate or invent disability (Autor and Duggan, 2003). For the employment and income variables, we pool data from the three census rounds to help separate cohort effects from life-cycle trajectories in these indicators[8]. For the education and age of first marriage variables, we use only the 1980 Census as we do not expect attainment to change after the age of 37 (age of youngest cohorts in the sample in 1980). The cognitive and physical disability variables are only available in the 2000 census.

State level data on disease-specific mortality rates for 1930-43 and a host of socioeconomic characteristics were gathered from a variety of sources (detailed in **Data Appendix)**. For this period, annual time series data by state record pneumonia mortality jointly with influenza mortality so we, like Jayachandran, et al (2010), work with this compound variable. Combining mortality rates from pneumonia and influenza has the potential advantage of lowering measurement error given that surveillance systems may have conflated influenza and pneumonia deaths given that a large proportion of influenza deaths were caused by secondary bacterial pneumonia.[9]

---

[8] So as to investigate the potential role of age or lifecycle effects, we estimated the equations by census year. The coefficient of interest was similar, which suggests that it is cohort variation that drives our estimates

[9] Additionally, decadal statistics that provide separate estimates for pneumonia and influenza show that pneumonia was dominant and, importantly, that the decline in pneumonia plus influenza mortality between 1930 and 1940 was entirely on account of pneumonia. This is consistent with influenza, a viral illness, being

Also in order to reduce measurement error, we use data on all-age pneumonia mortality rather than age specific pneumonia rates. It is well known that states varied markedly in terms of the accuracy of birth and death registration, rendering cross-state comparisons, particularly in the pre-antibiotic period, of infant mortality difficult (Linder and Grove, 1947). The fact that absolute contribution of (the decline in) infant pneumonia mortality to the all-age rate was larger than for other age groups (*Figures 3* and *4*), gives us the freedom to use these data in lieu of what is understood to be a noisier series (see **Data Appendix** for further details).

Descriptive statistics for the outcomes, disease, and controls (discussed below) are provided in *Appendix Table 1*. The **Data Appendix** provides detailed information on sources and variable definitions.

C. Research Strategy

C.1. Baseline framework

Infectious disease incidence tends to be selective. In the sample period, pneumonia, mortality and morbidity rates were higher among the poor, rural residents, and blacks (Britten, 1942). So, even if disease exposure had no causal effects on later life attainments, we would likely observe that higher levels of disease exposure in childhood are associated with lower socioeconomic status in adulthood. We avoid this problem by exploiting the nationwide arrival of sulfa drugs in 1937, which generated sharp cohort variation in pneumonia exposure (*Figure 3*) along with the cross-state convergence in pneumonia burdens (*Figure 4*)[10]. Our baseline specification is:

$$Y_{rstc} = \alpha + \beta * post_t * base\_pneumonia_s + \theta_{rs} + \eta_{rt} + \lambda_{rc} + e_{rstc}$$

where Y_{rstc} denotes an outcome recorded in adulthood for individuals of race (r), birth state (s) and birth year (t) observed in census year (c). The outcomes, listed earlier, are indicators either of human capital or of the returns to human capital. We define $post_t = 1$ for cohorts born in and after 1937. The pre-sulfa pneumonia mortality rate in the birth state is denoted $base\_pneumonia_s$ and may be thought of as capturing treatment intensity. Infectious disease mortality rates are commonly used to

---

unresponsive to sulfa drugs. The fluctuations in the combined series evident in *Figure 3, Panel A*, prior to 1937 are driven by influenza outbreaks (Jayachandran et al., 2010).

[10] We base our identification strategy on the arrival of the sulfa drug technology at the national level instead of the timing of their availability at the state level as any state specific differences in adoption rates are likely endogenous.

proxy disease exposure; for a formalization of the rationale for this, see Bozzoli et al. (2009). Our research strategy is similar in spirit to that employed by Acemoglu and Johnson (2007), Bleakley (2007), Cutler et al. (2010) and Lucas (2010).

The Greek letters represent race-specific fixed effects for birth state, birth year, and census year. We cluster standard errors at the birth state level to account for serial correlation in the outcomes (Bertrand et al, 2004). Regressions are weighted by cell size, although the estimates are not sensitive to this. We restrict our analysis to the time period 1930-1943 in order to reduce the possibility of confounding from other public health events or interventions, for example, the influenza epidemic of 1928-9 and the increasingly widespread use of penicillin after 1943. We estimate all models separately by gender, given that gender differences in the impacts of early life shocks have been noted in numerous other studies and that boys are thought biologically more sensitive to the early life health environment than girls (Gluckman and Hanson, 2005; Low, 2000; Waldron, 1983).

The estimated equation may be thought of as the reduced form of a system in which adult outcomes are allowed to depend upon pneumonia exposure at birth and the latter is instrumented with the sharp arrival of sulfa drugs, the impact of which varies across states as a function of their pre-sulfa pneumonia burden. The parameter β captures the full impact of sulfa-induced reductions in pneumonia exposure. This will include direct biological impacts of pneumonia infections on mental and physical growth in infancy as well as any subsequent investments made in response to these endowment changes.[11]

C.2. Threats to Inference

The main threat to inference with the differences-in-differences type of model above is the presence of unobservables that vary by birth state and birth year that are correlated with both pneumonia exposure in infancy and the outcomes. We therefore investigate the sensitivity of our results to a variety of controls for underlying trends in health and socioeconomic conditions in the birth state.

Diseases not treatable with sulfa drugs: A particular concern is that *post*base_pneumonia* may pick up sudden improvements in health arising independently of, but coincident with, sulfa

[11] It will in principle also include any indirect impacts flowing, for example, from the increase in household resources generated by lower pneumonia morbidity amongst parents of the index cohort. Since pneumonia infections amongst adults of childbearing age were relatively scarce (*Figure 2*), we expect such indirect effects to be small.

drug availability, such as state-specific public health interventions. A powerful check against this is control for trends in diseases that were not treatable with sulfa drugs. The premise is that omitted factors will not have discriminated between sulfa-treatable and sulfa-untreatable diseases. In particular, we interact *post* with pre-sulfa birth state specific mortality rates from tuberculosis, diarrhea (under the age of 2), malaria, heart disease and cancer and include these as controls in our model. In this way we allow the outcomes to evolve as discontinuous functions of these five other diseases. Diseases not treated by sulfa drugs therefore function like placebos in that we expect the coefficient on *post*base_(untreated disease)* to be close to zero.

Inclusion of communicable diseases (tuberculosis, diarrhea, malaria) helps control for state specific changes in sanitation, public health programs, housing, etc, that may have coincided with the arrival of sulfa drugs. Controlling for malaria, which was declining significantly during the study period (Barreca, et al, 2012), is likely to be particularly important in isolating the impact of pneumonia given that, like pneumonia, malaria enjoyed a higher prevalence in the South and that early life exposure to malaria has been shown to influence later life outcomes (Barreca, 2010; Bleakley, 2010; Cutler et al., 2010; Lucas, 2010; Venkataramani, 2012). Including diarrhea, a disease of childhood, helps control for unobservables specific to the young. The inclusion of non-communicable diseases is expected to control for factors such as health care quality and access.

Diseases treatable with sulfa drugs*:* Sulfa drugs led to marked declines in conditions other than pneumonia, most notably scarlet fever, erysipelas, meningitis and puerperal sepsis (Jayachandran et al., 2010). Thus, the coefficient on $post_t$*$base\_pneumonia_s$ may have loaded on to it the effects of reductions in these omitted sulfa-treatable diseases. However, scarlet fever, erysipelas and meningitis accounted for a negligible fraction of infant and all-age mortality. With regards to the former, in 1930, the number of infant deaths in 1000 live births from these diseases was 0.1, 0.3 and 0.2, compared with 8.9 from pneumonia (Linder and Grove, 1947).

In contrast, puerperal sepsis accounted for 40% of maternal deaths in 1930 and maternal mortality was high, at almost 7 deaths per 1000 births. Furthermore, large absolute reductions in maternal mortality occurred with the arrival of sulfa drugs (Jayachandran et al., 2010; Thomasson and Treber, 2008). Falling maternal mortality could confound our pneumonia estimates in several ways. At the population level, parents may perceive greater returns to early life investments in girls when maternal mortality rates are lower, potentially leading to improvements in long-run outcomes for women vis-a-vis men (Jayachandran and Lleras-Muney, 2009). At the individual level, it is *a priori* plausible that there are consequences for boys and girls from one's own mother dying. There may be

additional impacts flowing from changes in family size that arise because of changes in the probability of infertility, a potential complication of post-partum sepsis[12]. To account for these possibilities, we therefore include the interaction $post_t$*$base\_maternal\ mortality_s$.

State economic and infrastructure characteristics: To further address pre-existing trends, we control for birth state-birth year varying socioeconomic variables including logged state income per capita, state per capita public health spending, and the numbers of schools, hospitals and physicians per capita.[13] The inclusion of income per capita in the state of birth is pertinent given changes in economic fortunes during the Depression Era, even if these have been shown to have had little, if any, impact on long-run health and economic outcomes (Cutler et al, 2007). To allow for any remaining converging/diverging trends at the birth state level, we investigate specifications with birth state specific time trends and census division*birth year fixed effects. These additional specifications are motivated in part by convergence in economic development between the US South (which was particularly plagued by pneumonia) and other parts of the country during the 20th century (Mitchener and MacLean, 1999).

We note that including this rich set of controls may amount to "over-controlling." For example, the increased risk of infection from weakened immune systems as well as competing risks from different conditions may create population level correlations in disease rates. Thus, controlling for additional diseases may capture variation in disease trajectories that are in fact driven by the use of sulfa drugs rather than by unobserved confounding factors. Similarly, controlling for state trends is typically quite demanding of the data. We therefore sequentially add in each set of controls when estimating our model.

Mortality selection: Survival selection will tend to bias our estimates downwards if frail children are more likely to succumb to pneumonia than their healthier counterparts (Almond 2006; Bozzoli et al., 2009). With the advent of sulfa drugs, more of these children will have survived childhood and these innately less healthy individuals are likely to be less productive and healthy as adults. We expect this selection bias to be small since pre-intervention death rates among infants was nationwide was 0.7% (~7 in 1,000 live births, see *Figure 3, Panel B*).

C.3. Age of exposure

---

[12] Any such effects are necessarily indirect because maternal post-partum infections (puerperal sepsis) that were controlled with sulfa drugs could not be transmitted to infants.

[13] The estimates are not sensitive to whether we directly include the characteristic (X_{st}) or we include *post-37* multiplied by the pre-intervention level of the characteristic ($post_t$*X_s) to allow for discontinuous effects of state per capita income that may otherwise load on to the variable of interest.

In an extension of the baseline specification which models birth year exposure, we allow for returns to sulfa exposure at other ages, as well. Though far lower, pneumonia mortality rates were still considerable for those aged 1-5 (*Figures 1* and *3*).[14] To investigate whether exposure to pneumonia at other ages had impacts, we replace *post* with a vector of birth year dummies for every year in the sample and then graph the coefficients on $birth\ year_t$*$base\_pneumonia_s$. As we discuss later, this specification also provides a useful falsification test: the presence of sharp breaks in the coefficients in years other than around 1937 may suggest that the patterns seen in the data are due to some process other than the introduction of sulfa drugs.

D. Results

D.1. Main results

Refer to *Table 1*, where each cell reports estimates of the coefficient on $post_t$*$base\_pneumonia_s$ from a separate regression. The rows denote the outcome variables and the columns different sets of control variables. We find that exposure to sulfa drugs, and thereby, reduced exposure to pneumonia led to statistically significant improvements in the expected direction for every outcome for men. These results are remarkably robust to successive inclusion of controls for other diseases, birth state socioeconomic and infrastructure characteristics, birth state specific linear time trends, and census division of birth*birth year fixed effects, allaying any potential concerns that they are driven by some unobserved convergence process.

Using the specification in column 4 of *Table 1*, we estimate that shifting the pre-sulfa pneumonia death rate (*base_pneumonia*) from the 75th to the 25th percentile of its distribution (i.e., from 1.18 to 0.92 deaths per 1,000) results in the following impacts for men[15]: a 0.15 increase in years of schooling, a 1.28% point increase in the probability of completing high school, and a 1.69% point increase in the probability of completing college. The same change in *base* led to a 2.11% increase in family income, a 0.72 % point decrease in the probability of being below 200% of the federal poverty line, a 0.71 % point increase in the probability of being employed, and a 0.26% point decrease in the probability of reporting a disability limiting or preventing work (this last impact is

[14] In contrast to Almond (2006) and Kelly (2011), we do not expect impacts from fetal exposure because mothers of childbearing age experienced very low infection rates (Britten, 1942).

[15] This is a change of 0.26 deaths per 1000, similar to the change in pneumonia rates after 1937 for a state at the 75th percentile of the baseline (pre-1937) pneumonia mortality distribution. The average pre-post drop in pneumonia was, as it happens, similar, at 23% for the all-age population (for our sample: see *Appendix Table A1*) and 30% for infants (section 2). A one *standard deviation* change in pre-intervention pneumonia is 0.19 and could be plugged in instead.

twice as large and significant in column 5). It is also associated with a 0.55% point decrease in the probability of reporting a cognitive disability. This coefficient is driven to insignificance with the inclusion of state controls but the point estimates in columns 4 and 5 remain large. The point estimates for physical disability are large but not statistically significant in any specification. While the sulfa birth cohorts married later, these effects are not statistically significant. Of note, all of these estimates represent averages for an entire birth state*birth cohort and can therefore be considered as intent-to-treat effects. In Section IV, we provide rough estimates of impacts on those who actually contracted childhood pneumonia. There we also compare our results to estimates of returns to other health and educational interventions.

For most outcomes, the estimated coefficients for women (*Table 1b*) are of a similar magnitude in the specifications with birth state and birth year fixed effects and control diseases (columns 1 and 2). However, they are diminished and driven to insignificance by the inclusion of state socioeconomic variables and state specific trends. We discuss these gender differences in further detail below.[16]

D.2. Age of exposure

Figures 5a and *5b* plot the coefficients on *birth year*base_pneumonia* for men and women, respectively. In line with the results in *Table 1* the coefficient series for men show trend breaks in 1937, consistent with the timing of the arrival of sulfa drugs but this is more ambiguous for women. The effect sizes generally increase over the period 1937-1939, consistent with the introduction of sulfapyridine, the sulfa agent that was more potent in treating pneumonia than its predecessors, in 1938 and its widespread adoption for pneumonia treatment by 1939. There is no evidence of trend breaks in years other than those associated with the introduction of sulfa. The coefficients prior to 1937 hover around zero indicating that the returns to sulfa drug exposure were likely limited to the birth year. Overlaid on the pneumonia coefficient plots are coefficients estimated from an identical equation for heart disease, a control disease on the premise that it was not directly influenced by the arrival of sulfa drugs. We find no evidence of a trend break in heart disease mortality (or for any of

[16] We do not find any evidence of long run outcomes responding to the sulfa-led decline in maternal mortality rates. The coefficient on *post*base_MMR* tends to take the unexpected sign and is generally insignificant for men and women. So, either reductions in maternal mortality did not change parental investments or any enhanced investments were offset by the pathways discussed in Section IIC. The coefficients on the vector of sulfa-untreatable diseases (TB, diarrhea, malaria etc) similarly tend to be insignificant or of the unexpected sign (results available on request). These findings reinforce our claim that the estimated impacts of *post*base_pneumonia* arise from the introduction of sulfa drugs.

the other control diseases – available upon request), which provides evidence against the role of other economic or health system factors behind our main results.

The coefficient plots suggest that the birth year is a "sensitive period" for interventions designed to control infectious diseases and in particular pneumonia. As discussed earlier, this is consistent with (a) pneumonia morbidity and mortality rates being far higher for infants than for older children and (b) the developmental changes induced by infectious disease being greater the younger the child is at exposure.

D.3. Additional Robustness Checks

Collectively, we have so far shown that the estimates for men are robust to the inclusion of a variety of birth state*birth year disease environment and macroeconomic controls, as well as flexible birth region trends and linear birth state-specific trends. The flexible birth cohort specific coefficient plots presented above further add to our confidence that our results are not driven by unobserved birth state*birth year factors. In this section, we consider a number of additional robustness checks.

Estimation sample*:* Here we test the sensitivity of our findings to meaningful changes in the estimation sample. Similar to Jayachandran et al (2010), we re-estimate our model removing the 1936 and 1937 birth cohorts as there was a minor influenza epidemic during these years. The point estimates are actually somewhat larger, with the qualitative conclusions remain similar (*Table 2, Panel A*). In *Panel B* of *Table 2* we present estimates using only cohorts born between 1935 and 1941 (which excludes World War II birth cohorts). Again, our core findings generally remain robust to this specification change. We also re-estimated the equations excluding the 2000 Census (for reasons discussed in the **Data Appendix)** and the estimates were similar (available upon request).

Placebo and triple difference specification: We re-estimate our core model with a placebo intervention, defining post as 1934 rather than as 1937 onwards. Consistent with the coefficient plots, for no outcome do we find any evidence of long-run gains (*Table 2, Panel C*).

We also take advantage of age-specific incidence rates for pneumonia in a triple difference setup. As discussed above, infection rates and days of disability suffered by pneumonia patients in the pre-sulfa era (1934-36) peaked at birth and reached a minimum in the age group 10-24. The difference was a factor of six to ten (*Figure 1*). Consequently, we expect limited if any long run benefits of sulfa drugs to be evident for individuals who were first exposed when at lowest risk of contracting pneumonia (ages 10-15) compared with individuals exposed at birth. This insight forms the basis of our triple difference strategy. For this we extended the sample back to birth cohorts 1915 and onwards. We define *treated* as 1 for individuals who were exposed to sulfa at birth (i.e. born

in or after 1937) and as 0 for individuals who had their first exposure to sulfa at age 10-15. *Post* is now equal to 1 for treated individuals born on or after 1937 and untreated individuals who were 10-15 on or after 1937 (*Appendix Figure A1* provides a visual representation of this set up). The regressor of interest is now *post*treated*base_pneumonia* and we naturally include the underlying two-way interactions and main effects, additionally including birth state and birth year fixed effects, the control diseases, and birth state specific linear time trends. The estimates are very similar in magnitude and significance to those in *Table 1,* the only noteworthy difference being that we now see employment gains for women born after the introduction of sulfa drugs *(Table 2, Panel D).*

Selective migration*:* Since the US census records the state of birth we are confident that we correctly match outcomes to pneumonia mortality rates faced during infancy[17]. Here, we investigate potential endogeneity of migration of *parents* of children born in our sample in 1930-43, a possibility neglected in previous studies of the long reach of the birth year environment. This is relevant insofar as parents select the birth state of their children. In general, positively selected potential parents may have chosen to move to low pneumonia states in order improve the life chances of their births.[18] On its own, this creates a compositional effect at baseline that reinforces the tendency for low pneumonia states to have better outcomes, a possibility captured by state fixed effects. The introduction of sulfa drugs, on the other hand, by virtue of narrowing state differentials in pneumonia, will have weakened disease-led migration. As a result, the relative improvement in long run outcomes in low pneumonia states will be smaller after 1937. This raises the possibility that some of the post-sulfa convergence across states that we attribute to improvements in early life health (in part) reflects migration-led changes in population composition.

We investigate this possibility using information on 20-40 year olds in the 1930 and 1940 census files (i.e., the population group most likely to give birth during the sample period).[19] We

---

[17] Regardless, we estimated models that control for the current state of residence in addition to the birth state, and where we only focus on those who moved to states other than their birth state. In neither case were there substantive changes to our results (available on request). These checks also indicate that the impacts we identify are not driven by post-birth migration in the direction of better opportunities.

[18] Montalvo and Reynal-Querol (2007) analyze the *reverse* process, how levels of infectious diseases respond to migration patterns. The possibility that migration may respond to the infectious disease environment has been discussed by Mesnard and Seabright (2009) but we are not aware of many tests of this.

[19] We constructed a state level data set with the number of 20-40 year olds in each state in the years 1930, 1935 and 1940 (the 1940 census asks individuals to report their state of residence five years before enumeration, hence the 1935 data point). Selecting age 20-40 is appropriate because we are interested in endogeneity of the birth state of our cohorts born in the 1930s. As birth and death create changes in population that are concentrated at the two ends of the age distribution, looking at changes in the population aged 20-40 is expected to isolate the effects migration.

regress log population on *post*base_pneumonia*, and state and year fixed effects. The results are in *Table 3*. Here we see the expected weakening of migration along a state-pneumonia gradient: The positive coefficient in column 1 indicates that, post-37, higher pneumonia states had larger populations, consistent with smaller outflows from these states following sulfa-led convergence in disease levels. However, this effect is obliterated upon controlling for state income per capita at baseline (which our models control for), suggesting that migration induced by changes in the disease environment is not driving our results.

D.4. Gender Differences

We find larger impacts for men vis-à-vis women, similar to several previous studies (Banerjee et al, 2010; Crimmins and Finch, 2006; Cutler, et al, 2010; Lindeboom et al, 2010; Stein et al, 2005; Venkataramani, 2012). One possible explanation for this is that boys are biologically more sensitive to the early life health environment than girls (Gluckman and Hanson, 2005; Low, 2000; Waldron, 1983). In the case of pneumonia, evidence of this comes from the fact that pneumonia morbidity was 71% higher for boys than girls in national study conducted in 1928-29, based on case rates of 2.4% versus 1.4%, respectively (Britten, 1942). Therefore, young boys were likely to gain more from the arrival of new antibiotics than young girls as far as improvements to their early life endowments, which could explain why the former experienced long-run impacts that were much larger in magnitude.

Another possible explanation is that women faced lower returns to human capital, which led to reduced investment in young girls and, as a result, depressed long-run outcomes[20]. One piece of evidence that this explanation is unlikely to fully explain our results given that the average years of schooling, high school, and college attendance are quite similar for men and women in our sample.

[20] In their study on the long-run effects of malaria eradication in India, Cutler, et al (2010) propose that the reason they only find long-run impacts for men is because they far more likely than women to participate in the wage labor market, where productivity differences flowing from improved early childhood health are more easily observed than with non-market activities. Thus, focusing on effects on labor market-based outcomes may not reflect the full set of benefits accruing to women. In addition, they consider (and reject) the hypothesis that differential household composition may play a role as well, in that treated men marry younger women and treated women may marry older (and thereby untreated) men. As a result, the households of treated men will have more post-eradication born individuals than the households of treated women.
Neither of these explanations, however, likely applies to our data. First, over 50% of women in our sample participated in the labor market at the time of census enumeration (compared to 67% of men), and schooling attainments among men and women in the sample are quite similar. Second, we do not find any impacts on measures of disability for women, which are not tied directly to labor market outcomes and some component of which should be independent of household composition effects if they truly reflect health endowments.

To investigate this further we constructed, for each birth state, the aggregate female and male-female ratios for log family income, employment, education, and the estimated return to education for individuals who were 20-45 years of age in the 1940 census, around the time that our sample cohorts were born/entering school. We then interacted these variables (one at a time) with the treatment term, *post*base_pneumonia*. Similar to Venkataramani (2012), who also utilizes this strategy, we find no evidence that the long run returns to the sulfa health intervention in infancy were, for women, increasing in their absolute or relative socioeconomic opportunities (*Appendix Table A2*).

Note that this investment-based explanation is similar in spirit to the notion of complementarity that we explore in the next section. Indeed, we utilize a similar empirical strategy below when we assess whether the effect of early life investments varied by barriers to human capital accumulation faced by blacks. While we cannot definitively rule out the possibility of reduced investments in generating the gender differences we observe, the potential statistical power of our empirical approach is highlighted by the results in the next section. Ultimately, examining potential biological and/or behavioral explanations behind gender differences in long-run impacts of early life shocks will require richer, more detailed data, on available opportunities and investments than we have at our disposal.

III. Gradients in the Long-Run Impacts of Pneumonia: Institutional Segregation, Human Capital Investments, and Complementarity

In this section, we utilize cross-state variation in institutionalized black-white segregation to identify the role of subsequent investments in driving the link between early life endowments and adult socioeconomic status. We start by describing the institutional setting, data, and methods, and then move to our main results and specification checks.

A. Institutional Setting, Data, and Basic Research Strategy

It is well known that segregation-era policies in the American South limited the upward mobility of blacks via reducing access and the returns to vehicles of human capital investment (Aaronson and Mazumder, 2011; Card and Krueger, 1992; Chay et al., 2009; Donohue and Heckman, 1991; Smith, 1984; Welch, 1974). These policies had deep historical roots: following the Jim Crow Laws in the post-Civil War Reconstruction era, racial segregation became *de jure* in all public facilities in the US South from the late nineteenth century onward. While the mandate proposed "separate but equal" status for black Americans, in practice it systemized their economic

and social disadvantage. Segregation outside the South (henceforth "North" for expositional ease) was *de facto* and weaker. A North-South gradient in school access, school quality, opportunities and rates of return for blacks has been documented (e.g. Margo, 1990, Donohue and Heckman, 1991), and is reflected in racial gaps in education among adults in the 1940 US Census (*Table 4*).

Our strategy is to utilize this historically-determined variation in the cost of acquiring human capital across race to examine whether investments made after infancy are required to realize the full potential of improved early childhood endowments. Specifically, we estimate versions of the following specification:

$$Y_{stc} = \alpha + \beta_1 * post_t * base\_pneumonia_s * segregation_s + \beta_2 * post_t * base\_pneumonia_s + \beta_3 * post_t * segregation_s + \boldsymbol{\gamma} * \boldsymbol{X}_{s(t)} + \theta_s + \eta_t + \theta_s * \eta_t + \lambda_c + e_{stc}$$

where Y_{stc}, $post_t$, $base\_pneumonia_s$, and the fixed effects denoted by the Greek letters are the same as defined in Section IIC.1 above, $\boldsymbol{X}_{s(t)}$ is the vector of birth state-level disease, infrastructure, and socioeconomic controls, and $\theta_s * \eta_t$ reflects birth state specific linear time trends. For $base\_pneumonia_s$, we made the *a priori* choice of using race-averaged baseline mortality rates (which we employ in Section II) given known concerns about the comparability of race-specific mortality data across states prior to the 1940s (see Ewbank, 1987; we further discuss this choice in the **Data Appendix**). The new variable here is $segregation_s$, which represents an indicator for the presence or severity of institutionalized racial segregation in the birth state. This model is set up to examine, via the triple interaction term, whether the long-run effects of exposure to sulfa drugs varies by the nature of the institutional climate, our proxy for the net cost or benefit of investing in human capital. As we discuss below, we estimate this model separately for blacks and whites given cross-race variation in harm from segregation policies.

We use three different measures for $segregation_s$. The first is a simple indicator for birth in the US South (see **Data Appendix** for a definition of the states we include in the South), which accounts for the fact that *de jure* segregation was present in this region but not elsewhere. In order to more definitively isolate the role of race-specific institutions, we interacted $post\text{-}1937_t * base\_pneumonia_s$ with the share of the state population that were slaves in 1860 (Nunn, 2008). This is also a more continuous measure of racial segregation, varying substantially within the South while being zero in the North (*Table 4*). Our use of this variable is motivated by earlier work showing that states with a stronger slave history continued to possess extractive institutions that more strongly limited

opportunities for blacks decades later. For instance, such states had poorer quality public schools for blacks and later suffrage (Mariscal and Sokoloff, 2000; Engerman and Sokoloff, 2005). Consistent with this, the slave share in 1860 has also been shown to be predictive of race gaps in education (Bertocchi and Dimico, forthcoming; Sacerdote, 2005) and labor market productivity (Mitchener and MacLean, 2003) more than a half-century later.

Third, we examine interactions of *post*base_pneumonia* with state-level ratios of schooling attainments of young (20-40 year old) black men relative to white men calculated from 1940 census microdata (see the **Data Appendix**). Whereas the North/South and slave history variables capture institutional heterogeneity, the black-white schooling ratio serve as a direct indicator of educational choices made in response to these extractive institutions.

B. Threats to Inference

There are several threats to inference and interpretation with respect to our research strategy. Though we address these in further detail below, two deserve mention at the outset. First, our empirical strategy hinges critically on blacks having access to sulfa drugs. If the arrival of sulfa drugs did not lead to reduced pneumonia exposure among blacks because of poor access, particularly in the Southern states, then the statistical test of complementarity we outline above is a non-starter. Given the extent of institutionalized racism in the US South, one could easily imagine blacks having either delayed and/or poor access to new medical technology.

It is worth addressing this concern before proceeding further. To investigate whether blacks benefitted less (if at all) from the arrival of sulfa drugs, we examine the absolute change in pneumonia mortality after 1937 using data for the Southern states, where we would expect access issue to be most severe. Trends in pneumonia mortality rates for blacks and whites are presented in *Figure 6*. These data show, if anything, *sharper* drops in 1937 for the rates for blacks. Moreover, the extent of absolute cross-state convergence in pneumonia mortality after the arrival of sulfa drugs in the US South was larger for blacks than whites, particularly infants (*Figure 7*). Regressions confirm that the post-1937 absolute trend break is significantly larger in magnitude for blacks (*Appendix Table A3*), as is the extent of absolute convergence in pneumonia mortality across the US states after 1937.[21]

[21] *Appendix Table A3* displays estimates from race-specific models regressing the state and *race*-specific pneumonia mortality rates (in levels and logs) on *post*year*, controlling for the main effects of *post* and *year* and for state fixed-effects. This approach follows that used in Jayachandran, et al (2010). The coefficients on

Of note, these conclusions differ from Jayachandran, et al (2010), who argue that blacks may have been less likely to use sulfa drugs than whites due to either lack of access to hospitals and health providers (Almond and Chay 2006) or slower adoption of new drugs due to lower education (Glied and Lleras-Muney, 2008). Differences in approach including different sample windows, the use of *absolute* versus relative mortality rates, and a single versus double difference models largely explain these differential conclusions.[22] Moreover, blacks being able to access sulfa drugs to this extent is entirely plausible: although they experienced more limited access to good hospital care, sulfa drugs were initially available without prescription and widely used in the outpatient setting.[23] In addition, as we discuss in Section IIA, the cost of a course of treatment with this life-saving drug was relatively low, reducing the likelihood that poor blacks faced financial barriers to obtaining treatment.

Second, it may be that gradients in long-run impacts by indicators of racial segregation may actually reflect the role of other state socioeconomic characteristics rather than complementarity in early and late investments. While discontinuous time varying trends in the outcomes as a function of baseline institutional measures are accounted for by the $post_t$*$segregation_s$ term, cross state convergence in outcomes by the state specific trends, and fixed factors that may be correlated with institutions in the cross-section by the birth state fixed effects, the triple interaction term may still be contaminated by other social or economic processes that introduce gradients in the long-run effects of early life health. We argue that estimating our models separately by race helps address this possibility, as segregation-driven constraints to subsequent human capital investment only applied to blacks. Observing gradients for whites would therefore suggest that our results are driven by some alternate

---

*post*year* for the models examining mortality rate in levels indicates a significant trend break in 1937 for both races, which is larger in magnitude for blacks. To assess post-sulfa convergence in pneumonia across the states we regressed the level of the pneumonia mortality rate on *post*base*, defined as in model (1). The coefficient on *post*base* is -0.56 (s.e. 0.111) for blacks and -0.27 (s.e. 0.185) for whites (the convergence estimates are available upon request; a graphical representation, albeit for the South, is in *Figure 7*).

[22] Using state data for 1925-1943 and tuberculosis as a control disease, Jayachandran et al. (2010) estimate that the proportional change in mortality from sulfa-treatable diseases in 1937 relative to TB is larger for whites than for blacks. Their estimates do show that, among the treatable diseases (pneumonia, maternal mortality, scarlet fever) the white advantage is smallest for pneumonia. Our specification indicates a black advantage (i.e. stronger trend break) in pneumonia reduction. The difference between these two sets of estimates arises from our using the mortality level rather than its log, our focus on the 1930-1943 period, and our estimating the break for tuberculosis independently so as to avoid the assumption of common trends. See *Appendix Table A3*.

[23] Interestingly pediatric patient records from the Sydenham Hospital in Baltimore, Maryland, show that blacks and whites had equal access to sulfa drugs during the study era. However this particular Southern hospital may have been unusual in that its pediatric wards were racially integrated by the mid 1920s (Connolly, et al, 2012)

socioeconomic process[24]. Put differently, we appeal to cross-racial, in addition to cross-state, differences in the effective returns to human capital to achieve identification.

C. Results

To fix ideas, we start by estimating models similar to those presented in *Table 2,* but this time separately by race. As seen in *Table 5*, the results for white men mirror the coefficients for the full male sample presented in *Table 1*[25]. However, the sample for black men shows strikingly different coefficients. In particular, we find either zero or impacts that are not robust to specification for blacks vis-à-vis whites, although they do exhibit relatively larger reductions in the risks of cognitive disability and poverty.[26] Estimates from fully interacted models show that the black-white differences in schooling impacts are statistically significant, though, despite differences that are often large in magnitude, we unable to reject the null of equality for the other outcomes. As we find no robust impacts for white or black women (*Appendix Table A4*), we hereafter focus our solely on the male sample.

Table 6a presents the results for the models incorporating cross-state differences in institutionalized segregation for blacks. Collectively, the results provide a compelling explanation for the black-white differences noted in *Table 5*. The first panel of *Table 6a* examines gradients in sulfa drug impacts across birth in or outside the US South. For this specification, note that the coefficient on $post\text{-}1937_t*base\_pneumonia_s$ is now implicitly the coefficient for cohorts born in the North and the coefficient on $post\text{-}1937_t*base\_pneumonia*South$ is the difference in impact for Southern born cohorts relative to Northern cohorts.

In general, across the range of outcomes, the coefficient on $post\text{-}1937_t*base\_pneumonia*South$ is negative (positive for the "undesirable" outcomes), though the impact of reduced pneumonia exposure on poverty appears to be larger in the South. For instance, we observe a significant positive coefficient for black men in the North on income of 0.42. The interaction with South takes a significant coefficient of -0.38, indicating a complete erosion of these gains in the South. Put differently, moving from the 75th to 25th percentile of the pre-sulfa pneumonia mortality distribution is associated with a 10% increase in log income for Northern born black men but only a 1% increase

[24] This would not necessarily hurt our case if the reason for this is related to investment responses across races. However, we cannot be sure of this.

[25] For comparability purposes, we restrict the analysis to those born in those birth states where black and white births were observed for each birth year in our sample window.

[26] While the coefficient on cognitive disability loses significance upon the addition of controls for state trends, it maintains its magnitude and is consistently larger than the corresponding estimate for whites.

black men born in the south. Similarly, Northern blacks born in and after 1937 show an increase in the probability of completing college of 7% points but the corresponding probability for Southern blacks is almost exactly zero. Looking across the set of nine outcomes, several of the triple interactions are statistically significant, which is impressive given the demands that the triple interaction term, state level controls, and birth state specific time trends impose on the data and the fact that North-South is a crude measure of institutional racial disadvantage[27].

Panel B of *Table 6a* presents estimates of long-run impact gradients by the historical slave fraction variable. The pattern of results observed here is similar to those for the North-South variable. In particular, moving from the 75th to the 25th percentile of the nationwide pneumonia mortality distribution is associated with a 13% increase in income for black men born in states with no history of slavery. Increasing the slave fraction to the sample maximum (0.57) reduces the post-sulfa gain associated with the same reduction in pneumonia exposure to less than a 1% increase. This complete washout of impacts for those born in states with higher historical slave fractions is also seen for the education, employment, and work-related disability variables, with statistically significant gradients observed for the latter two.

Finally, this general pattern is borne out when considering gradients by birth state black-white educational attainment ratios, as well. Similar to the above, in *Table 6a, Panel C*, we find well-determined triple interactions with the race education ratio for income and employment, and the signs for the other outcomes are generally consistent with our story so far, showing stronger sulfa-related gains in states with a narrower race gap in education.[28]

---

[27] We extended this analysis by examining differences between the states in the Deep South (Alabama, Georgia, Louisiana, Mississippi, and South Carolina) and those in rest of the south, as segregation was more entrenched in the former. The estimates for income and employment show a consistent gradation of outcome returns with differences between the North and the Deep South being sharper than differences between the North and the entire South (results available upon request).

[28] We further estimated interactions with race-specific school quality using data for 18 Southern states (taken from Card and Krueger, 1992) for the years in which our sample cohorts were aged 6-16. The estimates show that high school and college completion rates for blacks were increasing in a quality score constructed from term length, the pupil-teacher ratio and teacher salaries, but the interaction terms are imprecisely estimated (results available on request). We also estimated interactions of *post*base_pneumonia* with the race differential in (log) income for adults in the 1940 census (see data appendix) and, again, we find negative coefficients on the interaction term for most outcomes and the interactions for college, employment and work-related disability are statistically significant (available on request). This is consistent with black investments in human capital being depressed by race differentials in school quality as well as employment segregation (Welch 1974; see section 1 above). Note that the marginal cohort, born 1937, was 27 when the Civil Rights Act of 1964 mandated equal employment opportunities.

In *Table 6b*, we repeat these analyses for white men. Crucially, for most outcomes, sulfa-led improvements for white men are not significantly different by measures of institutional segregation. Where there are significant differences differences, they are often in the *opposite* direction to that seen for blacks: for example, with the North-South and slave fraction indicators, we actually find increases employment for white men to be larger for those born in the South. This pattern establishes that impact gradients for blacks are unique and do not reflect a more general, pan-racial process that differentially drives the returns to early life outcomes.

D. Specification Checks

The differential results across race increase our confidence that the gradients we observe reflect the key role of subsequent investments in realizing the potential of improvement early life endowments. However, there remain alternate explanations for our findings, which we address in this subsection.

Measurement error: It is possible that the coefficients on the triple interaction term, which generally imply muted long-run impacts of infant health, simply reflect the fact that disease exposures in infancy were measured poorly for blacks born in the Southern states relative to Northern states. That is, one could be concerned that mortality and age reporting were less reliable in states with more extractive institutions.

Regarding the measurement error in mortality rates, one issue is that blacks constituted about 10% of the overall population and a much smaller proportion in some states during our study era. As a result, the race-averaged mortality rates we use may not accurately reflect the disease environment faced by blacks, with the resulting measurement error driving down the estimates on *post*base_pneumonia* for blacks more than for whites. If this process was worse for Southern-born cohorts, it could explain the fact that many of triple interaction terms tend towards null effects for these individuals.

However, measurement error is unlikely to be driving our findings for several reasons. First, we do find several economically and/or statistically significant effects for blacks in *Table 5* (the model without gradients estimated by race), suggesting that our race-averaged indicator is indeed picking up meaningful variation in the disease environment (we expound on this further in the **Data Appendix**). Moreover, not all of the coefficients on the triple interaction term in *Table 6* suggest a movement towards a null effect for Southern born cohorts. In particular, in most specifications we find evidence of *larger* reductions on the probability of living below the poverty line for those born

in states with (more severe) institutionalized segregation (though these differences are not statistically significant).

We also conduct an additional to address this measurement issue. First, we estimate models including interactions between *post-1937$_t$*base_pneumonia* and the number of years prior to the arrival of sulfa drugs that a given state became part of the national birth registration system. We think of this variable as a proxy for the quality of data collection systems, which may have improved the longer a given state was involved in collecting birth and death rates used for national measurements. As seen in *Table 7*, *Panel A*, our results remain unchanged even after controlling for this variable.

Age reporting tends to be larger at lower levels of education and for those without access to hospitals and other areas where demographic surveillance data is recorded and processed. In view of this, we investigated whether Southern blacks were more likely to show age heaping than whites. This is relevant as it would imply that treatment exposure, which is defined on birth year, is misspecified for this group, perhaps depressing the coefficient estimates.[29] *Figure 8* plots the distribution of birth year for the black sample. There is no evidence of age heaping, either in the North or South, for our post-1930 sample cohorts.

Decoupling of Morbidity and Mortality: We established in Section IIIA that blacks *did* benefit from sulfa drugs in terms of reductions on pneumonia mortality. However, it is still possible that there were barriers to utilizing sulfa drugs on the intensive margin. In particular, blacks may have been willing to pay for antibiotics to treat more severe, life-threatening illnesses, but may have deferred obtaining sulfa drugs with less-severe cases due to high transport and/or time costs of, or institutional barriers to, receiving outpatient medical therapy. Put differently, this implies that while mortality may have declined for blacks, improvements in morbidity (on average) may not have followed suit, particularly for those living in the US South. As such, our estimates of depressed impacts in the South could simply reflect antibiotic access issues on this more intensive margin rather than the failure to make complementary investments among children with better early life endowments.

We offer three pieces of evidence to suggest that this explanation likely does not account for the results in *Table 6*. First, as discussed above, the sign of the gradients we estimate generally suggest larger in magnitude impacts on the probability of living below the poverty line for those born in states with (more severe) institutionalized segregation. While not statistically significant, this

[29] We thank James Fenske for this suggestion amongst others.

finding suggests that Southern blacks did derive endowment benefits as a result of the arrival of sulfa drugs, though their ultimate impacts were more muted than blacks born in the North.

Second, we also estimate models where we include interactions between *post-1937$_t$*base_pneumonia$_s$* and the number of pharmacists per 1,000 black population residing in areas populated by blacks in each birth state as reported in the 1940 census (see **Data Appendix**; we include all of the relevant second-order interactions, as well). Pharmacists met the vast bulk of the demand for sulfa drugs during our study period, and were even able to sell the antibiotics without physician approval up through 1938 (Lerner, 1991; Lesch, 2007). Thus, the availability of pharmacists represents a relevant measure of the cost of obtaining sulfa drugs. If the gradients we observe in long-run impacts are actually driven by outpatient access to sulfa drugs - that is, the presence of segregationist institutions was associated with reduced access to pharmacists among blacks - then the inclusion of this new triple interaction term should reduce the magnitude of our estimates. As shown in *Table 7, Panel B,* our core findings remain robust to including this control.

Of course, the problem with this specification check is that the density of pharmacists in black populated areas need not reflect actual access. Our third piece of evidence abstracts from this possibility entirely by appealing to insights from biology. In particular, we take advantage of data on rheumatic fever, a disease of the heart, joints, and/or brain that can only occur after a preceding streptococcal bacteria infection, such as pharyngitis ("sore throat"), tonsillitis, or scarlet fever[30]. Rheumatic fever incidence and mortality rates declined throughout the 20th century, with the drop accelerated with the widespread arrival of penicillin in the mid 1940s (Bisno, 1990; Denny, 1994; Massell, et al, 1988). Given this, evidence of declining rheumatic fever rates among blacks in the penicillin-era would therefore support the notion that blacks were indeed accessing antibiotic therapy for the treatment of non-life threatening illness.

This is exactly what we find in the data. *Figure 9* plots rheumatic fever mortality data for blacks and whites between 1941 and 1951. For both blacks and whites we see an increasing rate of decline in rheumatic fever starting around 1944-1946, when penicillin started being mass-produced. Indeed, the mortality rates for blacks converge strongly towards those for whites in this series. As there were no known improvements supportive therapy or access to health care institutions that preferentially impacted blacks vis-à-vis whites during this time, we can attribute these declines to

[30] It is thought that antibodies produced following an infection caused by Group A streptococcal bacteria cross react with tissues around blood vessels, heart valves, and joints to cause rheumatic fever (Bisno, 1991).

increased rates of treatment of antecedent non-life threatening bacterial infections[31]. Collectively, the evidence appears to favor the contention that blacks received antibiotic treatment for infections, be they severe or not.

Differential survival selection: Given that black pneumonia mortality rates were higher in the pre-sulfa drug era and they experienced larger absolute declines in mortality post-sulfa, survival selection may have been greater among blacks. As the sulfa drug shock was positive, survivors will have been negatively selected which, in principle, may explain why we observe weaker economic gains in adulthood for blacks, particular for those born in the South. Our finding that it is blacks at the lower end of the outcome distribution that benefit (e.g. we observe reduced poverty risk without increases in average cohort income, an effect which appears to be larger in magnitude among those born in Southern states; see *Table 5* and *Table 6*) undermines this possibility since the lower part of the outcome distribution is where we would expect the impact of mortality selection to be most severe. This contention is further bolstered by recent research showing that the differences between survivors and non-survivors may have to be implausibly large in order to invoke mortality selection as an explanation for changing mean outcomes over time (Alderman, et al, 2012).

North-South migration: The Great Migration of blacks from the South to the North encompassed the period of this study, even if it slowed after 1930. Previous work shows that the relatively educated were more likely to migrate northward (Vigdor, 2002; Aaronson and Mazumder, 2011). If there were changes in North-South migration driven by economic opportunities that happened to coincide with the health shock in 1937, their impact will be absorbed by *post*south*. But what if migration out of the South was induced by its higher pneumonia levels? Prior to the introduction of sulfa, this will have enhanced divergence in disease levels between the South and the North and, in our specification, this will be captured in the coefficient on the two-way term, *base*south*. The introduction of sulfa stimulates a tendency towards convergence in disease levels which, given higher initial disease burdens in the South, implies South-North convergence. In our specification this would show as a positive coefficient on *post*base*south*, but what we find is a

[31] It has been suggested that the arrival of penicillin cannot fully account for the decline in rheumatic fever since a nontrivial number of cases were due to very mild streptococcal infections that escaped medical attention (Massell, 1988). In lieu of this, a mechanism of penicillin reducing the (mean) ability of the bacterial population to cause rheumatic fever has been proposed. Relevant to our discussion, if this is a dominant mechanism, one might imagine that it would be possible to treat one segment of the population (whites) causing spillover effects to the other (blacks). We view this as unlikely given that contraction of scarlet fever or streptococcal pharyngitis involves relatively close contact to the infected party (skin contact, inhalation of aerosolized particles, etc) - this is precisely what segregation polices sought to avoid, at least in the public sphere.

negative coefficient. Thus, since our case rests upon a North-South divergence in sulfa drug led gains, accounting for North-South migration would only strengthen our results.

We nevertheless extended the migration equation discussed in Section 5 to look specifically at black migration from the South to the North and found that, conditional upon income, there is no evidence of endogenous migration (*Appendix Table A5*).

Race differences in child labor: Post-1937 improvements in child health may have raised the returns to child, adolescent, and young adult labor alongside raising the returns to schooling (Bleakley, 2010b; Venkataramani, 2012). This is especially pertinent for Southern blacks, who lived in predominantly rural areas and for whom child labor laws had no significant impact on educational attainment, presumably because of a paucity of black schools and, related, states being more likely to exempt black children from these laws (Lleras-Muney, 2002). As such, another explanation for the results in *Table 5* could be that blacks chose to invest their child's time in labor over schooling in response to the positive health shock. To the extent that they did, this choice will have reflected the relatively low returns to schooling for blacks in this era. So it remains consistent with our hypothesis that institutionalized segregation explains the race gaps in socioeconomic returns flowing from a reduction in pneumonia..

IV. Discussion and Policy Implications

Our study makes two overarching contributions. First, we establish large impacts of infant exposure to pneumonia, a leading killer of children in the pre-antibiotic era US and in developing countries today, on indicators of socioeconomic status and disability in adulthood, exploiting the sharp reduction in pneumonia morbidity induced by the introduction of antibiotics in early twentieth century America. Second, using exploit variation in returns to human capital investment across race driven by institutionalized segregation, we provide evidence that gains flowing from reductions in infant pneumonia materialize to a lesser extent, if at all, if subsequent complementary investments are not made. In demonstrating this, we illustrate an additional, insidious mechanism by which racial segregation contributes to race inequality in the long run: by crippling the ability of better-endowed children to fully realize their potential.

Regarding the first point, it is important to understand the magnitude of the estimated impacts of reduced early life pneumonia as these have potential implications for child health and poverty alleviation efforts in the United States and elsewhere. We begin by noting that our estimates are effectively intent to treat impacts, which will understate the individual impact of birth year exposure to sulfa drug driven reductions in pneumonia morbidity because not all members of a

cohort were afflicted with pneumonia. To approximate average impacts for those who were, we adjust our estimates by the fraction of infants with pneumonia. As historical infection data, in the United States likely underestimates the true pneumonia morbidity burden, we take a case rate of 15 per 100 child-years, which approximates prevalence in today's developing countries. This would inflate the effect sizes by a factor of 6.7. For instance, the post-sulfa gain in income accrued by those afflicted by pneumonia from moving from a state at the 75th percentile of the pre-intervention pneumonia mortality rate distribution to a state at the 25th percentile is 14.1%; the equivalent gain for schooling would be 1.01 years. The fact that a spell of pneumonia led to median of 1 month of disability per patient in the year before sulfa drugs arrived in America (Britten, 1942) and that a child may have had recurrent spells underlines the plausibility of the severe scarring implied by these estimates.

These results are comparable in magnitude to recent estimates of the long-run impact of deworming or malaria eradication in developing countries.[32] These impacts are also similar to those flowing from interventions in the education sector.[33] However, it is important to look at benefits in relation to costs. As discussed above, a complete course of sulfa was between $28-$100 in 2008 dollars, varying with duration and dose. Consider a cost of $50. The income return at the cohort-level was 2.1%, which, applied to mean income for the sample cohorts (*Table 1*), implies a benefit-cost ratio of 16.6. Inflated by a pneumonia prevalence rate of 15, this implies a ratio of 111.2 for those who incurred the infection and were treated. We also computed the cost of an additional year of schooling so as to assess the effectiveness of the sulfa drug revolution relative to a variety of interventions evaluated in recent randomized control trials. The returns to pneumonia reduction via

[32] Malaria eradication is estimated to have led to a 15-27% increase in wage income (Bleakley, 2010) and about a 3-year increase in schooling (Lucas, 2010). Deworming in primary schools is estimated to have generated 2-3 additional years of schooling and a 21-29% increase in income (Baird et al., 2011).

[33] Aaranson and Mazumder (2011) show that the Rosenwald intervention, involving construction of black schools in the rural South with a coverage rate of 19.2%, led to an increase in schooling of 0.25 years (3.5% relative to a baseline of 7.1 years). Duflo (2001) estimates that for every additional primary school per 1000 children constructed in Indonesia in the mid-1970s, there was an increase in education of 0.12-0.19 years and an increase in wages of 1.5-2.7% for men. Our intent-to-treat estimates compare favorably to these intent-to-treat estimates: among men, the drop in pneumonia associated with moving from the 75th to the 25th percentile of the pre-sulfa distribution led, for post-sulfa cohorts, to an increase in schooling of 0.15 years (1.24% relative to a baseline of 12.1 years). The increase in wages, at 2.11%, is comparable to that flowing from school building in the cited studies The "units" of the interventions are not comparable but these figures are of interest as they compare program impacts for the sorts of programs that are currently being implemented in poorer countries.

antibiotics exceed those for merit scholarships, uniforms, and conditional cash transfers assessed in various settings.[34]

Certainly, these calculations are subject to a number of caveats. First, the cost and efficacy of antibiotics today are quite different than in the 1930s and 1940s. New agents, improvements in production processes, and improvements in ancillary care and health care resources all make life saving courses of antibiotics cheaper, although growing drug resistance may impact both these costs and the returns (see, for example, Garau, 2002). Second, the returns to pneumonia reduction in the United States (particularly from an intervention 80 years ago) need not apply to today's developing countries. Finally, our calculations do not include the returns from preventing infant deaths (longevity) or the direct value of better health, both of which are likely to be large (Murphy and Topel, 2005).

Regardless, the broad magnitude of these returns has implications for global health policy. For example, immunizations and antibiotics for pneumonia reach only a small fraction of infected infants (WHO 2011). This is despite the fact that interventions to reduce pneumonia morbidity and mortality have been shown to be successful in the developing world.[35] The Advanced Market Commitment (AMC) and the Health Impact Fund (HIF) are recent initiatives directed at garnering commitment of international resources to finance preventative and therapeutic innovations and their large-scale distribution in poor countries. The sustainability of these initiatives depends upon measurement of health impacts (Grootendorst, 2009). The standard approach in the global health literature is to look for immediate reductions in morbidity and mortality at the *patient* level, often summarized in quality or disability adjusted life years (for instance, see the contribution of Jamison et al. (2012) to the Copenhagen Consensus). This ignores the possibility that exposure to infectious disease influences human capital accumulation, with long run socioeconomic consequences, and it thereby leads to under-estimation of the returns to intervention. Accounting for dynamic gains in the health and socioeconomic domain at the population level helps construct an efficiency case for

---

[34] See http://aidthoughts.org/?p=1279, http://www.copenhagenconsensus.com/Default.aspx?ID=1632. The latter suggests a mean benefit-cost ratio of 9 for conditional cash transfers, although there is much variation.

[35] For example, a recent community-based intervention in India that involved treatment of pneumonia in children aged 0-4 years with co-trimoxazole showed a case-fatality rate of 0.8% in the treatment area compared with 13.5% in the control area. The cost of co-trimoxazole was US $0.025 per child per year or $2.64 per child saved (Bang et al., 1999). In addition, large-scale rollouts of pneumococcal vaccines have shown to be highly efficacious in reducing pneumonia incidence and mortality in developed and developing countries alike (Cutts et al., 2005).

investment in motivating new global health interventions that target young children (see Bhalotra and Pogge 2012).

The bulk of the empirical literature on the long-arm of early childhood stops here. The second contribution is that we go further and show that failure to make complementary investments in childhood and young adulthood can lead to much more muted gains from positive endowment shocks than those described above. From a policy perspective this suggests that early life investments need to be followed up with subsequent investments in order to be most potent. More specifically, in showing that institutional constraints that strongly limited investments in human capital, such as further schooling, lead to diminished impacts of positive infant health shocks for Southern blacks, our results are directly relevant to today's developing countries, where barriers to human capital accumulation (such as poor access to quality schools, imperfect information, and binding credit constraints) remain. Furthermore, these findings add yet another negative to the list of negative consequences that flow from extractive institutions (Acemoglu and Johnson, 2012). The long-reach of discriminatory policies via their ability to hold back the potential of talented individuals is yet another reason to lift these institutional or social constraints where they may exist.

A challenge going forward is to determine exactly which kinds of complementary investments matter in unlocking the potential of improved early life endowments. While we are able to provide evidence of complementarity that rests on ostensibly firmer identification than previous work, our results represent what are effectively "black box" estimates. That is, while we infer subsequent human capital investments are consequential by establishing gradients in long-run impacts across institutional settings that historically capped the opportunities of a disadvantaged population group, we do not observe these investments directly. Future work would do well to examine complementarity in the setting of discrete observed investments, though, as discussed previously, finding exploitable situations where both endowments and subsequent investments are exogenously assigned could provide difficult.

References

Aaronson, Daniel, and Bhashkar Mazumdar. "The Impact of Rosenwald Schools on Black Achievement." *Journal of Political Economy* 119, no. 5 (2011): 821-88.

Acemoglu, D, and JA Robinson. *Why Nations Fail: The Origins of Power, Poverty, and Prosperity*. New York: Crown Publishers, 2012.

Acemoglu, Daron, and Simon Johnson. "Disease and Development: The Effect of Life Expectancy on Economic Growth." *Journal of Political Economy* 115, no. 6 (2007): 925-65.

Adhvaryu, A, and A Nyshadham. "Endowments at Birth and Parents' Investments in Children." *Mimeo, Yale University* (2012).
Aizer, Anna, and Flavio Cunha. "The Production of Human Capital in Childhood: Endowments, Investments and Fertility." *Mimeo, Brown University* (2012).
Alderman, H, M Lokshin, and S Radyakin. "Tall Claims: Mortality Selection and the Height of Children in India." *Economics and Human Biology* 9, no. 4 (2012): 393-406.
Almond, Douglas, and Kenneth Y Chay. "The Long-Run and Intergenerational Impact of Poor Infant Health: Evidence From Cohorts Born During the Civil Rights Era." *mimeo, Columbia University* (2006).
Almond, Douglas, and Janet Currie. "Human Capital Development before Age 5." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David E Card: Elsevier, 2011.
Almond, Douglas, and Janet Currie. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (2011): 153-72.
Autor, David H, and Mark G Duggan. "The Rise in Disability Rolls and the Decline in Unemployment." *Quarterly Journal of Economics* 118, no. 1 (2003): 157-205.
Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. "Worms and Work: Long-Run Impacts of Child Health Gains." *mimeo, University of California, Berkeley* (2011).
Banerjee, Abhijit, Esther Dulfo, G Postel-Vinay, and T Watts. "Long-Run Health Impacts of Income Shocks: Wine and Phylloxera in Nineteenth Century France." *The Review of Economics and Statistics* 92, no. 4 (2010): 714-28.
Barker, DJP, and C Osmond. "Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales." *The Lancet* 321 (1986): 1077-81.
Barreca, A, P Fishback, and S Kantor. "Agricultural Policy, Migration, and Malaria in the United States in the 1930." *Explorations in Economic History* 49, no. 4 (2012): 381-98.
Barreca, Alan. "The Long-Term Economic Impact of *in Utero* and Postnatal Exposure to Malaria." *Journal of Human Resources* 45, no. 4 (2010): 865-92.
Becker, G, and N Tomes. "Child Endowments and the Quantity and Quality of Children." *Journal of Political Economy* 84, no. 4 (1976): S143-S62.
Bertocchi, G, and A Dimico. "The Racial Gap in Education and the Legacy of Slavery." *Journal of Comparative Economics* (Forthcoming).
Bertrand, M, E Dulfo, and S Mullainathan. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* 119 (2004): 249-75.
Bharadwaj, P, K Loken, and C Neilson. "Early Life Health Interventions and Academic Achievement." *American Economic Review* (Forthciming).
Bhutta, Zulfiqar A. "Dealing with Childhood Pneumonia in Developing Countries: How Can We Make a Difference." *Archives of Disease in Childhood* 92 (2007): 286-88.
Bisno, A. "The Resurgence of Rheumatic Fever in the United States." *Annual Reviews of Medicine* 41 (1990): 319-29.
Black, Robert E, Simon Cousens, Hope L Johnson, Joy E Lawn, Igor Rudan, Diego G Bassani, Prabhat Jha, Harry Campbell, Christa Fischer Walker, Richard Cibulskis, Thomas Eisele, Li Liu, and Colin Mathers. "Global, Regional, and National Causes of Child Mortality in 2008: A Systematic Analysis." *The Lancet* 375 (2010): 1969-87.
Bleakley, Hoyt. "Disease and Development: Evidence from Hookworm Eradication in the American South." *The Quarterly Journal of Economics* 122, no. 1 (2007): 73-117.
Bleakley, Hoyt. "Health, Human Capital and Development." *Annual Review of Economics* 2 (2010): 283-310.
Britten, Rollo H. "The Incidence of Pneumonia as Recorded in the National Health Survey." *Public Health Reports* 57, no. 40 (1942): 1479-94.

Card, David, and Alan Krueger. "School Quality and Black-White Relative Earnings: A Direct Assessment." *The Quarterly Journal of Economics* 107, no. 1 (1992): 151-200.
Case, A, A Fertig, and C Paxson. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24 (2005): 365-89.
Case, Anne, and Christina Paxson. "Early Life Health and Cognitive Function in Old Age." *American Economic Review: Papers and Proceedings* 99, no. 2 (2009): 104-09.
Chay, Kenneth Y, Jonathan Guryan, and Bhashkar Mazumder. "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon after Birth." *NBER Working Paper w15078* (2009).
Connolly, C, J Golden, and B Schneider. "A Startling New Chemotherapeutic Agent: Pediatric Infectious Disease and the Introduction of Sulfonamides at Baltimore's Sydenham Hospital." *Bulletin of the History of Medicine* 86 (2012): 66-93.
Crimmins, Eileen M, and Caleb E Finch. "Infection, Inflammation, Height, and Longevity." *Proceedings of the National Academy of Sciences* 103, no. 2 (2006): 498-503.
Cunha, F, and J Heckman. "The Technology of Skill Formation." *American Economic Review* 97, no. 2 (2007): 31-47.
Cunha, Flavio, James Heckman, and Susan Schennach. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78, no. 3 (2010): 883-931.
Currie, J, and E Moretti. "Biology as Destiny? Short- and Long-Run Determinants of Intergenerational Tranmission of Birth Weight." *Journal of Labor Economics* 25, no. 2 (2007): 231-64.
Cutler, David, Winnie Fung, Michael Kremer, Monica Singhal, and Tom Vogl. "Early-Life Malaria Exposure and Adult Outcomes: Evidence from Malaria Eradication in India." *American Economic Journal: Applied Economics* 2 (2010): 72-94.
Denny Jr, FW. "A 45-Year Perspective on the Streptococcus and Rheumatic Fever: The Edward H. Kass Lecture in Infectious Disease History." *Clinical Infectious Diseases* 19, no. 6 (1994): 1110-22.
Donohue, John J, and James J Heckman. "Continuous Versus Episodic Change: The Impact of Civil Rights Policy on the Economic Status of Blacks." *Journal of Economic Literature* 29, no. 4 (1991): 1603-43.
Dulfo, E. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment " *American Economic Review* 91, no. 4 (2001): 795-813.
Engerman, Stanley L, and Kenneth L Sokoloff. " The Evolution of Suffrage in the Americas." *Journal of Economic History* 65 (2005): 891-921.
Eppig, Christopher, Corey L Fincher, and Randy Thornhill. "Parasite Prevalence and the Worldwide Distribution of Cognitive Ability." *Proceedings of the Royal Society B* 277, no. 1701 (2010): 3801-8.
Evans, G.M., and Wilfrid F Gaisford. "Treatment of Pneumonia with 2-(P-Aminobenzenesulphonamido) Pyridine." *The Lancet Infectious Diseases* 232, no. 5992 (1938): 14-19.
Ewbank, Douglas C. "History of Black Mortality and Health before 1940." *The Milbank Quarterly* 65, no. Suppl 1 (1987): 100-28.
Farba, T, and AH Thomson. "The Burden of Pneumonia in Children in the Developed World." *Pediatric Respiratory Review* 6, no. 2 (2005): 76-82.
Gaisford, Wilfrid F. "Results of the Treatment of 400 Cases of Lobar Pneumonia with M&B 693." *Proceedings of the Royal Society of Medicine* 32, no. 9 (1939): 1070-76.

Gibberd, G F. "Prontosil and Similar Compounds in the Treatment of Puerperal Haemolytic Streptococcus Infections." *British Medical Journal* 2 (1937): 695-98.
Goldin, Claudia, and Lawrence Katz. *The Race between Education and Technology*. Cambridge: The Belknap Press of Harvard University Press, 2008.
Greengard, Joseph, William B Raycraft, and William G Motel. "Effects of Chemotherapy on Pneumonia in Infants under One Year of Age." *American Journal of Diseases of Children* 62, no. 730-742 (1943).
Grove, Robert D., and Alice M Hetzel. *Vital Statistics Rates in the United States 1940-1960*. Washington, DC: United States Government Printing Office, 1968.
Hodes, Horace L, William C Stifler, Ethel Walker, Maclyn McCarty, and Robert G Shirley. "The Use of Sulfapyridine in Primary Pneumococcic Pneumonia and in Pneumoccic Pneumonia Associated with Measles." *The Journal of Pediatrics* 14, no. 4 (1939): 417-46.
Jayachandran, Seema, and Adriana Lleras-Muney. "Life Expectancy and Human Capital Investments: Evidence from Maternal Mortality Declines." *The Quarterly Journal of Economics* 124, no. 1 (2009): 349-97.
Jayachandran, Seema, Adriana Lleras-Muney, and Kimberly V Smith. "Modern Medicine and the 20th-Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs." *American Economic Journal: Applied Economics* 2, no. 2 (2010): 118-46.
Johnson, R. "Long-Run Impact of School Desegregation and School Quality on Adult Attainments." *NBER Working Paper No. 16664* (2011).
Johnston, IDA, DP Strachan, and R Anderson. "Effect of Pneumonia and Whooping Cough in Childhood on Adult Lung Function." *The New England Journal of Medicine* 338, no. 9 (1998): 581-87.
Kelly, Elaine. "The Source of Asian Flu: *In Utero* Exposure to Pandemic Influenza and the Development of a Cohort of British Children." *Journal of Human Resources* 46, no. 4 (2011): 669-94.
Kiefer, David M. "Miracle Medicines: The Advent of the Sulfa Drugs in the Mid-1930s Gave Physicians a Powerful Weapon." *Today's Chemist at Work* 10 no. 6 (2001): 59-60.
Klugman, Keith P, and Charles Feldman. "Pneumococcal Infections." In *Bacterial Infections of Humans: Epidemiology and Control*, edited by A.S. Evans and P.S. Brachman. New York: Springer, 2009.
Lerner, BH. "Scientific Evidence Versus Therapeutic Demand: The Introduction of the Sulfonamides Revisited." *Annals of Internal Medicine* 115, no. 4 (1991): 315-20.
Lesch, John E. *The First Miracle Drugs: How the Sulfa Drugs Transformed Medicine*. New York, NY: Oxford University Press, 2007.
Lindeboom, Martin, F Portrait, and G.J. van den Berg. "Long-Run Effects on Longevity of a Nutritional Shock Early in Life: The Dutch Potato Famine of 1846-1847." *Journal of Health Economics* 29, no. 5 (2010): 617-29.
Linder, Forrest E., and Robert D. Grove. *Vital Statistics in the United States 1900-1940*. Washington, DC: United States Government Printing Office, 1947.
Long, Perrin H, and Eleanor A Bliss. "Observations Upon the Experimental and Clinical Use of Sulphanilamode in the Treatment of Certain Infections." *Journal of the American Medical Association* 108 (1937): 32-37.
Low, Bobbi S. *Why Sex Matters : A Darwinian Look at Human Behavior*. Princeton, NJ: Princeton University Press, 2000.
Lucas, Adrienne. "Malaria Eradication and Educational Attainment: Evidence from Paraguay and Sri Lanka." *American Economic Journal: Applied Economics* 2 (2010): 46-71.
Margo, RA. *Race and Schooling in the South: 1880-1950*. Chicago: University of Chicago Press, 1990.

Mariscal, Elisa, and Kenneth L Sokoloff. "Schooling, Suffrage, and the Persistence of Inequality in the Americas, 1800-1945." In *Political Institutions and Economic Growtn in Latin America: Essays in Policy, History, and Political Economy*, edited by Stephen Haber. Stanford: Hoover Institution Press, 2000.

Massell, C, CG Chute, AM Walker, and GS Kurland. "Penicillin and the Marked Decrease in Morbidity and Mortality from Rheumatic Fever in the United States." *New England Journal of Medicine* 318, no. 5 (1988): 280-86.

Mitchener, Kris James, and Ian W MacLean. "U.S. Regional Growth and Convergence, 1880-1980." *Journal of Economic History* 59, no. 4 (1999): 1016-42.

Moody, E Earl, and E.G. Knouf. "Pneumonia in Children: Treatment with Sulfapyridine." *California and Western Medicine* 53, no. 3 (1940): 116-23.

Neal, Derek, ed. *Why Has Black-White Skill Convergence Stopped?* . Edited by Eric Hanushek and Finis Welch, *Handbook of Economics of Education*. Amsterdam: Elsevier, 2006.

Rudan, Igor, Lana Tomaskovic, Cynthia Boschi-Pinto, and Harry Campbell. "Global Estimate of the Incidence of Clinical Pneumonia among Children under Five Years of Age." *Bulletin of the World Health Organization* 82 no. 12 (2004): 895-903.

Ruggles, Steven, J Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B Schroeder, and Matthew Sobek. "Intergrated Public Use Microdata Series: Version 5.0 [Machine-Readable Database]." Minneapolis: University of Minnesota, 2010.

Sacerdote, Bruce. "Slavery and the Intergenerational Transmission of Human Capital." *The Review of Economics and Statistics* 87, no. 2 (2005): 217-34.

Smith, Charles H, and Rosa L Nemir. "The Sulfapyridine Treatment of Pneumonia in Children." *Journal of the American Medical Association* 113, no. 21 (1939): 1857-60.

Smith, James P. "Race and Human Capital." *American Economic Review* 74, no. 4 (1984): 685-98.

Stein, Aryeh D, Jere R Behrman, Ann DiGirolamo, Ruben Grajeda, Reynaldo Martorell, Agnes Quisumbing, and Usha Ramakrishnan. "Schooling, Educational Achievement, and Cognitive Functioning among Young Guatemalan Adults." *Food and Nutritional Bulletin* 26, no. 2 (2005): S46-S54.

Thomasson, Melissa A, and Jaret Treber. "From Home to Hospital: The Evolution of Childbirth in the United States, 1928-1940." *Explorations in Economic History* 45, no. 1 (2008): 76-99.

Ungerleider, Harry E, Henry W Steinhaus, and Richard S Gubner. "Public Health and Economic Aspects of Pneumonia - a Comparison with Pre-Sulfonomide Years." *American Journal of Public Health* 33 (1943): 1093-102.

van der Poll, Tom, and Steven M Opal. "Pathogenesis, Treatment, and Prevention of Pneumococcal Pneumonia." *The Lancet* 374 (2009): 1543-56.

Venkataramani, Atheendar. "Early Life Exposure to Malaria and Cognition in Adulthood: Evidence from Mexico." *Journal of Health Economics* 31, no. 5 (2012): 767-80.

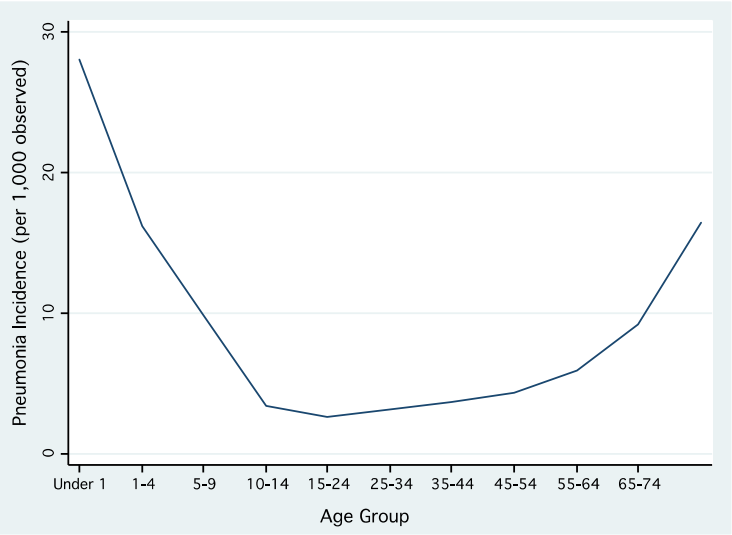
Waldron, I. "The Role of Genetic and Biological Factors in Sex Differences in Mortality." In *Sex Differentials in Mortality: Trends, Determinants and Consequences*, edited by Alan D. Lopez and Lado T. Ruzicka. Canberra and New York, NY: Dept. of Demography, PB Consulting Service, 1983.

Wegman, Myron E. "Infant Mortality in the 20th Century, Dramatic but Uneven Progress." *Journal of Nutrition* 131, no. Suppl (2001): 401S-08S.

Welch, Finis. "Black-White Differences in Returns to Schooling." *The American Economic Review* 63, no. 5 (1974): 893-907.

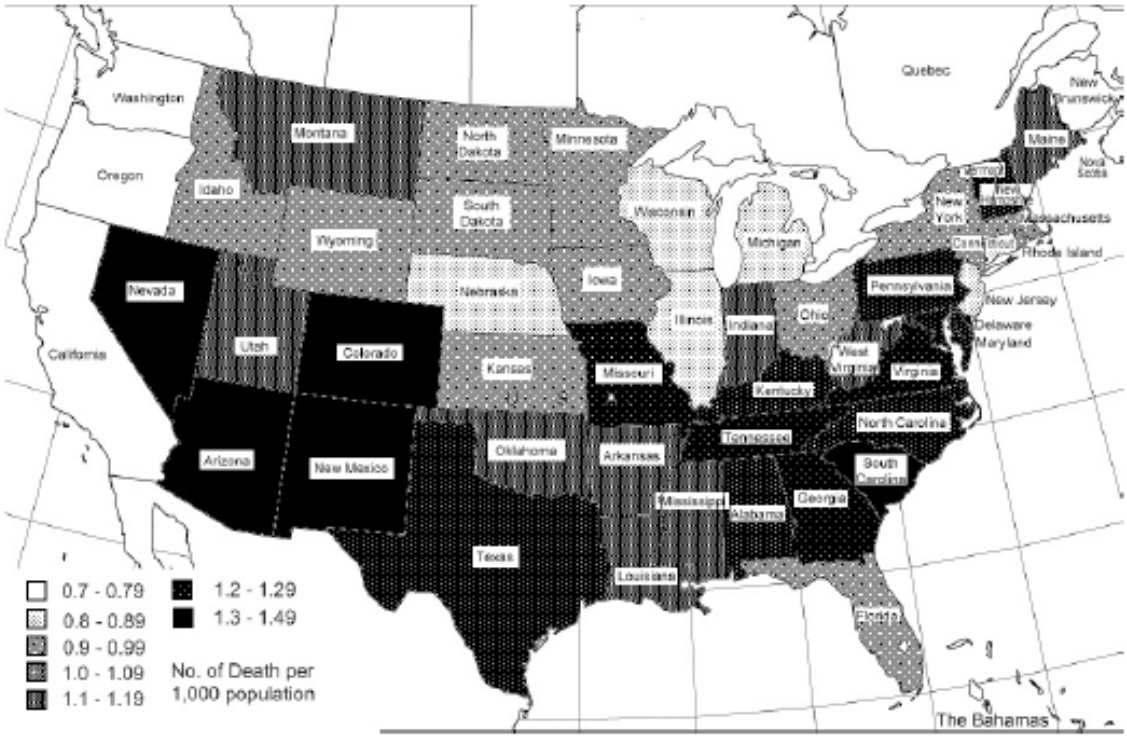
World Health Organization. *Pneumonia, Fact Sheet N331* 2011 [cited March 2011]. Available from http://www.who.int/mediacentre/factsheets/fs331/en/index.html.

Figure 1 – Age Distribution of Pneumonia in the 1930s



Source: Britten (1942)

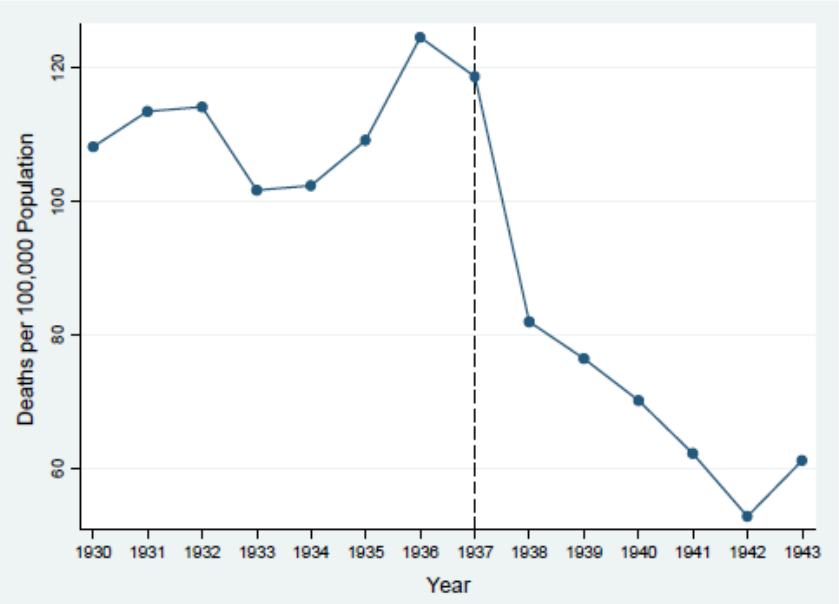
Figure 2 – Geographic Distribution of Pneumonia by Average Mortality Rates (1930-1936)



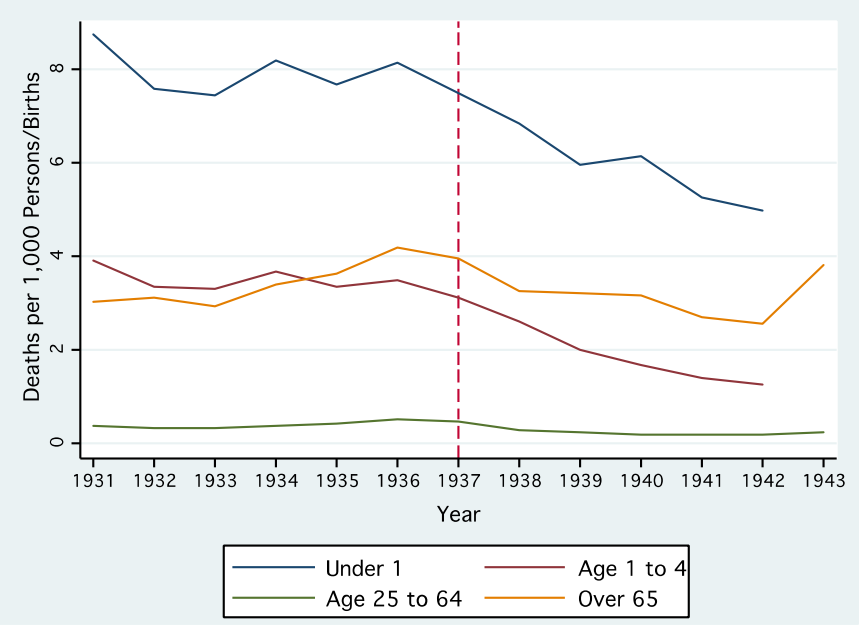
Source: Map drawn by authors using all-age pneumonia and influenza death rates (see **Data Appendix for** further details).

Figure 3 – National Trends in Pneumonia Mortality, 1930-1943

Panel A – All-age



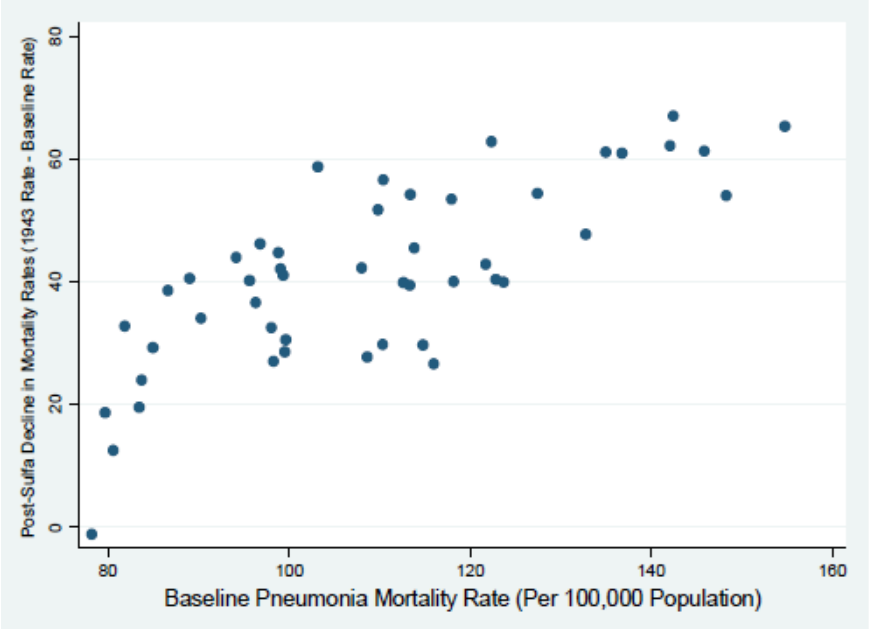
Panel B – Age specific



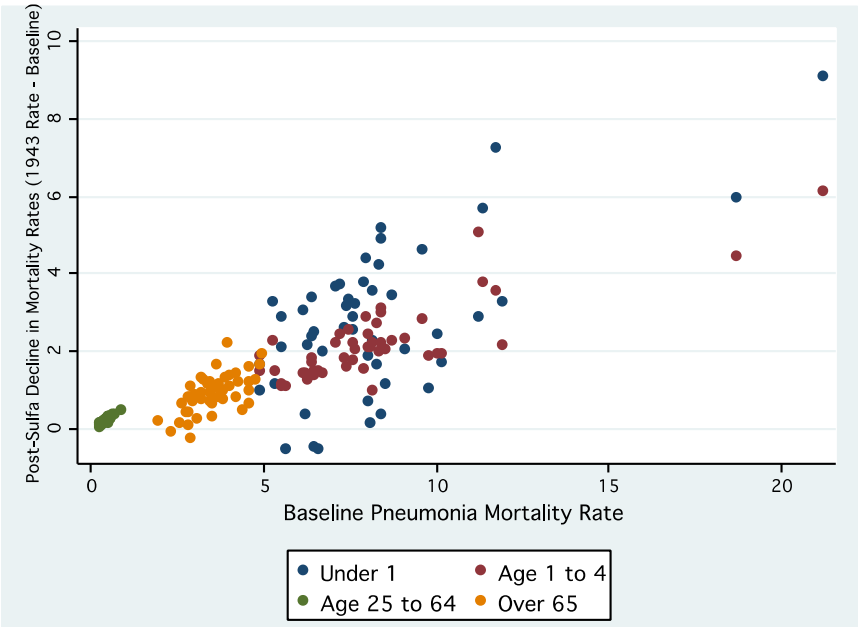
Source: US Vital Statistics.

Figure 4 – Convergence in Pneumonia Mortality Rates after 1937

Panel A – All age

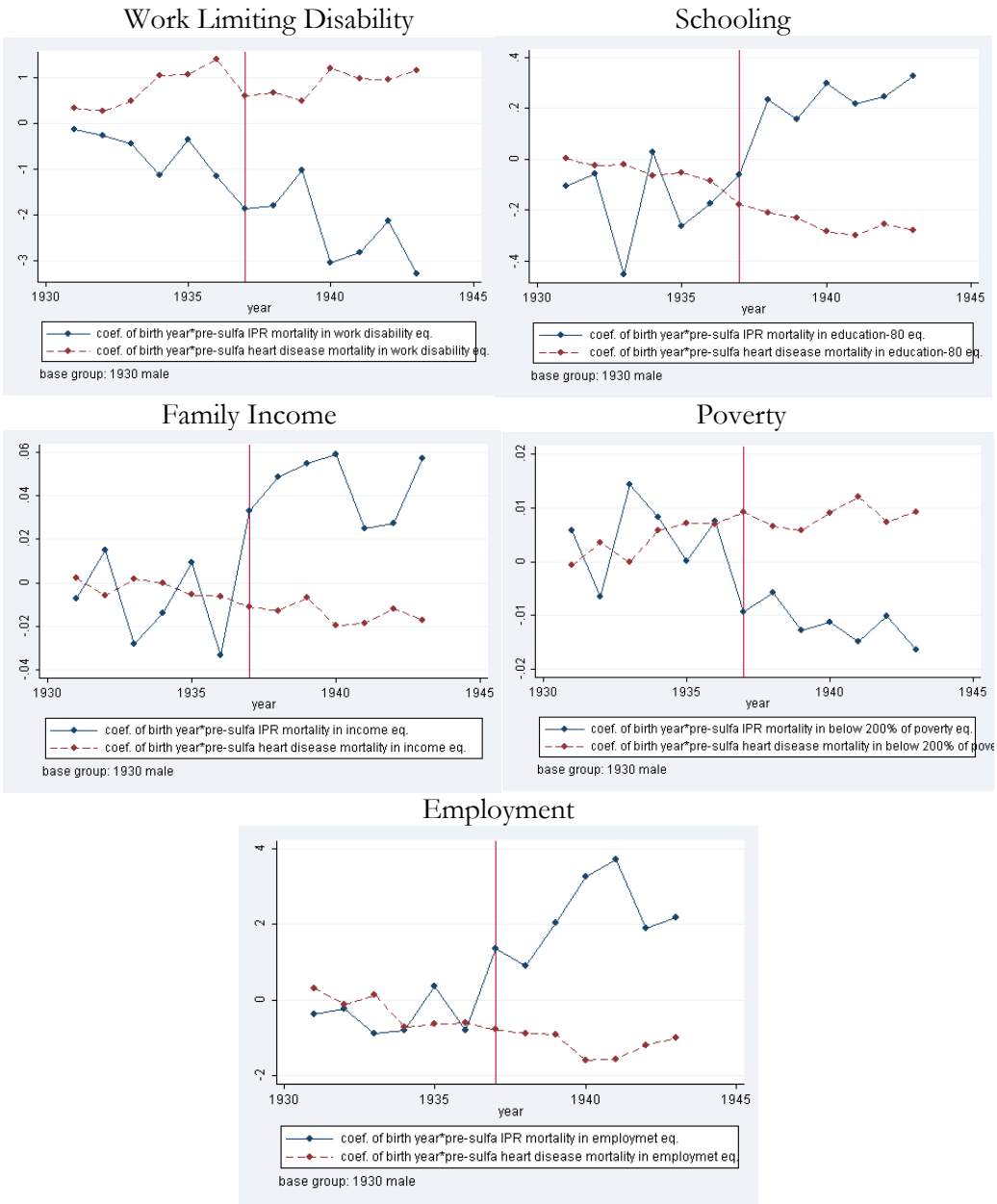


Panel B – Age Specific



Source: US Vital Statistics.

Figure 5a: Birth-Year Specific Coefficients on Baseline Pneumonia Mortality Rates, Men

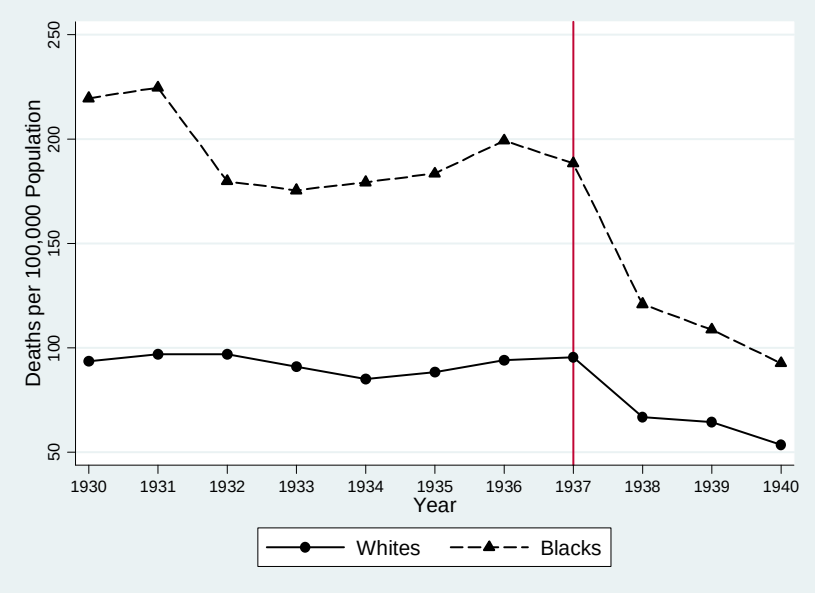


Notes: Each point reflects the estimate on the interaction between the marked birth year and the pre-intervention (base) level of the pneumonia mortality rate in the birth-state. This is conditional upon birth state and birth year fixed effects and the full set of controls for mortality from other diseases and state macroeconomic and infrastructure variables. The vertical line denotes the year sulfa drugs became available in the United States. Agents more efficacious against pneumonia became available in 1938. The dashed line represents similar estimates for heart disease, which we use as a control. We have confirmed that the other control diseases (diarrhea, TB, cancer) show no trend break.

Figure 5b: Birth-Year Specific Coefficients on Baseline Pneumonia Mortality Rates, Women

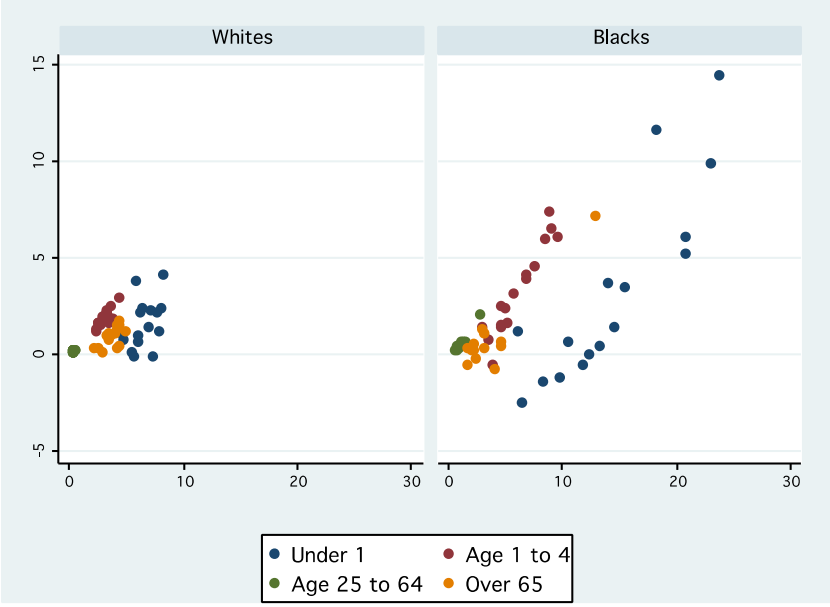
Notes: See notes to *Figure 5a*

Figure 6 – National Race Specific Trends in Pneumonia Mortality



Source: US Vital Statistics (gathered by Adriana Lleras-Muney).

Figure 7 – Convergence in Race-Specific Mortality Rates



Source: US Vital Statistics

Figure 8 – Cohort Size by Birth Cohort in the 1980 Census Sample

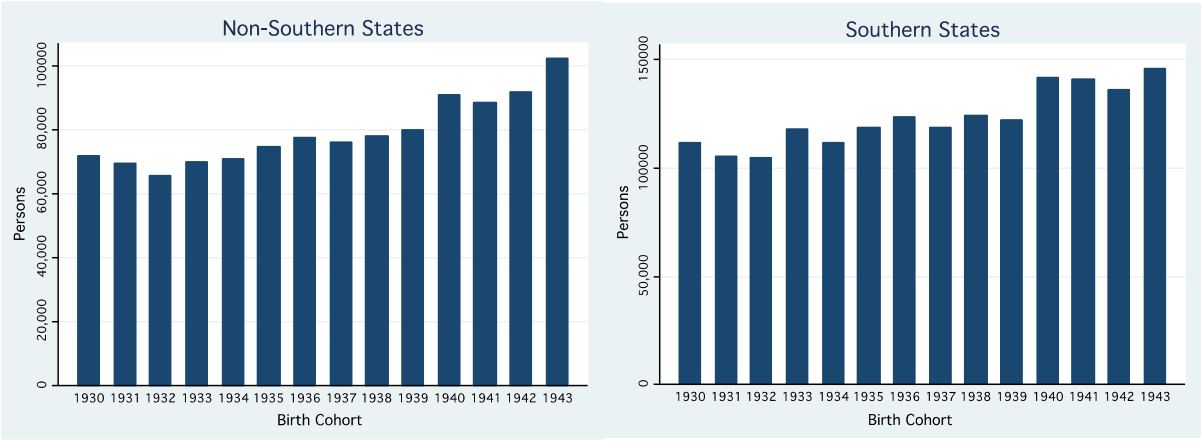
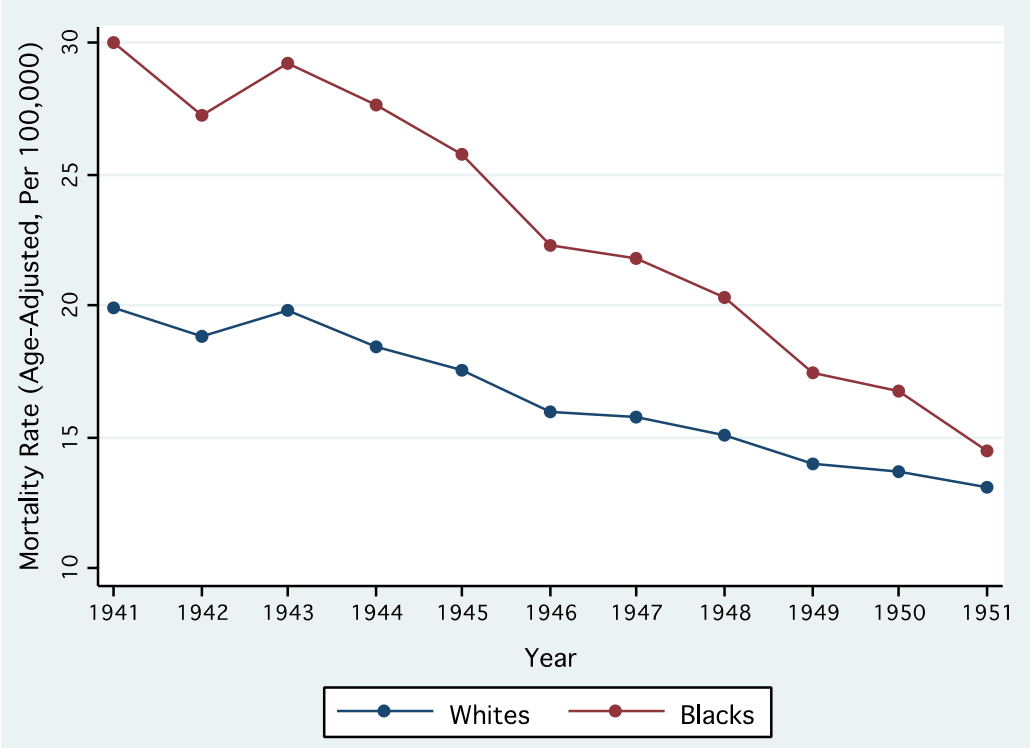


Figure 9 – Rheumatic Fever Mortality Rates by Race, 1941-1951



Source: US Vital Statistics

Table 1a – Estimates of Impacts on Adult Outcomes of Pneumonia Exposure in Infancy – Men

	(1)	(2)	(3)	(4)	(5)
Schooling	0.399***	0.328***	0.350***	0.565***	0.727***
(N = 1130, mean = 12.17)	(0.105)	(0.0829)	(0.108)	(0.0919)	(0.148)
College	-0.0138	-0.0167	0.00966	0.0649***	0.111***
(N = 1130, mean = 0.39)	(0.0109)	(0.0107)	(0.0155)	(0.0140)	(0.0246)
High School	0.0822***	0.0726***	0.0684***	0.0495***	0.0623***
(N = 1130, mean = 0.70)	(0.0189)	(0.0146)	(0.0185)	(0.0163)	(0.0139)
Employment	0.0157***	0.0154***	0.0254***	0.0272**	0.0290*
(N = 3336, mean = 0.68)	(0.00573)	(0.0045)	(0.00507)	(0.0108)	(0.0160)
Family Income (log)	0.0387***	0.0371***	0.0517***	0.0805***	0.109***
(N = 3336, mean = 10.57)	(0.0138)	(0.0124)	(0.0166)	(0.0186)	(0.0191)
Poverty	-0.0272***	-0.0271***	-0.0157**	-0.0277**	-0.0161*
(N = 3336, mean = 0.24)	(0.00540)	(0.00511)	(0.00768)	(0.0112)	(0.00884)
Marriage Age	0.0539	-0.0466	0.162	0.260	0.369
(N = 1231, mean = 21.56)	(0.167)	(0.142)	(0.213)	(0.291)	(0.311)
Work Limiting Disability	-0.0291***	-0.0259***	-0.018***	-0.01	-0.0233***
(N = 3336, mean = 0.16)	(0.0060)	(0.0059)	(0.0059)	(0.001)	(0.0086)
Cognitive Disability	-0.0140**	-0.0111**	-0.00315	-0.0211	-0.0253
(N =1102, mean = 0.076)	(0.00601)	(0.00546)	(0.00732)	(0.0127)	(0.0164)
Physical Disability	-0.0123	-0.00930	-0.00253	-0.00789	0.0107
(N =1102, mean = 0.22)	(0.00980)	(0.0101)	(0.0120)	(0.0259)	(0.0224)
Controls					
Birth State, Birth Year FE	Yes	Yes	Yes	Yes	Yes
Post*BaseRate(Control Diseases)	No	Yes	Yes	Yes	Yes
Birth State X Birth Year Variables	No	No	Yes	Yes	Yes
Birth State Linear Trends	No	No	No	Yes	Yes
Birth Region X Birth Year FE	No	No	No	No	Yes

Notes: *** = p<0.01, ** = p<0.05, * = p<0.10. Robust standard errors adjusted for clustering at the birth state level are in parentheses. Every row contains a different adult outcome and the set of controls is enlarged in moving from the first to the fifth column. Every cell contains the coefficient on postt*base_pneumonias from a different regression. The outcome variables are defined in the data appendix and their means and s.d. are in Appendix Table 1. FE denotes fixed effects. BaseRate(Control Diseases) includes pre-sulfa birth state averages for maternal mortality and mortality from heart disease, cancer, under 2 diarrhea, malaria and tuberculosis. Birth State*Birth Year Variables includes per capita log state income, state education expenditure, state health expenditure, school buildings, hospitals, and physicians, all by birth state and birth year. Region refers to census division region. N refers to the number of birth state x birth year x race x gender x census year cells. Each cell is weighted by its population in the regression analysis; unweighted regressions produce substantively similar results.

Table 1b - Estimates of Impacts on Adult Outcomes of Pneumonia Exposure in Infancy - Women

Women	(1)	(2)	(3)	(4)	(5)
Schooling (Years)	0.410***	0.396***	0.197*	-0.0183	-0.0382
(N=1137, mean = 11.95)	(0.0848)	(0.08)	(0.108)	(0.134)	(0.134)
College	-0.0321**	-0.0274**	-0.00722	0.00616	0.00277
(N=1137, mean = 0.33)	(0.0145)	(0.0117)	(0.0137)	(0.0195)	(0.0263)
High School	0.107***	0.0975***	0.0553**	-0.0186	-0.0166
(N=1137, mean = 0.71)	(0.0228)	(0.0187)	(0.0218)	(0.0190)	(0.0313)
Employment	-0.0102	-0.00825	0.0190***	0.0102	-0.00436
(N = 3378, mean = 0.52)	(0.0065)	(0.0058)	(0.0065)	(0.0112)	(0.0139)
Family Income (log)	0.0545***	0.0506***	0.0325**	-0.00507	0.0364
(N = 3378, mean = 10.44)	(0.0147)	(0.0125)	(0.0135)	(0.0197)	(0.0233)
Poverty	-0.0386***	-0.0328***	-0.0188**	-0.00724	-0.017
(N = 3378, mean = 0.30)	(0.0075)	(0.00623)	(0.00897)	(0.0124)	(0.0139)
Marriage Age (Years)	0.0966	-0.0218	0.0290	-0.101	0.311
(N = 1231, mean = 21.56)	(0.139)	(0.111)	(0.176)	(0.290)	(0.377)
Work Limiting Disability	-0.0249***	-0.0202***	-0.00366	0.01439*	0.0177**
(N = 3378, mean = 0.15)	(0.00691)	(0.00572)	(0.0066)	(0.00772)	(0.00798)
Cognitive Disability	-0.00903	-0.00731	0.00665	0.0202*	0.0115
(N = 1118, mean = 0.07)	(0.00605)	(0.00520)	(0.00768)	(0.0120)	(0.0131)
Physical Disability	-0.0104	-0.00623	0.00596	0.0347**	0.0313
(N = 1118, mean =0.23)	(0.00886)	(0.00882)	(0.0111)	(0.0152)	(0.0201)
Controls					
Birth State, Birth Year FE	Yes	Yes	Yes	Yes	Yes
Post*BaseRate(Control Diseases)	No	Yes	Yes	Yes	Yes
Birth State X Birth Year Variables	No	No	Yes	Yes	Yes
Birth State Linear Trends	No	No	No	Yes	Yes
Birth Region X Birth Year FE	No	No	No	No	Yes

See Notes to Table 1a.

Table 2 – Robustness Checks for Main Results

	Schooling (years)	College	High School	Employment	Family Income	Poverty	Marriage Age	Work Disability	Cognitive Disability	Physical Disability
Panel A – Removing the influenza uptick of 1936-1937										
Men	0.696***	0.0856***	0.0730***	0.0288*	0.0798**	-0.0287**	-0.800***	-0.0122	-0.0364**	-0.0501
	(0.143)	(0.0267)	(0.0151)	(0.0153)	(0.0314)	(0.0117)	(0.219)	(0.0118)	(0.0173)	(0.0301)
Women	-0.0551	0.0095	-0.044	0.0253*	-0.0209	-0.00925	-0.0769	0.0203**	0.00846	0.0466*
	(0.198)	(0.0288)	(0.0276)	(0.0149)	(0.0252)	(0.0128)	(0.215)	(0.0097)	(0.0125)	(0.0268)
The years 1936-37 are removed from the sample as there was an influenza epidemic during this period (see Jayachandran et al, 2010). Men: N = 965 for Schooling, College and High School, N = 2851 for Employment, Family Income, Poverty, and Work Disability, N = 1057 for Marriage Age, N = 944 for Cognitive and Physical Disability. Women: N = 970 for Schooling, College, and High School, N = 2879 for Employment, Family Income, Poverty and Work Disability, N = 1062 for Marriage Age, N = 952 for Cognitive and Physical Disability										
Panel B- Narrow Cohort Window										
Men	0.413***	0.0498**	0.0570**	0.0152	0.108***	-0.0287**	0.425	-0.00956	-0.00463	0.0252
	(0.149)	(0.0197)	(0.0269)	(0.0127)	(0.0203)	(0.0141)	(0.442)	(0.0171)	(0.0180)	(0.0344)
Women	0.0142	0.00976	0.0222	-0.000621	-0.0264	-0.0000105	-0.209	0.00992	0.0299**	0.0329*
	(0.163)	(0.0279)	(0.0235)	(0.0159)	(0.0250)	(0.000161)	(0.406)	(0.00949)	(0.0137)	(0.0191)
Model estimated for 1935-1941 birth cohorts. Men: N = 616 for Schooling, College and High School, N = 1824 for Employment, Family Income, Poverty, and Work Disability, N = 616 for Marriage Age, N = 604 for Cognitive and Physical Disability. Women: N = 619 for Schooling, College, and High School, N = 618 for Marriage Age, N = 1847 for Employment, Family Income, Poverty and Work Disability, N = 613 for Cognitive and Physical Disability										
Panel C - Placebo Intervention in 1934										
Men	-0.0915	-0.0200	-0.00772	-0.00957	-0.00921	0.00821	0.0138	0.00878	0.00487	0.0253**
	(0.113)	(0.0137)	(0.0197)	(0.00605)	(0.0153)	(0.00661)	(0.206)	(0.00589)	(0.00675)	(0.0126)
Women	-0.0866	0.020*	0.0229	-0.0207***	-0.0300**	0.0034	0.383**	-0.00258	-0.00719	-0.0045
	(0.0897)	(0.0119)	(0.0224)	(0.00713)	(0.0149)	(0.0084)	(0.166)	(0.00495)	(0.00757)	(0.0129)
The placebo involves defining *post* as 1 for cohorts born in and after 1934 instead of in and after 1937. See Table 1 for sample sizes for each variable.										
Panel D - Triple Difference Estimates										
Men	0.493***	0.0477**	0.0822**	0.0259**	0.0731***	0.00377	0.0818	-0.00798	0.0140	-0.0987
	(1.75)	(0.0144)	(0.0336)	(0.0104)	(0.0242)	(0.0412)	(0.235)	(0.00654)	(0.0133)	(0.0156)
Women	0.247	0.0170	0.0864*	0.0276**	0.0164	0.0052	-0.365	0.00279	0.0112	0.00915
	(0.236)	(0.0175)	(0.0491)	(0.0106)	(0.0242)	(0.0178)	(0.384)	(0.0103)	(0.0118)	(0.00133)
The triple difference term is post*base_pneumonia*treated where post and base_pneumonia are defined as in the baseline model and treated is defined as 1 if first exposed to sulfa at age 0-1 and as 0 if first exposed at age 10-15. Men: N = 2210 for Schooling, College and High School, N = 6472 for Employment, Family Income, Poverty, and Work disability, N = 2203 for Marriage Age, N = 2119 for Cognitive and Physical Disability. Women: N = 2226 for Schooling, College, High School, and Marriage Age, N = 6589 for Employment, Family Income, Poverty and Work Disability, N = 2168 for Cognitive and Physical Disability.										
See Notes to Table 1. The baseline specification on which these variations are estimated is that in column 4 of Table 1.										

Table 3 – Assessing Selective Migration After the Arrival of Sulfa Drugs

	(1)	(2)
Full Sample	0.100*	0.0308
	(0.0518)	(0.0713)
High School Educated Sample	-0.208	-0.127
	(0.142)	(0.201)

Notes: See Section IIID.3. The dependent variable is the log of the population aged 20-40. Each cell contains the coefficient on post*base_pneumonia from a separate regression.

Data are from the US Census Microdata from 1930 and 1940 and collapsed to the state*year level. Data for the year 1935 were gleaned from questions on state of residence 5 years before the enumeration in the 1940 census. There are 48 states, with N = 144. Column 1 contains only state and year fixed effects. Column 2 adds in controls for baseline income, baseline diarrheal mortality, and baseline heart disease mortality interacted with post-1937.

Table 4 – Differences in Infrastructure and Socioeconomic Outcomes by Race Across US Regions

	Non-South		South	
	White	Black	White	Black
Slave Fraction	0		0.31	
	0		(0.18)	
Pupil-Teacher Ratio			27.95	32.16
			(1.97)	(4.52)
Total Years Schooling	12.83	12.01	11.89	8.52
	(0.48)	(1.59)	(0.85)	(1.17)
Fraction Completing High School	0.36	0.34	0.29	0.07
	(0.05)	(0.29)	(0.08)	(0.05)
log(Earned Income)	6.77	6.39	6.67	6.00
	(0.22)	(0.38)	(0.26)	(0.33)

Notes: The cells report region*race means, with standard deviations in parentheses. Slave fraction is the proportion of the state population enslaved in 1860 (see Section IIIA and Data Appendix for details). Pupil-Teacher Ratio (taken from Card and Krueger, 1992) is only available for Southern states. The education and income variables are for adult males aged 20-45 in the 1940 Census. There are 31 states in Non-South and 17 in South.

Table 5 – Estimates for Men by Race

	(1)	(2)	(3)		(4)	(5)	(6)
White Men				**Black Men**			
Schooling	0.443***	0.620***	0.970***	**Schooling**	-0.153††	-0.143††	-0.100††
(N = 515, mean = 12.58)	(0.0996)	(0.111)	(0.204)	(N = 492, mean=11.54)	(0.249)	(0.203)	(0.231)
College	0.0140	0.0752***	0.138***	**College**	-0.0254†	0.0428	0.0304
(N = 515, mean = 0.40)	(0.0152)	(0.0157)	(0.0302)	(N = 492, mean = 0.28)	(0.0193)	(0.0529)	(0.0591)
High School	0.0864***	0.0490**	0.0689***	**High School**	0.0197	0.0317	0.0376
(N = 515, mean = 0.75)	(0.0175)	(0.0184)	(0.0202)	(N = 492, mean = 0.62)	(0.0403)	(0.0562)	(0.0579)
Employment	0.0188**	0.0229*	0.0300	**Employment**	-0.0158	0.0332	0.0000261
(N= 1510, mean = 0.73)	(0.00758)	(0.01312)	(0.0190)	(N = 1422, mean = 0.61)	(0.0239)	-0.0433	-0.0378
Family Income (log)	0.0560***	0.0808***	0.116***	**Family Income (log)**	0.011	0.124	0.0223
(N= 1510, mean = 10.71)	(0.0162)	(0.0219)	(0.0315)	(N = 1422, mean = 10.39)	(0.0597)	(0.111)	(0.088)
Poverty	-0.0128*	-0.0240**	-0.0126	**Poverty**	-0.0181	-0.0572**	-0.0350
(N = 1914, mean = 0.17)	(0.00751)	(0.0115)	(0.00807)	(N = 1422, mean = 0.35)	(0.0180)	(0.0259)	(0.0223)
Marriage Age	0.173	0.467	0.883**	**Marriage Age**	-0.290	-0.281	0.131
(N = 515, mean = 21.83)	(0.270)	(0.346)	(0.346)	(N = 490, mean = 21.18)	(0.582)	(1.033)	(1.288)
Work Limiting Disability	-0.0215***	-0.0165	-0.0339***	**Work Limiting Disability**	0.00756	-0.0084	0.00713
(N= 1510, mean = 0.13)	(0.00666)	(0.0109)	(0.00696)	(N = 1422, mean = 0.19)	(0.0160)	(0.0350)	(0.0410)
Cognitive Disability	0.00142	-0.0150	-0.00587	**Cognitive Disability**	-0.0406*†	-0.0655	-0.0356
(N = 499, mean = 0.06)	(0.00725)	(0.0133)	(0.0210)	(N = 464, mean = 0.10)	(0.0227)	(0.0607)	(0.0623)
Physical Disability	-0.00174	-0.00963	0.0314	**Physical Disability**	0.0138	-0.0165	0.0140
(N = 499, mean = 0.19)	(0.0129)	(0.0304)	(0.0239)	(N = 464, mean = 0.26)	(0.0467)	(0.0581)	(0.0801)
Controls							
Birth State, Birth Year FE	Yes	Yes	Yes	Birth State, Birth Year FE	Yes	Yes	Yes
Post*BaseRate(Control Diseases)	Yes	Yes	Yes	Post*BaseRate(Control Diseases)	Yes	Yes	Yes
Birth State X Birth Year Variables	Yes	Yes	Yes	Birth State X Birth Year Variables	Yes	Yes	Yes
Birth State Linear Trends	No	Yes	Yes	Birth State Linear Trends	No	Yes	Yes
Birth Region X Birth Year FE	No	No	Yes	Birth Region X Birth Year FE	No	No	Yes

See Notes to Table 1. Each cell represents an estimate from a separate regression. The controls included are the same as those in column 4 of Table 1. For comparability purposes, we restrict our estimation sample to states only those states where black births occured. The symbols † and †† denote statistically significant differences in the black and white coefficients at the 5 and 10%, respectively. We computed this by estimating models where the treatment exposure variable, as well as all of the controls, were interacted with the race fixed effect.

Table 6a – Heterogeneity in the Long-Run Impacts of Pneumonia by Indices of Racial Segregation, Black Men

	Schooling	High School	College	Employment	Family Income	Poverty	Marriage Age	Disability work	Cognitive Disability	Physical Disability
Panel A - North/South										
post*base	0.636	0.0952	0.267*	0.0978	0.421**	-0.0342	-0.151	-0.159*	-0.0908	-0.0816
	(0.582)	(0.1030)	(0.0137)	(0.0838)	(0.175)	(0.0635)	(2.734)	(0.0856)	(0.111)	(0.139)
post*base*south	-0.946	-0.0566	-0.274*	-0.0957	-0.383*	-0.0227	-0.0432	0.191**	0.0265	0.099
	(0.654)	(0.113)	(0.0145)	(0.0907)	(0.197)	(0.0701)	(3.156)	(9.42)	(0.121)	(0.149)
Panel B - Slave Fraction										
post*base	0.231	0.0520	0.202*	0.146**	0.465***	-0.0762	-1.103	-0.105	0.112	-0.0848
	(0.573)	(0.0789)	(0.112)	(0.0635)	(0.113)	(0.0506)	(1.801)	(0.0668)	(0.0973)	-0.112
post*base*slave	-0.774	-0.0256	0.438	-0.322**	-1.026***	0.0594	2.438	0.294*	0.114	0.159
	(1.373)	(0.191)	(0.264)	(0.145)	(0.281)	(0.125)	(4.178)	(0.147)	(0.209)	(0.259)
Panel C - B-W Education Ratio										
post*base	-0.0371	0.0184	0.134	0.107*	0.377***	-0.071	-0.362	-0.0736	-0.0181	-0.0371
	(0.468)	(0.0964)	(0.0955)	(0.0612)	(0.104)	(0.0465)	(1.698)	(0.0587)	(0.0824)	(0.0948)
post*base*bweduc	0.862	-0.105	0.720*	0.559**	1.794***	-0.112	-0.428	-0.518	0.253	0.238
	(2.598)	(0.540)	(0.511)	(0.326)	(0.652)	(0.257)	(9.622)	(0.321)	(0.415)	(0.492)
N	492	490	492	1422	1422	1422	490	1422	464	464
N (Slave Fraction)	433	433	433	1268	1268	1268	432	1268	422	416

Notes: Each pair of cells (post*base and post*base*segregation) represents estimates from a separate regression. In Panel A, the coefficient on post*base is the coefficient for black men born in Northern states and the coefficient on post*base*south is the change in this coefficient for black men born in the South. Similarly post*base in Panel B is the coefficient for states with no slave history (essentially Northern states) but now there is variation in slave share that allows heterogeneity in impacts within the Southern states. In Panel C we have normalized the ratio so that post*base is the coefficient for the notional state with equal years of education for black and white men and the interaction term shows how outcomes vary as this ratio varies across states. The model estimated here is similar to that estimated in Table 1 (with the same controls as in column 4 of Table 1), except we additionally estimate the triple interaction post-37*base_pneumonia*segregation index, where segregation index includes the dummy for birth in the South (south), the population share of slaves in 1860 (slave), and the ratio of schooling between young black and white men in the 1940 census (bweduc). Each of these indicators is discussed in the Data Appendix and summary statistics are in Table A3. All models include double interactions post-37*segregation and base_pneumonia*segregation (which is subsumed in the birth state fixed effect).

Table 6b - Heterogeneity in the Long-Run Impacts of Pneumonia by Indices of Racial Segregation, White Men

	Schooling	High School	College	Employment	Family Income	Poverty	Marriage Age	Disability work	Cognitive Disability	Physical Disability
Panel A - North/South										
post*base	0.529***	0.0397	0.0552**	-0.00659	0.0761**	-0.0293*	-0.155	-0.0104	-0.0219	-0.0146
	(0.152)	(0.0281)	(0.0238)	(0.0123)	(0.0308)	(0.0173)	(0.352)	(0.0145)	(0.0196)	(0.0429)
post*base*south	0.171	0.0272	0.0544	0.0755***	0.0249	0.00955	1.661**	-0.00657	0.0260	0.0457
	(0.320)	(0.0609)	(0.0467)	(0.027)	(0.0638)	(0.0238)	(0.729)	(0.0195)	(0.0344)	(0.0633)
Panel B - Slave Fraction										
post*base	0.566***	0.0646**	0.0705***	0.00363	0.0853***	-0.0300**	0.323	-0.0172	-0.0165	-0.0156
	(0.144)	(0.0265)	(0.0194)	(0.0124)	(0.0269)	(0.0129)	(0.370)	(0.0164)	(0.0172)	(0.0408)
post*base*slave	0.360	-0.0936	0.0640	0.151***	0.0144	0.0596	1.816	-0.00241	0.0358	0.00338
	(0.587)	(0.111)	(0.0704)	(0.0513)	(0.131)	(0.0379)	(1.502)	(0.0455)	(0.0528)	(0.138)
Panel C - B-W Education Ratio										
post*base	0.646***	0.0530***	0.0770***	0.0216*	0.0889***	-0.0281***	0.393	-0.0184	-0.0168	-0.0119
	(0.106)	(0.0190)	(0.0171)	(0.0116)	(0.0215)	(0.00990)	(0.307)	(0.0113)	(0.0133)	(0.0307)
post*base*bweduc	0.402	0.139	0.00409	0.123*	0.193	-0.209***	-3.426**	-0.0378	-0.0616	0.022
	(0.778)	(0.153)	(0.132)	(0.0670)	(0.125)	(0.0532)	(1.646)	(0.0533)	(0.0798)	(0.180)
N	515	515	515	1510	1510	1510	515	1510	499	499
N (Slave Fraction)	447	447	447	1326	1326	1326	447	1326	441	441

See Notes to Table 6a. As in Table 5, we restrict our estimation sample to states only those states where black births occurred.

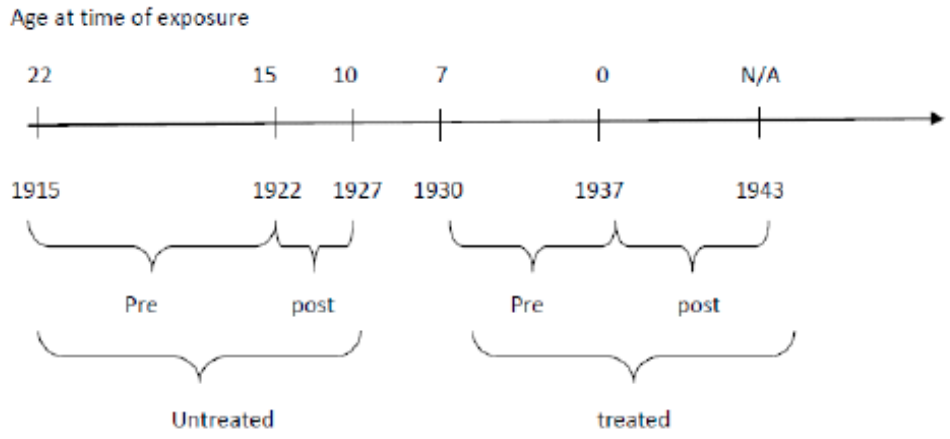
Table 7 – Robustness Checks for Impact Gradient Models

	Schooling	High School	College	Employment	Family Income	Poverty	Marriage Age	Disability work	Cognitive Disability	Physical Disability
Panel A - Controlling for Time of Entry into National Death Registration System										
post*base	1.745	0.104	0.672***	0.287*	0.700**	-0.276**	-0.115	0.023	-0.202	0.0380
	(1.108)	(0.163)	(0.232)	-0.153	(0.311)	(0.115)	(4.473)	(0.164)	(0.158)	(0.265)
post*base*slave	-0.814	-0.104	-0.400*	-0.396**	-0.762**	0.0308	6.097	0.264	0.0616	0.0968
	(1.466)	(0.206)	(0.232)	(0.170)	(0.333)	(0.161)	(4.472)	(0.156)	(0.217)	(0.284)
Panel B - Controlling for Pharmacists per Capita in Black Counties										
post*base	-0.432	0.0939	0.260	0.169	0.640**	-0.290**	0.742	-0.133	-0.115	-0.0898
	(0.971)	(0.166)	(0.208)	(0.104)	(0.246)	(0.114)	(4.025)	(0.129)	(0.209)	(0.229)
post*base*slave	-0.404	0.0519	-0.545	-0.487**	-1.590***	0.516**	0.622	0.183	0.127	0.0248
	(1.613)	(0.309)	(0.344)	(0.201)	(0.424)	(0.187)	(6.980)	(0.170)	(0.359)	(0.393)
N	433	433	433	1268	1268	1268	432	1268	422	416
N (Panel B)	302	302	302	906	906	906	302	906	302	302

Notes: Each pair of cells represents results from a different regression. All regressions are for the black sample. Panel A includes interactions between post, post*base, and a measure for the number of years prior to 1936 that the birth state joined the National Death Registration System to the model from Table 6a. Panel B includes interactions between post, post*base, and the number of pharmacists per capita operating in black counties (those counties with greater than 10% of the population black, with population weighted averages computed for the entire state). Each robustness check is conducted using the models with the slave fraction variable, but the results hold for the models using the south dummy and black-white education ratio, as well. Please see the notes for Table 6a for further details.

Appendix Figures

Figure A1 – Triple Difference Specification: Treatment Definitions on Cohort and Age of Exposure



Appendix Tables

Table A1 – Descriptive Statistics

Census Variables	Men	White Men	Black Men	Women	White Women	Black Women
Schooling	12.14 (1.03)	12.58 (0.78)	11.54 (1.04)	11.94 (0.75)	12.17 (0.61)	11.64 (0.80)
High School	0.71 (0.15)	0.77 (0.09)	0.62 (0.17)	0.71 (0.14)	0.77 (0.09)	0.64 (0.15)
College	0.35 (0.13)	0.40 (0.09)	0.28 (0.14)	0.28 (0.10)	0.30 (0.08)	0.25 (0.11)
Log Family Income	10.58 (0.44)	10.72 (0.38)	10.39 (0.44)	10.45 (0.42)	10.61 (0.33)	10.23 (0.43)
Poverty	0.24 (0.13)	0.17 (0.05)	0.35 (0.13)	0.29 (0.15)	0.20 (0.06)	0.42 (0.15)
Employed	0.68 (0.23)	0.73 (0.23)	0.61 (0.24)	0.52 (0.19)	0.52 (0.17)	0.52 (0.21)
Marriage Age	21.56 (2.16)	21.83 (0.73)	21.18 (3.18)	19.40 (1.52)	19.63 (0.55)	19.10 (2.21)
Work Disability	0.16 (0.09)	0.13 (0.05)	0.19 (0.12)	0.15 (0.09)	0.12 (0.05)	0.19 (0.12)
Cognitive Difficulty	0.08 (0.07)	0.06 (0.04)	0.10 (0.09)	0.07 (0.06)	0.05 (0.02)	0.10 (0.09)
Physical Difficulty	0.22 (0.11)	0.18 (0.06)	0.26 (0.16)	0.22 (0.11)	0.18 (0.05)	0.28 (0.13)

Birth State Baseline Mortality Rates (Per Thousand, N = 48 States)		Birth State X Birth Year Socioeconomic Variables (N = 669)	
Pneumonia	1.06 (0.19)	Log Income Per Capita	6.18 (0.50)
Tuberculosis	0.64 (0.37)	Log Hospitals Per 1,000	-2.83 (0.46)
Diarrhea	8.22 (5.64)	Log Physicians Per 1,000	0.14 (0.36)
Cancer	0.96 (0.31)	Log Schools Per 1,000	0.72 (0.65)
Heart Disease	2.09 (0.63)	Log Educational Spending Per Capita	4.27 (0.79)
Maternal mortality	6.35 (1.24)		
Malaria	0.032 (0.074)		

Notes: Figures provided are means with standard deviations in parentheses. The means for the census variables are based on the 2019930 men, 2137468 women, 1821471 white men, 198459 black men, 1897973 white women and 249495 black women born between 1930 and 1943 who are part of the 1980, 1990 and 2000 5% US census samples available from IPUMS-USA. Census family income figures have been converted to 2000 dollars. Baseline mortality rates reflect average mortality rates for each birth state over the period 1930-1936. Rates for pneumonia, tuberculosis, malaria, heart disease and cancer reflect deaths per 1,000 total population. Diarrhea and maternal mortality rates are per 1,000 live births.

Table A2 – Heterogeneity in Impacts for Women by Measures of Labor Force Opportunities in the Sulfa Drug Era

	Schooling	High School	College	Employment	Family Income	Poverty	Marriage Age	Disability work	Cognitive Disability	Physical Disability
post*base	-0.0213	-0.0229	0.00544	0.997	-0.00504	-0.00655	-0.136	1.419*	0.0214*	0.0371**
	(0.152)	(0.0193)	(0.0199)	(1.058)	(0.0194)	(0.0123)	(0.298)	(0.753)	(0.0113)	(0.0153)
post*base*wage ratio	-0.221	0.377	0.0189	6.394	-0.0574	-0.0952	2.156	-0.796	-0.141	-0.111
	(1.753)	(0.228)	(0.257)	(16.87)	(0.270)	(0.197)	(3.656)	(10.61)	(0.175)	(0.260)

Notes: Each cell pair reflects estimate for a separate regression. The estimation sample includes women only. These regressions are similar to those in Table 1b, except we now interactions between post*base and an birth state specific indicator for the female to male wage earnings ratio for adults aged 25-45 in the 1940 Census (all the relevant double interactions are accounted for, as well). This measure is to proxy for the opportunities available to women while the sulfa cohorts were young adults. To interpret the regressions, post*base refers to the estimated impacts at the mean female-male wage ratio and post*base*wage ratio represents the marginal (additional) impact for a one standard deviation change in this variable. See the main text for further details. All models include the same controls as in column 4 of Table 1b. Of note, the substantive results are unchanged if we use the female-male ratio in years of school or occupation index scores.

Table A3 – Estimates of Trend Breaks in Mortality Rates After 1937

	Levels			Logs		
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	Pneumonia	TB	MMR	Pneumonia	TB	MMR
1930-1943						
Blacks	**-0.114*** **	0.0409***	-0.250***	-0.0814***	0.0172***	-0.0548***
	(0.0188)	(0.00957)	(0.0904)	(0.0101)	(0.0064)	(0.0102)
Whites	**-0.0773*** **	0.00513**	-0.235***	-0.0984***	-0.000346	-0.0963***
	(0.00962)	(0.0023)	(0.0351)	(0.00949)	(0.00507)	(0.00778)
1925-43						
Blacks	-0.0862***	0.0289***	-0.369***	-0.0676***	0.00532	-0.0689***
	(0.0169)	(0.00823)	(0.0819)	(0.00853)	(0.00507)	(0.01)
Whites	-0.0640***	0.0105***	-0.262***	-0.0888***	0.00406	-0.103***
	(0.0096)	(0.00223)	(0.0392)	(0.00874)	(0.00403)	(0.00639)

Notes: Each dependent variable*time period*race cell represents a separate regression estimate of the coefficient on post*year. Each model controls for post (=1 if the year is 1937 or greater), year (1937 set to 0) and state fixed effects. The first column shows that the trend break for absolute pneumonia mortality is significantly larger for blacks than whites, but slightly smaller when using logged mortality.

For comparison with Jayachandran et al. (2010), we show similar estimates for tuberculosis (TB, a control disease) and maternal mortality (MMR, a sulfa-treatable condition). We show these estimates in levels as well as logs and for the sample period used in Jayachandran et al., which is 1925-1943. While extending the sample period back to 1925 lowers the coefficients for blacks more than for whites (lower panel), the black advantage is preserved. The evolution of TB, included in the regression as a control disease in Jayachandran, et al (2010) differs starkly by race. For maternal mortality (MMR) we see a white advantage when log mortality is used, consistent with Jayachandran et al. For the purposes of the paper, the upshot is that there is no evidence that the sulfa-induced drop in pneumonia after 1937 was smaller for blacks

Table A4 – Estimates for Women by Race

	(1)	(2)	(3)		(4)	(5)	(6)
White Women				**Black Women**			
Schooling	0.275***	-0.0560	0.0525	**Schooling**	-0.198	-0.220	-0.381
(N = 544, Mean = 12.16)	(0.0855)	(0.174)	(0.264)	(N = 499, mean=11.64)	(0.363)	(0.293)	(0.374)
College	0.000167	0.00627	0.0215	**College**	-0.0483*	0.0285	0.0304
(N = 544, Mean = 0.29)	(0.0124)	(0.0206)	(0.0366)	(N = 499, mean=0.30)	(0.0261)	(0.0370)	(0.0409)
High School	0.0732***	-0.00928	0.00770	**High School**	0.0217	-0.0547	-0.0562
(N = 544, Mean =0.76)	(0.0218)	(0.0199)	(0.0405)	(N = 499, mean=0.63)	(0.0498)	(0.0606)	(0.0849)
Employment	0.0106	0.0101	-0.00344	**Employment**	0.0139	-0.0409	-0.0204
(N = 1593, Mean = 0.52)	(0.0083)	(0.0122)	(0.0181)	(N = 1464, mean = 0.52)	(0.0204)	(0.0284)	(0.0211)
Family Income (log)	0.0334**	-0.0170	0.0236	**Family Income (log)**	0.0190	0.0818	0.118*
(N = 1593, Mean = 10.61)	(0.0159)	(0.0255)	(0.0333)	(N = 1464, mean = 9.39)	(0.0389)	(0.0652)	(0.0689)
Poverty	-0.0209**	-0.00995	-0.0266*	**Poverty**	-0.00899	-0.000978	-0.00949
(N = 1593, Mean = 0.20)	(0.00893)	(0.0156)	(0.0154)	(N = 1464, mean = 0.42)	(0.0193)	(0.0234)	(0.0336)
Marriage Age	-0.144	-0.0730	0.319	**Marriage Age**	0.459	0.482	0.486
(N = 544, Mean = 19.7)	(0.215)	(0.314)	(0.460)	(N = 524, mean 19.1)	(0.584)	(0.921)	(1.112)
Work Limiting Disability	-0.00654	0.0137	0.0205	**Work Limiting Disability**	0.0141	0.0115	-0.00981
(N = 1593, Mean = 0.12)	(0.00697)	(0.00975)	(0.0126)	(N = 1464, mean = 19.70)	(0.0138)	(0.0302)	(0.0221)
Cognitive Disability	0.00976	0.0200	0.00464	**Cognitive Disability**	0.0223	0.0175	0.00543
(N = 525, Mean = 0.05)	(0.00742)	(0.0140)	(0.0173)	(N = 480, mean = 0.10)	(0.0270)	(0.0343)	(0.0337)
Physical Disability	0.0190*	0.0447***	0.0404	**Physical Disability**	-0.0195	0.119	0.0637
(N = 525, Mean = 0.18)	(0.0111)	(0.0136)	(0.0244)	(N = 480, mean = 0.28)	(0.0344)	(0.0971)	(0.0775)
Controls							
Birth State, Birth Year FE	Yes	Yes	Yes	Birth State, Birth Year FE	Yes	Yes	Yes
Post*BaseRate(Control Diseases)	Yes	Yes	Yes	Post*BaseRate(Control Diseases)	Yes	Yes	Yes
Birth State X Birth Year Variables	Yes	Yes	Yes	Birth State X Birth Year Variables	Yes	Yes	Yes
Birth State Linear Trends	No	Yes	Yes	Birth State Linear Trends	No	Yes	Yes
Birth Region X Birth Year FE	No	No	Yes	Birth Region X Birth Year FE	No	No	Yes

See Table 5 notes for further details.

Table A5 – Northward Migration of Blacks between 1935 and 1940 as a Function of Baseline Pneumonia Mortality

	(1)	(2)
Full Sample	-0.0510**	-0.00103
(N = 42305)	(0.0329)	(0.0113)
Some High School and Beyond	-0.0760**	0.00396
(N = 15328)	(0.0485)	(0.00658)

Notes: Probit regressions of the probability of moving Northward between 1935 and 1940 among those living in Southern states in 1935 (See Data Appendix for definitions of North and South) as a function of the baseline state level pneumonia mortality rate (defined previously). Controls in column 1 include age and sex. Column 2 adds baseline state per capita income, under 2 diarrheal mortality rates, and heart disease mortality. Each cell represents a probit marginal effect from a separate regression. For the subgroup analysis, "Some High School and Beyond" (defined as completing 8th grade or better) was chosen as it is near the sample median and a natural cutpoint. The results are substantively similar if high school completion is used.

Data Appendix

Outcome variables

All of our outcomes data were taken from the 5% United States Census Microdata samples for 1980, 1990 and 2000. These data are publicly available via the Integrated Public Use Microdata Series – USA project (Ruggles et al., 2010) at http://usa.ipums.org/usa/. As discussed in Section II, we use averages of the individual data by birth state, birth year, census year, race and sex. For income, poverty and employment, we utilize cell level data from all three censuses. Models for years of schooling, high school completion, and college completion use only the 1980 census to avoid duplicating the data given that years of schooling seldom change after the age of 37, the age in 1980 of the youngest cohort (born 1943) in our sample. Physical and cognitive disability measures are only available in the 2000 census.

There is some concern in the literature that the 2000 census microdata sample may be subject to inaccuracies in age reporting (Alexander et al., 2010). While this problem primarily pertains to those over the age of 65, all of whom were born at least two years prior to the start of the sulfa era, we still assessed whether our results remained the same if the 2000 census was excluded. The substantive results were unchanged.

Specific sources and construction of each of our outcome variables is as follows:

Schooling (HIGRADE in IPUMS) – In the 1980 census, years of schooling are designated for the pre-high school years. Later census files group those completing ninth grade and under into three categories and top code those who progress beyond college. This favors our choice of using only the 1980 census data for this variable.

High School and College – We computed these using the *Schooling* measure above. Specifically, we assigned *High School* = 1 for those individuals who completed grade 12 and above. *College* = 1 for those individuals who reported completing 4 years of college.

Logged total family income (*FTOTINC* in IPUMS) – Nominal total pre-tax money income earned by the respondent's family unit in the previous calendar year.

Poverty - Indicator for whether family income is 200% below the federal household poverty line or less. We constructed this using the *POVERTY* variable in IPUMS (which specifies the percent above the poverty line for a given reported level of income).

Marriage Age (*AGEMARR* in IPUMS) – Reflects the mean age of first marriage as reported in the 1980 Census (this question is not available in later samples).

Employed (EMPSTAT in IPUMS) - Individual employment = 1 if the individual reports current employment and 0 otherwise.

Work limiting disability (DISABWRK in IPUMS) – Indicates a physical or mental health condition that causes difficulty working, limits the amount or type of work, or prevents working altogether. The disability cannot be transient (e.g., pregnancy) and must have been present for at least six month prior to survey. We coded any limitation in the ability to work (either certain limitations or the inability to work altogether) as representing disability.

Cognitive disability (DIFFREM in IPUMS) – Denotes whether an individual has difficulty with cognitive tasks due to a physical or mental illness.

Physical disability (DIFFPHYS in IPUMS)- Denotes whether the respondent has a condition that limits basic tasks of daily living that involve movement (walking, running, lifting, etc).

Mortality Data

State time series data on mortality rates from influenza and pneumonia, under-2 diarrhea, heart disease, cancer and tuberculosis, and the maternal mortality ratio are obtained from various volumes of the US Vital Statistics (Grove, 1968; Linder, 1947; United States Bureau of the Census, 1930-1943). We combined and extended the data series collected by Grant Miller (http://www.nber.org/data/vital-statistics-deaths-historical/), and by Seema Jayachandran, Adriana Lleras-Muney, and Kimberly Smith (http://www.aeaweb.org/articles.php?doi=10.1257/app.2.2.118). We obtained the malaria data from the NBER. We used these data to create birth state-specific pre-sulfa era rates for each disease by averaging the cause-specific mortality rates between 1930 and 1936. (The choice of the time period over which we compute baseline rates does not change our substantive results).

Race specific state mortality data for 18 states (with blacks comprising >10% of the population) were generously provided by Adriana Lleras-Muney. They were originally gleaned from yearly US Vital Statistics volumes (http://www.cdc.gov/nchs/products/vsus.htm).

Data on rheumatic fever mortality rates by race were taken from the 1940-1960 US Vital Statistics compendium (Grove, 1968).

State level socioeconomic and infrastructure variables

State time series data on logged state per capita income were downloaded from the Bureau of Economic Analysis website (http://www.bea.gov/regional/spi/). Data on the number of schools, doctors, hospitals, and educational expenditures per capita were taken from Adriana Lleras-Muney's website (http://www.econ.ucla.edu/alleras/research/data.html). These data were originally collected from various volumes of the *Biennial Survey of Education* (schools and expenditures) and the American Medical Association's *American Medical Directory* (doctors and hospitals). We used linear interpolation for each state to calculate education and health infrastructure values for 1940-1943, as Lleras-Muney's data was only collected through to 1939. For state per capita health expenditures, we used data at: http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/6304?archive=ICPSR&q=6304, collected from various reports from the US Census bureau. The only year in our sample for which this data source had information for was 1932, so we used data for that year, interacted with *post* in our regression models.

The measure of pharmacists we use as a specification check in Section IIID was taken from 1940 IPUMS Census microdata. Specifically, we computed the number of individuals reporting pharmacist as their occupation per 1,000 black population living in counties where the black population share was greater than 10%. We then computed state averages weighting by county black population.

Indicators of institutionalized racial segregation

North-South - We define South to include Alabama, Delaware, District of Columbia, Florida Georgia, Kentucky, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, West Virginia, Arkansas, Louisiana, Texas and Missouri. For expositional ease, our use of the term *North* refers to all regions of the US other than the Southern states.

Slave Fraction - The state-specific share of slaves in the population in 1860 (Nunn, 2008). Available for all states and is zero for non-southern states.

Race education ratio - Ratio of black to white years of education (using the IPUMS variable *HIGRADE*), each being the state average for individuals age 20-45 in the 1940 census.

On measurement error in age-specific mortality rates

As discussed in the Section IIB, we use all-age mortality rates in our analysis given well-known measurement error in infant deaths in the 1930s (Linder and Grove, 1947). This error is thought to tend towards systematic underreporting of mortality rates with a good deal of variation in the amount of error across states.

Appendix Table A6 replicates the results from *Table 1a* using pneumonia mortality rates for children under 5 instead of all-age pneumonia mortality. To given the former the best possible chance, we compute the baseline mortality rates for the period 1934-1936 as, by this time, all states had joined the US Birth and Death Registration system and/or were actively improving births and death recording systems. The estimates generally follow the same signs as in *Table 1b*. However, the implied effect sizes are all smaller than magnitude than those obtained when using all-age mortality rates. Only the estimate on years of schooling is statistically significant.

These findings confirm our *a priori* suspicions regarding measurement error in age specific mortality measurements. To be clear, what they do *not* suggest is that the association between all age pneumonia and adult outcomes we find is driven primarily by changes in adult pneumonia (say, through reduced susceptibility to income shocks in infancy). The reason for this, as we mention in Section IIB, is that pneumonia morbidity was highest among infants (*Figure 1*) and the bulk of the post-sulfa decline in pneumonia mortality was driven by reductions in infant deaths (*Figure 3, Panel B*).

Moreover, when we rerun our models using infant pneumonia mortality rates computed over the period 1930-1936 as our baseline, which we expect to be even more noisily measured given the above discussion, the coefficients we obtain are even smaller (and none are statistically significant – available upon request). This is consistent with our measurement-error based explanation for the results in *Appendix Table A6*.

On measurement error in mortality rates for blacks

The key concern in our exploration of gradients in the long-term impacts of early life pneumonia is that our measures of the (constraints to) returns to human capital investments – that is, indicators of institutionalized segregation – may be correlated with the severity of measurement error in mortality rates. As discussed in the text, we opt to use race-average mortality rates given that mortality in rural areas may have been subject to underreporting (Ewbank, 1940) and a large fraction of blacks lived in predominantly rural Southern states. Also, attribution of death to specific diseases, as well as

recording of deaths in general, may have been less accurate for blacks given their more limited access to hospitals.

In *Appendix Table A7*, we compare estimates of the impact of early life pneumonia using race-averaged and race-specific mortality rates in the sample of blacks for which we have both sets of data. Scaling the impacts by differences in the dispersion in the mortality rate measures we find that, impacts on long-run outcomes are often smaller in magnitude and less likely to be statistically significant (at least, for the poverty variable) when race-specific mortality measures are used, though the two sets of results are not drastically different from each other in the qualitative sense. Given this, and our *a priori* concerns about race-specific mortality data, we believe our use of race-averaged mortality rates is appropriate and persist with this in our analysis.

Table A6 – Pneumonia Impact Results Using Age-Specific Mortality Rates

	Schooling	High School	College	Employment	Family Income	Poverty	Marriage Age	Disability work	Cognitive Disability	Physical Disability
Panel A - Age-Specific Mortality Rates										
post*base	0.312**	0.0244	0.0158	-0.0158	0.0339	-0.0189	0.586*	0.00268	0.0191	0.0302
	(0.147)	(0.0221)	(0.0240)	(0.0152)	(0.0282)	(0.0127)	(0.292)	(0.0104)	(0.0165)	(0.0302)
	[0.109]	[0.85%]	[0.55%]	[-0.55%]	[1.19%]	[-0.66%]	[0.205]	[0.09%]	[0.67%]	[0.11%]
Panel B - Age-Averaged Mortality Rates										
post*base	0.565***	0.0495***	0.0649***	0.0272**	0.0805***	-0.0277**	0.260	-0.01	-0.0211	-0.00789
	(0.0919)	(0.0163)	(0.0140)	(0.0108)	(0.0186)	(0.0112)	(0.291)	(0.001)	(0.0127)	(0.0259)
	[0.152]	[1.34%]	[1.75%]	[0.73%]	[2.17%]	[-0.74%]	[0.070]	[-0.27%]	[-0.57%]	[-0.21%]

Notes: Panels A and B present results from the regression in Column 4 of Table 1a. Panel A uses age-specific mortality rates to compute the baseline pneumonia rate (specifically, the number of pneumonia deaths for children under the age of 5 per 100 live births). Panel B uses the same age-averaged mortality rate used throughout the paper. Robust, cluster corrected standard errors are in parentheses. The square brackets contain estimates of the effect sizes when the mortality rate in a birth state moves from the 75th to the 25th percentile of the baseline pneumonia mortality distribution. See Table 1a for further details and sample size information. All models include the controls in Column 4 of Table 1.

Table A7 – Impacts for Blacks Using Race-Specific Mortality Rates

	Schooling	High School	College	Employment	Family Income	Poverty	Marriage Age	Disability work	Cognitive Disability	Physical Disability
Panel A - Race-Specific Mortality Rates										
post*base	-0.0996	0.0201	0.0193	-0.0287	0.0129	-0.00793	0.531	0.0154	-0.0377	0.00898
	(0.109)	(0.0334)	(0.0214)	(0.0186)	(0.0563)	(0.00612)	(0.703)	-0.0163	(0.0324)	(0.0403)
	[-0.0777]	[1.57%]	[1.51%]	[2.23%]	[1.01%]	[-0.62%]	[0.414]	[1.20%]	[-2.95%]	[0.70%]
Panel B - Race-Averaged Mortality Rates										
post*base	-0.208	0.0224	0.0299	-1.219	0.0398	-0.0301	-0.442	1.440	-0.0911	-0.0218
	(0.155)	(0.0650)	(0.0428)	(2.744)	(0.112)	(0.0149)	(1.212)	(3.951)	(0.0609)	(0.0745)
	[-0.058]	[0.62%]	[0.88%]	[-0.34%]	[1.11%]	[-0.84%]	[-0.123]	[0.40%]	[-2.55%]	[-0.61%]

Notes: Panels A and B present results from Table 5 (black sample only) using race-specific and race-averaged mortality rates, respectively, to calculate the baseline pneumonia mortality rate. Panel B differs from Table 5 in that the sample is restricted only to those states for which we have (more accurate) race-specific mortality data. Robust, cluster corrected standard errors are in parentheses. The square brackets contain estimates of the effect sizes when the mortality rate in a birth state moves from the 75th to the 25th percentile of the baseline pneumonia mortality distribution. See Table 5 for further details and sample size information. All models include the controls in Column 4 of Table 1.