The Economic Effects of the 1918 Influenza Epidemic*

Elizabeth Brainerd

Department of Economics
Williams College
Williamstown, MA 01267
Phone: 413-597-4375
Email: Elizabeth.Brainerd@williams.edu

Mark V. Siegler

Department of Economics
California State University, Sacramento
Sacramento, CA 95819-6082
Phone: 916-278-7079
Email: msiegler@csus.edu

June 2002

Abstract: The 1918-19 influenza epidemic killed at least 40 million people worldwide and 675,000 people in the United States, far exceeding the combat deaths tolls experienced in the U.S. in the two World Wars, Korea, and Vietnam combined. Besides its extraordinary virulence, the 1918-19 epidemic was also unique in that a disproportionate number of its victims were men and women age 15 to 44, giving the age profile of mortality a distinct 'W' shape rather than the 'U' shape usually observed in an influenza epidemic, and leading to an extremely high death rate in the prime working ages. This paper examines the impact of this exogenous shock on subsequent economic growth using data on U.S. states for the 1919–1930 period. Controlling for numerous factors including initial income, density, human capital, climate, the sectoral composition of output, geography, and the legacy of slavery, the results indicate a large and robust positive effect of the influenza epidemic on per capita income growth across states during the 1920s. Although the positive influence of mortality on growth is consistent with the predictions of the standard neoclassical growth model and some endogenous growth models, at least some of the growth from 1919-1921 to 1930 likely represents a return to trend rather than a change in trend.

Key Words: Influenza, Flu, 1918, Economic Growth.
JEL Classifications: I10, N12, O40.

*The authors would like to thank Ralph Bradburd, Kevin Hoover, Oscar Jorda, Steve Perez, Lara Shore-Sheppard, Anand Swamy, participants in the UC Davis Macro Brown Bag seminar and the Williams Economics Department seminar for helpful comments and suggestions. We would also like to thank Laura Bennett and Thomas Sproul for their excellent research assistance.

I. Introduction

In his Presidential Address to the Economic History Association, Neal (2000) argued that his fellow economic historians would do the "economics profession, and the society at large, a big favor if we focused an increasing share of our research efforts on shocks, rather than on longer periods of 'normal' economic change."[1] The 1918 influenza epidemic certainly qualifies as a large shock. In the last four months of 1918 and the first six months of 1919, at least 40 million people worldwide died from this influenza epidemic.[2] This death toll exceeds the cumulative twenty-year toll from the AIDS epidemic.

In the United States, Crosby (1989) estimated that 675,000 Americans died from influenza and pneumonic complications and that about 550,000 of these were "excess deaths" of Americans who would have otherwise lived during a normal year.[3] These "excess deaths" surpass the number of combat deaths in the U.S. Armed Forces during World War I, World War II, the Korean War and Vietnam combined.[4] However, this epidemic has been almost completely ignored by economists and economic historians. A comprehensive search of *EconLit* found only two journal articles relating to the 1918 influenza epidemic, and none of the leading economic history textbooks even mention the epidemic.[5]

---

[1] Neal (2000, p. 332).
[2] The most recent estimate of the worldwide number of deaths due to the epidemic is 40 to 50 million (Potter 2001).
[3] Crosby (1989, pp. 206-207).
[4] Using U.S. Department of Defense and U.S. Coast Guard estimates, Ellis (2001, p. 209) reports 426,704 battle deaths during World War I, World War II, the Korean War, and the Vietnam War.
[5] A search on June 12, 2002 found only the articles by Noymer and Garenne (2000) and Bloom and Mahal (1997b) using the keywords "flu," "influenza," and "1918" separately. In addition, the textbooks by Atack and Passell (1994), Walton and Rockoff (2001), and Cain and Hughes (1997) fail to mention the epidemic.

The 1918 influenza epidemic is an important episode to study not only because of its sheer magnitude, but also because economists know little about how large population shocks affect economic growth: economic theory offers ambiguous predictions regarding the relationship between negative population shocks and economic growth, and the other major historical pandemics provide inconclusive evidence on the issue, as we review below. The importance of understanding the relationship is further underscored by the massive loss of life due to AIDS in many developing countries; in part due to lack of evidence the effect of the AIDS epidemic on economic growth in these regions is far from a resolved issue. Although we emphasize below the differences between the influenza epidemic and the AIDS epidemic, the two are clearly linked by the almost incomprehensible scale of the deaths recorded in both crises.

The purpose of this paper is to examine the effects of the 1918-1919 influenza epidemic on subsequent economic growth using data from a sample of U.S. states. We find that the epidemic is positively correlated with subsequent economic growth in the United States, even after taking into account differences in population density, levels of income per capita, climate, geography, the sectoral composition of output, human capital accumulation, and the legacy of slavery. Our results suggest that one more death per thousand resulted in an average annual increase in the rate of economic growth over the next ten years of at least 0.2 percent per year. This effect is larger than that suggested by simulations using both a standard neoclassical or endogenous AK growth model. However, we find that flu deaths in 1918 and 1919 among prime-age adults are a significant predictor of business failures in 1919 and 1920. This implies that one reason for the positive association between flu deaths and economic growth from 1919-1921 to

1930 is that the epidemic caused substantial business failures which caused the economy to be below trend on average between 1919 and 1921. In other words, some of the growth from 1919-1921 to 1930 is not a change in trend, but only a return to trend after this large temporary shock.

II. An Overview of the 1918-1919 Influenza Epidemic

A. Worldwide

Few health crises in world history rival the severity of the 1918-1919 influenza epidemic in terms of death tolls. Only the Plague of Justinian (100 million lives lost over 50 years), the Black Death of 1348-1351 (62 million lives), and the current AIDS epidemic (25 million) approach or exceed the mortality caused by the influenza epidemic.[6] This section outlines what is currently known about the biological nature of this influenza virus, its geographic spread and demographic consequences in the United States, and its diffusion and effect worldwide.

The influenza epidemic swept the world in three waves: the first in spring 1918, the second deadly wave in fall 1918, and a third wave that further afflicted some regions in early 1919. The precise origin of the epidemic is unknown, but the first recorded outbreak occurred in March 1918 among army recruits at Camp Funston, Kansas.[7] The virus spread quickly across the United States and reached Europe by early 1918, apparently with the arrival of American troop ships. The influenza epidemic swept

[6] See Potter (2001) for estimates of the death toll during the Plague of Justinian and the Black Death. The estimate from the AIDS epidemic is UNAIDS (2001, 2000).

[7] Recent research indicates that the epidemic likely originated in a large French mass transit camp in 1916, where influenza victims exhibited the distinctive symptoms that characterized the later epidemic (Oxford et al. 1999). The authors of this study hypothesize that the virus smoldered for several years before exploding in 1918.

across Europe and had reached India, Australia and New Zealand by June (Patterson and Pyle 1991).

The virulence and mortality rate of this first wave of influenza only slightly exceeded normal levels and it therefore attracted little attention. The epidemic was, however, characterized by two traits that link it clearly to the second deadly wave in the fall, and which distinguish this influenza epidemic from any epidemic before or since: it disproportionately killed young adults, and victims died with excessive accumulation of bloody fluid in their lungs, often with severe complications due to pneumonia (Crosby 1989).

The second wave began in August 1918 with new, deadly outbreaks of influenza occurring nearly simultaneously in Brest, France; Freetown, Sierra Leone; and Boston, Massachusetts. Undoubtedly fueled by the troop movements and demobilization surrounding the end of World War I, the virus spread explosively around the world in the second wave, with worldwide mortality from the epidemic peaking in October and November 1918. Only extremely isolated areas of the world – such as eastern Iceland – escaped the influenza epidemic altogether, and even the strictest quarantines proved ineffective in preventing the epidemic in most regions (Patterson and Pyle 1991). A third wave affected some areas of the world in early 1919, principally England and Wales, Australia, and other countries in the southern hemisphere.

Why was the fall wave so deadly – with mortality rates 5 to 20 times higher than normal – and why did it primarily strike young adults? These questions remain unanswered, despite much recent research on the 1918 influenza epidemic by microbiologists. The extraordinarily virulent fall strain of the virus is believed to have

occurred through genetic mutation or recombination of the spring strain, which significantly enhanced the transmissibility and deadliness of the virus (Reid et al. 2001). Recent analysis of frozen tissue samples from known influenza victims buried in permafrost in Artic regions has identified the genetic structure of the virus as a form of swine and avian influenza strains (Reid et al. 1999, 2000; Taubenberger et al. 1997), but the virulence and age pattern of mortality of the influenza epidemic remain unexplained. One hypothesis argues that the high death rate experienced by young men in the United States in the epidemic is consistent with an interaction between influenza and tuberculosis, since tuberculosis incidence is higher for men than for women and is a disease of young adults (Noymer and Garenne 2000). However, this hypothesis fails to explain the pattern of deaths in other countries in which male and female influenza deaths were approximately equal – as in Europe and Japan – or in which female deaths exceeded male deaths, as in India (Mills 1986; Rice 1990; Japan Statistics Bureau 1945). Reid et al. (2001) speculate that the unusually low mortality of the elderly during the 1918 influenza epidemic[8] may have resulted from heightened immune status due to exposure to a similar influenza strain that circulated before 1850, but this hypothesis remains unproven. As these authors point out, the high death rates of young adults cannot be explained by the poor living conditions of those involved in the war, because the death rates among young people affected by the war were the same as those unaffected by the war.

B. The United States

Although the United States escaped the influenza epidemic with a relatively low average death rate by world standards (5.8 deaths from influenza and pneumonia per

[8] See Luk et al. (2001) for evidence on this.

1,000 population), the impact of the epidemic varied widely across regions and had a profound demographic impact on the country. As noted previously, it is estimated that 675,000 people died of influenza and pneumonia during the epidemic; given a population estimate of 103 million in the United States on July 1, 1918 (Linder and Grove 1943), the influenza epidemic killed approximately 0.66 percent of the U.S. population. As illustrated in Figure 1, this translated into a decline in life expectancy of 11.8 years for both men and women in 1918.

In a typical influenza epidemic the majority of the victims are young children and the elderly, giving the age profile of mortality a distinct 'U' shape. A distinguishing characteristic of the 1918 epidemic was that it disproportionately killed men and women age 15 to 44, so that the age profile of mortality instead followed a 'W' pattern. This is illustrated in Figure 2, which plots the age-specific death rates per 100,000 population for men and women in the 1918 epidemic and the average rates from 1914 to 1916, along with the same data for whites and nonwhites separately. It is evident from Figure 2 that over one percent of males ages 25 to 34 died as a result of this epidemic. For both whites and nonwhites the male mortality rate in the 15-44 age group exceeds the female mortality rate by 50-75 percent in 1918, in contrast to the non-epidemic years in which the death rates by gender are virtually identical. The death rate for nonwhites also exceeds that of whites, although the 'W' pattern characterizes the age-specific death rates of both races. All countries for which age-specific death rates are available also recorded a 'W'-shaped age distribution of mortality; this is true, for example, in India, Australia, New Zealand, and South Africa (Mills 1986; Rice 1990; Union of South Africa 1921).

Other than the age, sex, and racial distribution of deaths, little else is known about the pattern of deaths across different subgroups of the population in the United States. In some countries such as South Africa, New Zealand and Australia, the European population experienced much lower death rates than did the indigenous populations who suffered severely;[9] American Indians in the United States are also reported to have experienced much higher death rates than the rest of the population (Crosby 1989). Whether these ethnic differences in death rates were due to differences in socioeconomic status or due to weaker immunity to the influenza virus is an unresolved issue. However, a door-to-door Public Health Service survey of more than 100,000 individuals conducted in nine cities during the summer of 1919 suggests that the mortality rate of whites "was nearly twice as great among the 'very poor' as among the 'well-to-do' and those classified as in 'moderate' circumstances."[10] Other observers argue that the socioeconomic status of influenza victims indicates few differences in mortality rates across income groups (Crosby 1989; Rice 1988).

It does seem clear that the influenza epidemic did not simply kill the weakest members of each cohort. Numerous eyewitness accounts by doctors and other medical personnel attest that influenza killed the most robust individuals in the population; for example the Acting Surgeon General of the Army remarked that the influenza epidemic

[9] For example, the death rate among Maoris in New Zealand was 42.4 per 1000 population, compared with 5.8 for the European population (Rice 1990).

[10] The results of this survey are reported and summarized in Sydenstricker (1931, p. 160). These results should obviously be weighed with extreme caution. Since the surveys were not conducted until after the flu epidemic, it is unclear whether the poverty was a cause or a result of flu mortality within each household. In addition, the categorization of economic status was not well defined and certainly not comparable across surveyors or cities. Sydenstricker (1931, p. 156) reported that "each enumerator was instructed to record at the time of her visit to the household her impression of its economic condition in one of four categories – 'well-to-do,' 'moderate,' 'poor,' or 'very poor' . . . They were purposely given no standards for comparison . . . [with] the intention being to have tem record their own impressions naturally."

"kills the young vigorous, robust adults;" public health specialists agreed as well that most influenza victims were those who "had been in the best of physical condition and freest from previous disease" (quoted in Crosby 1989 pp. 215-16).

The geographic spread of the influenza epidemic in the United States appears to have been arbitrary, with few discernible patterns evident in the areas affected lightly or severely by the epidemic. The states with the highest death rates from the epidemic – Pennsylvania, Maryland and Colorado – shared few common features climatically or economically, and in some cases neighboring states, cities or even counties experienced highly dissimilar mortality rates during the epidemic. For example, describing the geographic incidence of influenza mortality in Indiana, Ensley (1983, p. 7) writes, "…there is no discernible regional pattern in the severity of the epidemic…. The Northern area had the county with the highest mortality rate (Lake, 8.31), as well as the county with one of the lowest rates (Adams, 1.60)…. Unlike previous epidemics which traveled on a slow east-west axis, the Spanish Lady struck in a sudden, random fashion…." Cities in the same region experienced markedly different influenza mortality rates; for example the death rate from influenza and pneumonia in Chicago was 5.2 per 1,000 population in 1918, compared with a rate of 2.8 per 1,000 in Grand Rapids, Michigan which is less than 200 miles away. In Camden, N.J. the death rate was 12.6 per 1,000; in nearby Philadelphia the death rate was 9.3 per 1,000 (U.S. Bureau of the Census 1922).

The state mortality rates are illustrated in Figure 3, which is a map of the Registration Area states indicating state death rates with various degrees of shading. Few regional patterns are evident in Figure 3, although it appears that the eastern states

experienced somewhat higher death rates on average than did states in other regions. However, the death rates from influenza and pneumonia in Montana and Colorado (9.9 and 10.2 deaths per 1,000 population, respectively) were nearly as high as or exceeded the highest-mortality states in the east: New Jersey (10.0), Maryland (10.4) and Pennsylvania (11.2). The lowest-mortality states were Oregon and Minnesota (5.2 and 5.6, respectively). It is unlikely that differing effectiveness of the public health services across states can explain the variation in state mortality rates, because the public health measures taken by local authorities proved completely ineffective at halting the spread of the virus.[11] Some authors have conjectured that exposure to the first wave of the influenza virus in the spring of 1918 conferred immunity to the second wave, which may explain some of the regional patterns in influenza mortality, but evidence on this issue is unavailable.

These impressionistic observations are also consistent with the statistical evidence on the issue. There appears to be no relationship whatsoever between state-level mortality rates from influenza and pneumonia in 1918-19 and levels of real personal per capita income in either 1910 or 1919-21; the simple linear correlation between flu deaths in 1918-19 and the level of real personal per capita income in 1910 is 0.028, while the correlation between flu deaths and the level of real personal per capita income in 1919-21 is 0.084. Based on this evidence, state-level mortality rates appear to be randomly distributed and do not seem to be related to the level of economic development, climate or geography.

[11] For example, in some cities and towns residents were required to wear surgical masks to protect themselves from the virus. Because the influenza virus can penetrate even tightly woven cloth, this

III. Theoretical Predictions

Theoretical models of economic growth offer conflicting predictions of whether an influenza epidemic, and the accompanying negative shock to population and the labor force, should increase or decrease the rate of growth over the medium and long run. The first difficulty is accurately modeling the nature of such a shock. There was certainly a large negative shock to the population and the labor force in 1918 and 1919, but there is strong reason to believe that this shock affected population and labor force growth rates substantially beyond these two years. Since a large proportion of the deaths occurred among those aged 15 to 44, the epidemic undoubtedly adversely affected family formation and fertility for years after the epidemic. However, data limitations make it difficult to precisely estimate this effect.[12]

Standard neoclassical growth models are based on the assumption that capital accumulation is subject to diminishing marginal returns, thus implying that the long-run rate of growth of income per person depends exclusively on exogenous improvements in technology. However, other factors can influence the rate of growth of income per person (and per worker) during the transition between steady-state growth paths. In the standard Solow (1956) model, the growth rate of income per person (g_y) is approximately equal to:

$$g_y = \left[\alpha * \left(\frac{sY}{K} - \delta - n\right)\right] + [(1-\alpha) * g] \qquad (1)$$

---

measure failed to prevent transmission (Ensley 1983).

[12] Fertility rates decreased substantially across the 1920s. Keyfitz and Flieger (1968) estimate that the total fertility rate of women ages 15 to 49 decreased from 3.378 children in 1919-1921 to 2.547 in 1929-1931. However, there are undoubtedly many reasons for this decline besides the flu epidemic. These factors include urbanization, increasing incomes, restrictions on immigration, increased opportunities for young women, etc.

where $\alpha<1$ is the diminishing-returns-to-capital parameter from the Cobb-Douglas production function, s is the savings rate, Y is total output, K is the capital stock, δ is the depreciation rate of capital, n is the rate of growth of the population (and labor force), and g is the growth of the efficiency of labor due to new technology. That is, the growth rate of output per worker is a weighted average of the growth rates of capital per worker and the efficiency of labor. Rearranging and simplifying equation (1) yields:

$$g_y = g + \left[\alpha * \left(\frac{sY}{K} - (n + g + \delta) \right) \right] \qquad (2)$$

Along the steady-state growth path the term in parentheses above is zero, implying that the growth of output per worker is equal to the growth of labor efficiency. That is, $g_y = g$ along the steady-state growth path. However, it is clear that a decrease in n increases the rate of growth of income per person through capital deepening during the transition. Therefore, a flu epidemic should increase the rate of economic growth during the transition between steady-state growth paths in a standard Solow model.

Mankiw, Romer, and Weil (1992) augmented the standard neoclassical model by including human capital (H) in the production function, along with physical capital (K) and labor (L). In this case, the predicted effect of the flu epidemic on growth is ambiguous since both labor and human capital would be destroyed by the epidemic. The decrease in the rate of growth of L would cause income per person to increase through physical capital deepening. That is, the increase in the K/L ratio would increase the rate of growth of income per person (Y/L) in the medium-run. However, the decrease in the level of human capital would reduce the rate of growth of income per person. The net effect depends on the relative magnitudes of these two effects.

The essence of endogenous growth models is the proposition that investment in the broad sense, which includes investment in physical and human capital, and the production of new knowledge through research and development, does not experience diminishing marginal returns. However, here again, some models predict that an epidemic should increase growth while other models suggest that growth should be negatively affected. Some of the simplest endogenous growth models involve no more than specifying an alternative aggregate production function without diminishing returns to capital. This could be due to externalities that cause the social rate of return to investment to exceed the private rate or learning-by-doing could also cause there to be constant returns. In this case, the aggregate production function can be written as $Y = AK$. In this and similar models, an even temporary decrease in the rate of population growth leads to a permanently higher level of income per person and a higher growth rate of income per person in the medium-run because capital is not subject to diminishing returns.

Another strand of new growth theory seeks to explain productivity growth by emphasizing the production of new ideas; i.e., technological advance comes from things that people do. Romer (1990) first formalized this idea, and the model developed by Grossman and Helpman (1991) is emblematic of this strand of endogenous growth theory. In their model, the rate of innovation and economic growth depends positively on the size of the population. A larger market size increases the profits to innovators in an economy characterized by imperfect competition and monopoly profits. In this model, a flu epidemic that decreases the rate of market expansion also decreases the rate of growth of income per capita.

Jones (2002) developed a model that combines many of the insights discussed above. Production is assumed to be a function of both physical capital (K), human capital (H), and the total stock of available ideas (A). Each individual in the economy is endowed with one unit of time and divides this unit to produce goods, ideas, and human capital. Jones shows that the growth rate of output per worker (g_y) between any two points in time is approximately equal to:

$$g_y = \frac{\alpha}{1-\alpha} g_{\varepsilon K} + g_{\ell Y} + \Psi \Delta \ell_h + \gamma g_{\tilde{\ell} A} + \gamma \tilde{n} \qquad (3)$$

where $\alpha<1$ is the diminishing-returns-to-capital parameter. $g_{\varepsilon K} = \frac{s}{n + g_k + \delta}$ where s is the savings rate for physical capital, n is the population growth rate, g_k is the growth rate of the K/L ratio, and δ is the depreciation rate of physical capital. The other parameters in the model include: $g_{\ell Y}$ (the growth rate of the fraction of the labor force that works to produce goods), ψ (returns to schooling), $\Delta \ell_h$ (annual change in the proportion of the labor force in school), $g_{\tilde{\ell} A}$ (growth of the share of labor employed in research and development), and $\tilde{n}$ (labor force growth rate in the countries on the knowledge frontier). γ is shorthand for a combination of parameters from the ideas production function, but is calibrated to the postwar U.S. economy to be between 0.20 and 0.33.[13]

Along a steady-state growth path, all of the terms in equation (3) are zero with the exception of the last one. That is, in the long-run, growth in output per worker depends exclusively and positively on the labor force growth rate in the countries on the knowledge frontier. However, higher population growth has a negative effect on growth

during the transition between steady-states through the standard neoclassical channel described above. The net effect of a negative population and labor force shock during the transition between steady-states is ambiguous in this model and depends on the relative size of the parameters.

Finally, it may be the case that a market-clearing growth model is not the appropriate way to analyze the effects of a flu epidemic. One data limitation discussed in Section V is that we are forced to examine the growth of real personal per capita income across U.S. states between 1919-21 and 1930. That is, we examine the growth effects immediately after the epidemic. It may be the case that the epidemic caused only temporary disruption to the U.S. economy perhaps as a result of reduced consumer confidence, as well as business failures and bankruptcies caused by the deaths of hundreds of thousands of bread-winners and business owners. States with the highest mortality rates may have experienced the biggest decreases in per capita income by 1919-1921. What we may be witnessing is not a change in trend, but only a return to trend across the 1920s after this temporary shock. We examine this possibility in Section V.E.

IV. Related Literature

Like the ambiguous theoretical predictions of neoclassical and endogenous growth models, empirical studies of demographic catastrophes of the past also provide conflicting evidence on the relationship between population health shocks and growth.[14]

[13] Jones (2002, p. 230).

[14] It has been reasonably well established that long-term improvements in life expectancy are positively correlated with economic growth (see, for example, Arora 2001; Bhargava et al. 2001; Mayer 2001). In contrast to this literature, this paper focuses on the effect of a large, one-time population mortality shock on economic growth.

The most relevant empirical evidence examines the effect of the Black Death in Western Europe and the impact of the 1918 influenza epidemic in India.

The Black Death – a combination of bubonic and pneumonic plagues – killed roughly one-quarter of the West European population between 1348 and 1351, and recurring epidemics continued to inflict high death tolls on the continent over the next quarter-century. The conventional view of the effect of the plague is that it sharply reduced the size of the working population, leading to a rapid increase in real wages for the laboring classes that persisted into the fifteenth century. The impact on per capita income is less clear, however, since wheat prices and rents also fell in the aftermath of the Black Death (Hirshleifer 1987; see also Robbins 1928). Bloom and Mahal (1997a) re-examine the effect of the plague on the wages of unskilled agricultural laborers in England during the epidemics that occurred between 1310 and 1449, and find a positive but statistically insignificant relationship between real wages and population growth in England in this period. The authors conduct a similar exercise for France, again finding a positive but insignificant effect of the plague on wages, and conclude that the evidence fails to support the hypothesis that the Black Death resulted in higher wages for the laboring classes. Given the limited data available to study the relationship between plague mortality and wages (n=6 and n=7 for England and France, respectively) the effect of the Black Death on wages and per capita income remains an unresolved issue.

In the same study Bloom and Mahal also examine the effect of the 1918-19 influenza epidemic on acreage sown per capita in India across 13 Indian provinces. India experienced one of the highest death tolls in the world from the influenza epidemic, with an estimated 17 to 18 million deaths; the epidemic also most severely affected the West,

Northwest and Central regions of India, with significantly lower rates of influenza mortality in the east and south of the country (Mills 1986). As in the case of the plague, the authors find no relationship between the magnitude of the population decline and changes in acreage sown per capita across provinces.

Two other demographic catastrophes merit discussion as well, although the parallels with the influenza epidemic are less clear than in the case of the Black Death: the AIDS epidemic and the recent mortality crisis in the countries of the former Soviet Union. Focusing first on the AIDS epidemic, populations in many developing countries are being decimated by this illness, which has led to declines in life expectancy exceeding 20 years and decreases in population growth rates of 0.6 – 1.5 percent in some countries (Haacker 2001). While the magnitude of the population shock will ultimately be at least as severe as that of the influenza epidemic in many countries, and AIDS – like the 1918 influenza epidemic – is primarily affecting prime-age adults, the AIDS epidemic differs from the influenza epidemic in important ways that likely have significant implications for its effect on economic growth. First, in contrast to the influenza epidemic which claimed victims within a matter of days of infection, AIDS is a slowly evolving disease which can be associated with long periods of reduced productivity, high medical expenditures, and extended periods of care by family members for infected individuals. Second, as discussed previously the influenza epidemic was an exogenous shock in that its mutation into a deadly form in the summer of 1918 and its geographic distribution within and across regions appear to have been randomly determined. The spread and severity of the AIDS epidemic, on the other hand, is likely to be related to

income levels which complicates econometric analyses of its impact with endogeneity problems.

Analysts disagree over the likely magnitude of the effect of AIDS on economic growth. Research simulating the effect of AIDS on growth in African countries suggests that AIDS will reduce GDP by 15-25 percent and per capita income by 0-10 percent in 2010 relative to a no-AIDS scenario (Cuddington 1993a, 1993b, Cuddington and Hancock 1994). Challenging these dire scenarios, Bloom and Mahal (1997b) use data from 51 countries for 1980 through 1992 to empirically test the correlation between AIDS incidence and per capita GDP growth, using two-stage least squares to address the endogeneity problem, and find a statistically insignificant coefficient on the AIDS variable. While the authors conclude on this basis that the impact of AIDS on growth is overestimated, it is also possible that the 1980-1992 period was too early in the epidemic to fully assess the effect of AIDS on growth.

The only other demographic crisis of a similar order of magnitude to the influenza epidemic is the tremendous increase in mortality experienced by the working-age population in the countries of the former Soviet Union in the early 1990s. Using 1989 as a baseline, approximately 3.5 million excess deaths were recorded between 1990 and 1999 in Russia alone, and similar increases in death rates occurred in the Baltics, Belarus and Ukraine.[15] Because this mortality crisis occurred simultaneously with a massive decline in GDP across the region – and was almost surely related at least in part to this collapse – it is difficult to distinguish the effects of the reduced labor force size on growth. One approach is to apply the parameters of an economic growth model

[15] See Bobadilla, Costello and Mitchell (1997) and Becker and Bloom (1998) for comprehensive edited collections of papers on various aspects of mortality in the former Soviet Union.

incorporating health from other cross-country work to Russian demographic data to calculate the effect of reduced life expectancy and labor force growth on GDP growth. Using this strategy, Bloom and Malaney (1998) conclude that Russia's mortality crisis contributed only a small amount – 1/3 of 1 percentage point – to the decline in Russia's GDP from 1990 to 1995.

To summarize, the predictions of neoclassical and endogenous growth models are ambiguous regarding the impact of an exogenous shock to population and labor force on economic growth. The empirical evidence is inconclusive as well, and is problematic in that other historical episodes of mortality shocks are either affected by endogeneity between mortality and income – such as AIDS and the Russian mortality crisis – or lack sufficient statistical evidence to draw clear conclusions. In none of these cases, however, does the evidence suggest a large negative effect of a population shock on growth. While data on the 1918 influenza epidemic in the United States are far from perfect, a study of this episode appears to provide a unique opportunity to analyze the effects of a large exogenous shock to population and labor force size on economic growth.

V. Empirical Evidence from U.S. States

A. Data

One advantage of using data from U.S. states, instead of cross-country data, is that U.S. states are relatively homogenous. As a result, we do not need to include the array of variables to control for institutional differences that have been so difficult to control for in cross-country growth studies.

The dependent variable in the specifications is the growth of real personal per capita income from 1919-1921 to 1930. Easterlin (1957) constructed nominal estimates of state income per capita at twenty-year intervals. Barro and Sala-i-Martin (1992) used these estimates, after deflating the nominal estimates by an aggregate consumer price index, in their study of long-run convergence across U.S. states. However, their procedure assumes that prices were the same across all U.S. states, which was likely not the case. Lindert (1978), however, has constructed real estimates of personal income per capita after taxes from the nominal Easterlin estimates.[16] Instead of using a national consumer price index, Lindert created regional price indexes from a variety of sources to deflate the nominal estimates for each state. More importantly, he provides estimates for both the period immediately after the epidemic (1919-21) and 1930 as well, thus allowing us to examine growth over the 1920s (in order to avoid the difficulties associated with the Great Depression, we omit the 1930s from the empirical analysis).[17]

One limitation of the per capita income data is that we only have observations for two points in time: 1919-1921 and 1930. As a result, we first observe real personal per capita income immediately after the epidemic. It would obviously be preferable to have an annual measure for the years both before and after the epidemic, but this is simply not available at the state level. Another limitation is that the personal income measure

[16] More recently, Mitchener and McLean (1999) have published state personal per capita estimates at six census years adjusted for differences in prices and labor input per capita across states. These data are not used for two reasons. Mitchener and McLean (1999, p. 1025) constructed "relative price indexes that measure, for a particular year, how the price level for any given state deviates from the U.S. average. . ." However, this method does not allow for comparisons over time. Second, they only report estimates for 1920 and 1940, making it impossible to examine only the 1920s before the Great Depression. For our purposes, the choice of the price index, whether based on regional prices like Lindert (1978) or aggregate prices like Barro and Sala-i-Martin (1992), is not likely to matter. Mitchener and McLean (1999, p. 1026) show that "part of the regional dispersion in incomes in the Easterlin data for 1880 and 1900 is due to the effect of differences in regional price levels, but this effect is trivial by 1920 and thereafter."

attributes capital income to the state of the asset holder instead of the state in which the business activity actually occurs. However, Barro and Sala-i-Martin (1992, p. 239) report that the post-World War II results using gross state product (where capital income is attributed to the state where the business activity occurs) and personal income "are nearly equal."

The primary explanatory variable is the number of influenza and pneumonia deaths per 1,000 population in each state in 1918 and 1919 reported by the U.S. Bureau of the Census *Mortality Statistics*. Both influenza and pneumonia are combined in the Census figures because they usually occurred together during the epidemic, and because "it is not believed to be best to study separately influenza and the various forms of pneumonia . . . for doubtless many cases were returned as influenza when the deaths were caused by pneumonia, and vice versa."[18] The primary limitation of the mortality data is that not all states are included in the sample. In 1918, the Registration Area, the area from which the Census Bureau received transcriptions of all death certificates, contained only 78.6 percent of the total estimated U.S. population.[19] With the exception of data from a few cities outside of this area, only 30 states are included in the Registration Area at the outset of the epidemic.[20] While almost 80 percent of the U.S. population is included in the sample, most of the least populous states are excluded. As a result, states in the South and West are underrepresented in the sample although states from all Census regions are included in the sample.

---

[17] See Lindert (1978, Appendix G, pp. 381-390) for a complete description. While data for 1929 would have been preferred, the first year of the Great Depression, 1930, was relatively mild.
[18] U.S. Bureau of the Census, *Mortality Statistics 1919*, p. 28.
[19] Ibid., p. 30.
[20] In 1918, the following states were omitted from the Registration Area: Alabama, Arizona, Arkansas, Delaware, Florida, Georgia, Idaho, Iowa, Mississippi, Nebraska, Nevada, New Mexico, North Dakota, Oklahoma, South Dakota, Texas, West Virginia, and Wyoming.

Economic theory and previous empirical work guided the selection of the other explanatory variables. Density, the number of persons per square mile in each state in 1910, was included to ensure that flu and pneumonia deaths were not simply acting as a proxy for density. Education levels (the share of the population age 6 to 20 in school in 1910) and the share of the population that is foreign-born control for differences in labor force skills across states. We control for convergence by including initial income – real income per capita in 1919-1921 – in the regressions, and also include controls for geographic characteristics of the states such as climate. In addition, a dummy variable for Southern states is included to account for the legacy of slavery; Wright (1986) argued that the legacy of slavery prevented convergence of the U.S. South before the imposition of federal minimum wage laws during the New Deal. Finally, Barro and Sala-i-Martin (1992) find that the beginning of the decade agricultural share of personal income is an important control variable in their study of convergence across U.S. states during the 1920s. They find that states with large agricultural shares grew more slowing during the 1920s as a result of decreasing farm prices and land values. Moreover, this is an important control variable in our study because one might expect agricultural states to be more rural and to suffer lower influenza mortality rates as a result. The precise sources and definitions of all of the variables are listed in Appendix 1.

B. Results

As is evident from Table 1, the size and significance of the flu coefficient is remarkably robust and stable across various empirical specifications. Column (1) reports the general unrestricted model including all of the explanatory variables described above. Density, initial income, the initial agricultural share and the proportion of the population

foreign-born in 1910 are also statistically significant. Note that the flu variable is not just a proxy for density since it remains positive and significant even with density included. The flu coefficient is also robust when the percent urban is included in addition to density.[21]

The specification in Column (2) was obtained by using PcGets (ver. 1.0); a powerful econometric modeling package designed to implement the general-to-specific approach to econometric modeling often associated with the London School of Economics.[22] PcGets automatically selects an undominated, congruent model even though the precise formulation of the econometric relationship is not known *a priori*. While the size of the flu coefficient is quite similar to the general specification in column (1), the standard error is smaller so that it is now significantly different from zero at the one-percent significance level. Initial income, density, foreign-born, and the initial agricultural share are the other robust variables in the restricted model.

The specifications in Columns (3) through (5) illustrate that the size and significance of the flu coefficient is not greatly affected by the inclusion or omission of the other explanatory variables. The regression in Column (3) includes both the level of real income per capita in 1920 and flu deaths. The insignificant and economically small coefficient on initial income suggests that there was no convergence taking place across states in the 1920s without controlling for the initial agricultural share.[23] Column (4) is

[21] With both urban and density included in the specification, the flu coefficient drops from 0.219 to 0.167, but it still remains significant at the 5 percent level.

[22] Phillips (1988) and Hendry (1995) explain the general-to-specific methodology in detail. Hoover and Perez (1999, 2001) show that the general-to-specific modeling strategy often reaches a specification near the true data generating process and out performs many other specification search procedures.

[23] These results are consistent with those reported by Barro and Sala-i-Martin (1992, Table 1, line 14, p. 231) who also find evidence of convergence conditional on the initial agricultural share. It evident that the 1920s were a very difficult time for farmers and that states with larger agricultural sectors grew more slowly during the 1920s.

consistent with Claudia Goldin's (1998) work on the importance of the high school movement and human capital accumulation in the early twentieth-century. Column (5) reports the simple ordinary least-squares relationship between growth and flu deaths. The flu coefficients range between 0.219 and 0.235 and are all statistically significant at the one-percent level with the exception of the most general specification reported in Column (1).

The results reported in Table 1 strongly suggest that influenza and pneumonia deaths are nearly orthogonal to all of the other explanatory variables in the sample. This result formalizes what Crosby (1989, p. 66) suspected when he concluded that "the states with the highest excess mortality rates – Pennsylvania, Montana, Maryland, and Colorado – had little indeed in common economically, climatically or geographically." Regardless of the specification, the size and significance of the flu coefficient remains quite constant.

The results shown in Table 2 replace the total state-level mortality rate from flu and pneumonia with the state-level mortality rates of those in the prime-working ages, defined as those ages 10-44 in 1918-1919.[24] As is evident from Table 2, the size of the flu coefficient is even larger and still statistically significant at or close to the one-percent level in all specifications. The coefficient on flu-prime age ranges from 0.242 in the specification suggested by PcGets [Column (2)] to 0.399 in the simple ordinary least-squares equation [Column (5)]. While the initial agricultural share remains negatively and significantly related to subsequent growth, there is far less support for convergence once flu-prime age is included. The specific model suggested by PcGets does not include initial income, and initial income is insignificant in the most general specification and in

Column (4). Once again, there is some evidence that the lagged schooling share is positively related to subsequent economic growth.

The key finding in Tables 1 and 2 is that both the total mortality rate from influenza and pneumonia in 1918 and 1919, and particularly the mortality rate of those of prime working age, is significantly and positively related to the growth in real income per capita from 1919-1921 to 1930 across U.S. states.

C. Tests of Robustness

There is always the concern that the regressions reported in Tables 1 and 2 contain specification errors, such as omitted variable bias, that could cause the positive and statistically significant impact of the flu on growth to be spurious. However, the diagnostic tests reported in the lower panels of Tables 1 and 2 suggest that none of the equations suffer from non-spherical errors, as there is no evidence of non-normality or autoregressive conditional heteroskedasticity. In addition to these tests based on the residuals themselves, we also include a general test for misspecification: the RESET (regression error specification tests) based on the work of Ramsey (1969). While a significant RESET test could indicate the evidence of omitted variable error or functional form misspecification, there is no evidence for this as reported in Tables 1 and 2.

To further examine the robustness of our results, we replicate the general unrestricted regressions in Table 1 for each decade in the post-World War II period using both contemporaneous influenza and pneumonia death rates and 1918-1919 rates. If influenza and pneumonia deaths are correlated with an omitted variable that generally causes growth, then we should find that the flu variable is significant in other time

[24] While Figure 2 shows that mortality rates peaked in the 15 to 44 age group, different age groupings between the *Census of Mortality* statistics and the *Bureau of the Census* population statistics meant that the

periods as well. The results reported in Table 3 include influenza and pneumonia deaths for the two years immediately proceeding each decade. For example, the 1950s regression includes 1948 and 1949 death rates from influenza and pneumonia, while the 1960s regression includes 1958 and 1959 death rates and so on. Although the flu coefficient is significant at the ten- percent level for the 1950s, it is far from significant for any of the other decades in the postwar period.

The results reported in Table 4 include 1918-1919 influenza and pneumonia death rates in each of the postwar regressions. The 1918-119 flu coefficients are close to zero and far from statistically significant in all of the postwar decades. If influenza mortality rates are correlated with any omitted variable that generally causes growth, such as an agglomeration effect that reflects the interconnectedness of the population, then the flu coefficients in Tables 3 and 4 should be significantly different from zero. We find no evidence that this is the case, thus lending support to the conclusion that the results for the 1920s are not spurious.

D. Simulation Results

The empirical results imply that one more influenza and pneumonia death per 1,000 population increases the average annual rate of growth at least 0.2 percent per year over the 1920s. Is the magnitude of this result economically plausible?

In order to examine this question, we conduct simulations using neoclassical and endogenous growth models to ask whether these models can generate positive growth effects as large as those suggested by the empirical results. For reasonable parameter values, it is difficult to generate a positive growth effect more than one-half as large as

10 to 44 age group had to be examined instead.

that suggested in Tables 1 and 2.[25] However, all of these models assume continuous market clearing and do not allow for cyclical disturbances associated with temporary aggregate supply or aggregate demand shocks.

E. The Epidemic and Recession

It is certainly possible that the epidemic caused real per capita incomes to fall by 1919-1921, so that the observed increase to 1930 is partly or largely a return to an unchanged long-run trend. Unfortunately, there is still great uncertainty regarding the timing and severity of business cycles in the years during and after the epidemic, even at a national level. According to the NBER business cycle chronology, there was a cyclical peak in August 1918 and a trough in March 1919. These dates are almost exactly coincident with the epidemic that began in August 1918 and had nearly run its course by March 1919. There is another peak in January 1920 and a trough in July 1921. Although there is debate regarding how far below trend the economy was in 1919 to 1921, there is no doubt that the economy was generally below trend during these years.[26] Moreover, what matters to us are not the national aggregates, but the differential impact across U.S. states. To measure this differential impact, we use one of the few comprehensive data

[25] We assume that the population growth rate is 1.8 percent per year before the epidemic (the actual growth rate from 1900-1918). Population growth is assumed to diminish to 1.1 percent in 1918 to roughly 0.0 percent in 1919, and remains at 1.5 percent throughout the 1920s. An AK model, with the savings rate set to 15 percent and capital lasting about 9 years on average, increases the average annual growth rate from about 2 percent per year to 2.4 percent across the 1920s. The empirical results predict about a roughly one-percent increase instead of a 0.4 percent increase. The only way to generate a large enough growth effect is for the savings rate to increase as well. For example, if the savings rate increases form 15 to 20 percent, then it is possible to generate growth effects as large as the empirical results suggest. It is certainly possible that the epidemic increased the level of precautionary savings in the economy. We are attempting to quantify this possible effect. Models with diminishing marginal returns predict even smaller growth effects. The complete simulation results are available from the authors.

[26] See Romer (1988) and Balke and Gordon (1989) for a discussion of this period.

sets on business conditions across U.S. states during this time, Dun's business failure rate data.[27]

The results reported in Table 5 show that prime-age influenza mortality is significantly and positively related to business failure rates in 1919 and 1920, but not 1921. However, there is a significant relationship between prime-age mortality and business failures over the entire 1919 to 1921 period. While the prime-age death rate is a significant predictor of business failures, note that total flu mortality is not. The first column of Table 6 replicates the general specification from Table 2, but also includes the business failure rate from 1919 to 1921 as an explanatory variable. The size and significance of the flu coefficient is very similar to Table 2, and while the business failure rate is positive it is insignificant. However, when the prime-age flu mortality rate is excluded from the specification in Column (2) of Table 6, then higher business failure rates from 1919 to 1921 are associated with higher growth rates from 191-1921 to 1930. Columns (3) and (4) begin with the restricted model suggested by PcGets and yields the same basic results. While these results are not conclusive, they suggest that prime-age influenza mortality rates were a contributing factor causing business failures. States with higher business failures from 1919 to 1921 grew more rapidly thereafter, thus strongly suggesting that at least some of the observed growth from 1919-1921 to 1930 was simply a return to trend following the temporary disruption caused by the flu epidemic.

[27] The total number of firms listed is well over 1,000,000 and included most manufacturing, retailing, wholesale, transportation and contracting firms in the U.S. However, the professions, farms, railroads, amusements, one-person services and firms in the "FIRE" sector (finance, insurance, and real estate) are excluded. A failure is defined as a closure leading to or likely to lead to a loss to creditors. Mergers and acquisitions are excluded.

VI. Conclusion

The death toll exacted by the 1918-1919 influenza epidemic was one of the highest ever recorded during a health crisis in world history. The epidemic disproportionately claimed young adults, and, although fueled by the troop movements that accompanied the end of World War I, the geographic distribution of influenza mortality across the United States appears to have been largely random. This exogenous shock to population size provides a unique opportunity to examine the impact of a large population decline on subsequent economic growth in an episode that is largely free of the endogeneity between economic growth and mortality that affects other historical episodes of population health shocks.

This study finds a large, positive effect of mortality from the 1918-1919 influenza epidemic on growth over the next decade across U.S. states in the Registration Area; the results are even stronger when the prime-age death rate rather than the overall death rate is used in the regression analysis. Identical tests of the relationship between influenza mortality and economic growth in subsequent decades across U.S. states indicate that the positive coefficient on influenza mortality is not simply acting as a proxy for the interconnectedness of a region's population which may positively contribute to growth.

While these results may be counterintuitive to one's *a priori* expectations, they are consistent with the predictions of neoclassical and endogenous growth models in which capital deepening is an important component of economic growth in the medium term. However, the magnitude of the estimated coefficients is larger than is suggested by simulations using these market-clearing growth models. Instead, we argue that the influenza epidemic was likely a contributing factor to the immediate post-WWI

recessions. U.S. states with higher influenza mortality also had higher business failures between 1919 and 1921 and were shocked further from trend as a result. At least some of the faster growth between 1919-1921 and 1930 in states with higher influenza mortality likely reflects not a change in trend, but a return to trend after this negative shock. Nevertheless, this epidemic was a large shock that had substantial macroeconomic effects.

References

Arora, Suchit. "Health, Human Productivity, and Long-Term Economic Growth." *The Journal of Economic History* 61 (3), 2001: 699-749.

Atack, Jeremy, and Peter Passell. *A New Economic View of American History*, Second Edition. New York, NY: W.W. Norton and Company, 1994.

Balke, Nathan S., and Robert J. Gordon. "The Estimation of Prewar Gross National Product: Methodology and New Evidence." *Journal of Political Economy* 97, 1989: 38-92.

Barro, Robert J., and Xavier Sala-i-Martin. "Convergence." *Journal of Political Economy* 100 (2), 1992: 223-251.

Becker, Charles M. and David Bloom, guest editors, Special Issue: The Demographic Crisis in the Former Soviet Union, *World Development* 26 (11), November 1998.

Bhargava, Alok, Dean T. Jamison, Lawrence J. Lau, and Christopher J.L. Murray. "Modeling the Effects of Health on Economic Growth." *Journal of Health Economics* 20 (3), 2001: 423-440.

Bloom, David E. and Ajay S. Mahal. "AIDS, Flu, and the Black Death: Impacts on Economic Growth and Well-being." In David E. Bloom and Peter Godwin, eds., *The Economics of HIV and AIDS: The Case of South and South East Asia*. Delhi: Oxford University Press, 1997a, 22-52.

_____. "Does the AIDS Epidemic Threaten Economic Growth?" *Journal of Econometrics* 77 (1), 1997b: 105-24.

Bloom, David E. and Pia N. Malaney. "Macroeconomic Consequences of the Russian Mortality Crisis." *World Development* 26 (11), 1998: 2073-2085.

Bobadilla, José Luis, Christine A. Costello, and Faith Mitchell, editors. *Premature Death in the New Independent States.* Washington, D.C.: National Academy Press, 1997.

Cain, Louis P., and Jonathan R. T. Hughes. *American Economic History*, Fifth edition. Addison-Wesley Publishing, 1997.

Crosby, Alfred W. *America's Forgotten Pandemic, The Influenza of 1918.* Cambridge: Cambridge University Press, 1989.

Cuddington, John T. "Modeling the Macroeconomic Effects of AIDS, With an Application to Tanzania." *The World Bank Economic Review* 7 (2), 1993a: 173-189.

_____. "Further Results on the Macroeconomic Effects of AIDS: The Dualistic, Labor-Surplus Economy." *The World Bank Economic Review* 7 (3), 1993b: 403-417.

Cuddington, John T. and John D. Hancock. "Assessing the Impact of AIDS on the Growth Path of the Malawian Economy." *Journal of Development Economics* 43 (2), 1994, 363-368.

Easterlin, Richard A. "State Income Estimates." In *Population Redistribution and Economic Growth: The United States, 1870-1950*, Vol. 1, edited by Everett S. Lee, Ann Ratner Miller, Carol Brainerd, et al., 702-759. Philadelphia, PA: The American Philosophical Society, 1957.

Ellis, Beth R., ed. *The World Almanac and Book of Facts, 2001*. Mahwah, New Jersey: World Almanac Books.

Ensley, Philip C. "Indiana and the Influenza Pandemic of 1918." *Indiana Medical History Quarterly* 9 (4), 1983: 3–15.

Goldin, Claudia. "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century." *Journal of Economic History* 58 (2), 1998: 345-74.

Grossman, Gene M., and Elhanan Helpman. *Innovation and Growth in the Global Economy*. Cambridge: MA, MIT Press, 1991.

Haacker, Markus. "The Economic Consequences of HIV/AIDS in Southern Africa." IMF Working Paper, December 2001.

Hirshleifer, Jack. *Economic Behaviour in Adversity*. Chicago: University of Chicago Press, 1987.

Hoover, Kevin D., and Stephen J. Perez. "Data Mining Reconsidered: Encompassing and the General-to-Specific Approach to Specification Search." *The Econometrics Journal* 2, 1999: 1-25.

_____. "Truth and Robustness in Cross-country Growth Regressions." California State University, Sacramento manuscript, 2001.

Japan. Statistics Bureau, Management and Coordination Agency. *Historical Statistics of Japan, Volume 1*. Tokyo: Japan Statistical Association, 1945.

Jones, Charles I. "Sources of U.S. Economic Growth in a World of Ideas." *American Economic Review* 92 (1), 2002: 220-239.

Keyfitz, N., and W. Flieger. *World Population: An Analysis of Vital Data*. Chicago: University of Chicago Press, 1968.

Linder, Forrest E. and Robert D. Grove. *Vital Statistics Rates in the United States 1900-1940*. Washington, D.C.: United States Government Printing Office, 1943.

Lindert, Peter H. *Fertility and Scarcity in America.* Princeton: Princeton University Press, 1978.

Luk, Jeffrey, Peter Gross and William W. Thompson. "Observations on Mortality During the 1918 Influenza Pandemic." *Clinical and Infectious Diseases* 33, 2001: 1375-8.

Mankiw, N. Gregory, David Romer, and David Weil. "A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* 107 (2), 1992: 407-437.

Mayer, David. "The Long-Term Impact of Health on Economic Growth in Latin America." *World Development* 29 (6), 2001: 1025-1033.

Miller, Ann Ratner, and Carol P. Brainerd. In *Population Redistribution and Economic Growth: The United States, 1870-1950*, Vol. 1, edited by Everett S. Lee, Ann Ratner Miller, Carol Brainerd, et al. Philadelphia, PA: The American Philosophical Society, 1957.

Mills, I. D. "The 1918-1919 Influenza Pandemic – The Indian Experience." *The Indian Economic and Social History Review* 23 (1), 1986: 1-40.

Mitchener, Kris James, and Ian W. McLean. "U.S. Regional Growth and Convergence, 1880-1980. *Journal of Economic History* 59(4), 1999: 1016-1042.

Neal, Larry. "A Shocking View of Economic History." *Journal of Economic History* 60 (2), 2000: 317-331.

Noymer, Andrew, and Michel Garenne. "The 1918 Influenze Epidemic's Effects on Sex Differentials in Mortality in the United States." *Population and Development Review* 26 (3), 2000: 565-581.

Oxford, John et al. "Who's That Lady?" *Nature Medicine* 5 (12) 1999: 1351-1352.

Patterson, K. David and Gerald F. Pyle. "The Geography and Mortality of the 1918 Influenza Pandemic." *Bulletin of the History of Medicine* 65 (1), 1991: 4-21.

Phillips, Peter C. B. "Reflections on Econometric Methodology." *Economic Record* 64, 1988: 334-59.

Potter, C.W. "A History of Influenza." *Journal of Applied Microbiology* 91, 2001: 572-579.

Ramsey, J.B. "Tests for Specification Error in Classical Least Squares Regression Analysis." *Journal of the Royal Statistical Society* B31, 1969: 250-271.

Reid, A. H., T. G. Fanning, J. V. Hultin, and J. K. Taubenberger. "Origin and Evolution of the 1918 'Spanish' Influenza Virus Hemagglutinin Gene." *Proceedings of the National Academy of Sciences of the USA* 96, 1999: 1651-1656.

_____. "Characterization of the 1918 'Spanish' Influenza Virus Neuraminidase Gene. *Proceedings of the National Academy of Sciences of the USA* 97, 2000: 6785 – 6790.

Reid, Ann H., Jeffrey K. Taubenberger, and Thomas G. Fanning. "The 1918 Spanish Influenza: Integrating History and Biology." *Microbes and Infection* 3, 2001: 81-87.

Rice, Geoffrey W. "Australia and New Zealand in the 1918-19 Influenza Pandemic." *Occasional Papers on Medical History of Australia* 4, 1990: 67-74.

_____. *Black November: The 1918 Influenza Epidemic in New Zealand.* Wellington, New Zealand: Allen & Unwin, 1988.

Robbins, Helen. "A Comparison of the Effects of the Black Death on the Economic Organization of France and England." *Journal of Political Economy* 36 (4), 1928: 447-479.

Romer, Christina D. "World War I and the Postwar Depression: A Reinterpretation Based on Alternative Estimates of GNP." *Journal of Monetary Economics* 22, 1989: 91-115.

Romer, Paul M. "Endogenous Technological Change." *Journal of Political Economy* 95 (5), 1990: S71-S102.

Solow, Robert M. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70 (1), 1956: 65-94.

Sydenstricker, Edgar. "The Incidence of Influenza among Persons of Different Economic Status during the Epidemic of 1918." *Public Health Reports* 46(4), 1931: 154-70.

Taubenberger, J. K, A. I. L. Reid, A. E. Krafft et al. "Initial Genetic Characterization of the 1918 'Spanish' Influenza Virus." *Science* 275, 1997: 1793-1796.

UNAIDS. *Report on the Global HIV/AIDS Epidemic*. December 2001 and June 2000.

Union of South Africa. *Official Yearbook of the Union and of Basutoland, Bechaunaland Protectorate and Swaziland*. (Pretoria, 1921).

U.S. Bureau of the Census. *Mortality Statistics 1919.* Washington, D.C.: Government Printing Office, 1921.

_____. *Mortality Statistics 1920.* Washington, D.C.: Government Printing Office, 1922.

_____. *Historical Statistics of the United States, Colonial Times to 1957.* Washington, D.C.: Government Printing Office, 1960.

_____. *Statistical Abstract of the United States 2001*. (121st edition) Washington, D.C.: Government Printing Office, 2001.

U.S. National Office of Vital Statistics. *Vital Statistics-Special Reports* 43. Washington, D.C.: Government Printing Office, 1956.

Walton, Gary M., and Hugh Rockoff. *History of the American Economy*, Ninth Edition. Cincinnati, OH: South-Western College Publishing, 2001.

White, H. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity." *Econometrica* 48, 1980: 817-838.

Wright, Gavin. *Old South, New South: Revolutions in the Southern Economy*. New York, NY: Basic Books, 1986.

Appendix 1: Data Sources

U.S. States: Interwar Period

Data Name	Description	Source
Growth	Average annual growth rate of real personal income per capita after taxes between 1919-21 and 1930, computed as the log differences and expressed as a percent.	Lindert (1978), Table G-6, p. 390.
Flu	Influenza and pneumonia deaths per 1,000 population in 1918 and 1919.	U.S. Bureau of the Census, *Mortality Statistics 1920*, p. 30.
Flu - prime age	Influenza and pneumonia deaths per 1,000 population aged 10-44 in 1918 and 1919.	U.S. Bureau of the Census, *Mortality Statistics 1918*, and *Mortality Statistics 1919*, Table 8 for both years for the numerator. The denominator [state populations at various ages from Miller and Brainerd (1957), Table L-2].
Initial income	Real personal income per capita after taxes in 1919-21 (1960 consumer dollars).	Lindert (1978), Table G-6, p. 390.
Climate	Number of average annual cooling degree days defined as the number of degrees the air temperature is above 65 degrees Fahrenheit multiplied by the number of days. The average of the entire length (years) of recorded temperatures was used.	*Statistical Abstract of the United States* (2001), Table No. 377. If a state has more than one reporting station, then an arithmetic average of all the stations in that state was computed.
Density	Persons per square mile in 1910.	Total population in 1910 from the *United States Historical Census Data Browser* (http://fisher.lib.virginia.edu/census/).

Foreign	Percent of persons foreign born in 1910.	*United States Historical Census Data Browser* (http://fisher.lib.virginia.edu/census/).
School	Percent of persons age 6-20 in school in 1910	*United States Historical Census Data Browser* (http://fisher.lib.virginia.edu/census/).
South	Dummy variable equal to one if the state was a member of the Confederacy, and zero otherwise.	
Initial agricultural share	Farm income as a percent of personal income in 1920.	Barro and Sala-i-Martin (1992). Data (AGRY20) from http://www.columbia.edu/~xs23/data.htm.
Business failure rate	Number of business failures divided by number of business concerns, expressed as a percent.	*Statistical Abstract of the United States* (1921), pp. 292-293 (original source: *Dun's Review*, NY: NY).

Figure 1. U.S. Male and Female Life Expectancy at Birth, 1900-1945

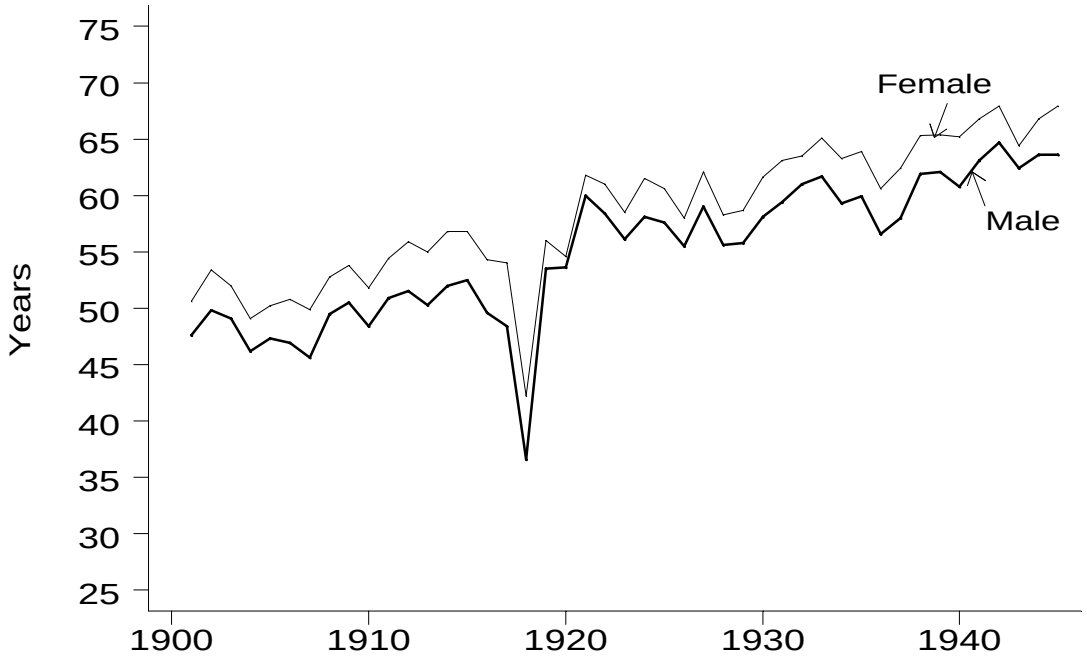

Figure 2a. Age-Specific Death Rates from Influenza and Pneumonia, United States (Deaths Per 100,000 Population in Each Age Group)

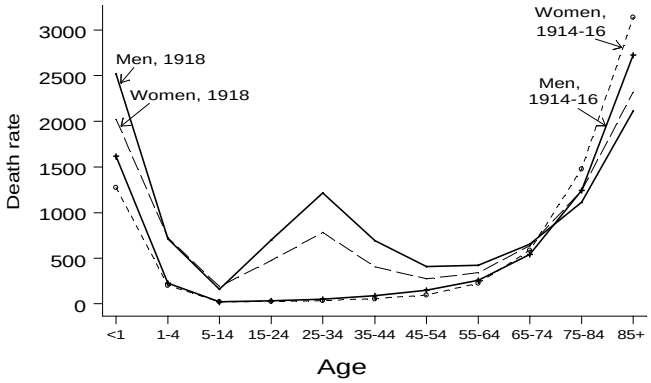

Figure 2b. Men

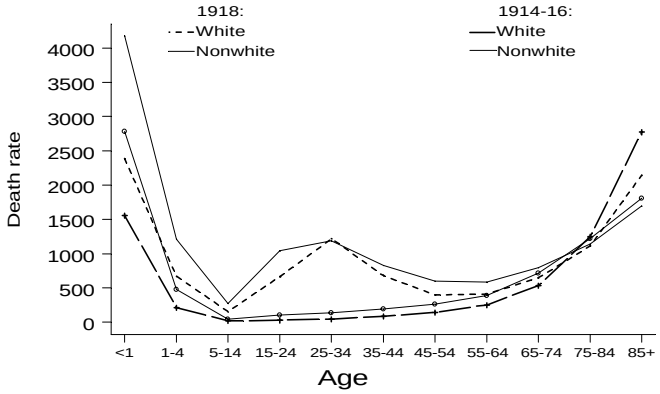

Figure 2c. Women

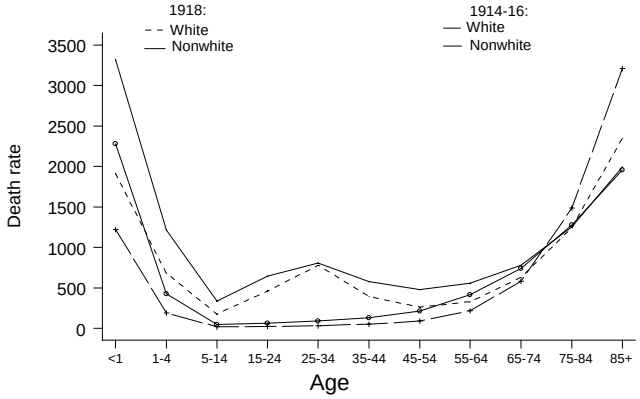

Figure 3. Influenza and Pneumonia Death Rates, 1918-1919

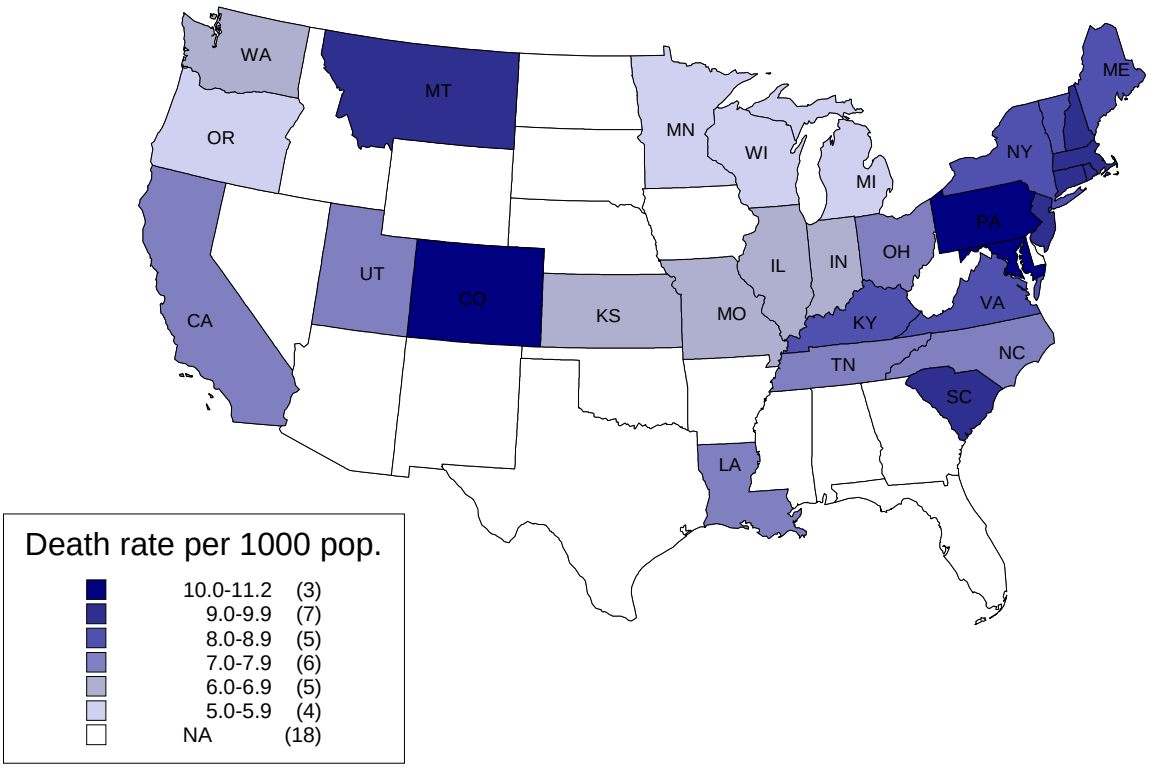

Table 1
Dependent Variable: Average Annual Growth in
Real Personal Per Capita State Income from 1919-21 to 1930

	(1)	(2)	(3)	(4)	(5)
Constant	-2.448		-2.818***	-4.992**	-2.031***
	(3.411)		(0.921)	(1.955)	(0.629)
Flu	0.219**	0.221***	0.222**	0.226***	0.235***
	(0.078)	(0.065)	(0.083)	(0.081)	(0.083)
Initial income	-0.002**	-0.002***	0.001	-0.001	
	(0.001)	(0.0005)	(0.001)	(0.001)	
Climate	0.0004				
	(0.0007)				
Density	-0.003**	-0.0036**			
	(0.001)	(0.0014)			
Foreign	0.080***	0.081***			
	(0.027)	(0.020)			
School	0.045			0.076**	
	(0.044)			(0.032)	
South	-0.235				
	(0.646)				
Initial agricultural share	-0.064**	-0.053***		-0.048*	
	(0.029)	(0.016)		(0.027)	
R^2	0.617	0.576	0.220	0.417	0.155
Adj. R^2	0.471	0.508	0.163	0.324	0.125
Schwarz criterion	2.829	2.476	2.858	2.794	2.826
Durbin-Watson	1.129	1.020	0.886	1.428	1.001
Jarque-Bera normality (p-value)	0.541	0.715	0.168	0.973	0.159
ARCH (p-value)	0.824	0.874	0.931	0.641	0.918
Ramsey RESET (p-value)	0.996	0.487	0.325	0.911	0.801
N	30	30	30	30	30

Notes: White (1980) heteroskedasticity-consistent standard errors are in parentheses. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level, and *** denotes significance at the 1 percent level. See Appendix 1 for the precise definitions and sources of the variables above.

Table 2
Dependent Variable: Average Annual Growth in
Real Personal Per Capita State Income from 1919-21 to 1930

	(1)	(2)	(3)	(4)	(5)
Constant	2.433		-4.587***	-1.651	-3.442***
	(2.833)		(0.910)	(1.493)	(0.869)
Flu-prime age	0.277***	0.242***	0.296**	0.212**	0.399***
	(0.088)	(0.050)	(0.112)	(0.086)	(0.117)
Initial income	-0.0023		0.002***	-0.0015	
	(0.0014)		(0.001)	(0.0012)	
Climate	-0.0004				
	(0.0004)				
Density	-0.005**	-0.004***			
	(0.002)	(0.001)			
Foreign	0.040				
	(0.031)				
School	-0.006			0.043*	
	(0.035)			(0.024)	
South	0.132				
	(0.693)				
Initial agricultural share	-0.110***	-0.092***		-0.086***	
	(0.026)	(0.011)		(0.024)	
R^2	0.729	0.666	0.404	0.626	0.287
Adj. R^2	0.625	0.641	0.360	0.566	0.261
Schwarz criterion	3.182	2.710	3.288	3.048	3.354
Durbin-Watson	1.426	1.253	1.459	1.346	1.303
Jarque-Bera normality (p-value)	0.783	0.732	0.680	0.793	0.697
ARCH (p-value)	0.250	0.663	0.716	0.630	0.015
Ramsey RESET (p-value)	0.936	0.728	0.557	0.793	0.585
N	30	30	30	30	30

Notes: White (1980) heteroskedasticity-consistent standard errors are in parentheses. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level, and *** denotes significance at the 1 percent level. See Appendix 1 for the precise definitions and sources of the variables above.

Table 3
Postwar Regressions
Beginning of the Decade Influenza and Pneumonia Death Rates
Dependent Variable: Average Annual Growth in
Real Personal Per Capita State Income

	1950s	1960s	1970s	1980s	1990s
Constant	3.093**	3.860*	1.030	-2.172	1.794
	(1.525)	(2.023)	(1.979)	(4.467)	(2.302)
Flu	0.649*	-0.010	-0.012	1.250	0.568
	(0.328)	(0.449)	(0.507)	(1.143)	(0.499)
Initial income	-0.0004***	-0.0003***	-0.00010	-0.00017*	-0.00002
	(0.00007)	(0.00007)	(0.0000)	(0.00008)	(0.00003)
Climate	0.00008	0.00008	0.00002	-0.0002*	-0.0002***
	(0.00009)	(0.00006)	(0.00006)	(0.0001)	(0.00006)
Density	-0.0003	0.0007*	-0.0003	0.0018***	0.00005
	(0.0006)	(0.0004)	(0.0003)	(0.0005)	(0.0003)
Foreign	0.031*	0.0146	-0.055	0.059	-0.045***
	(0.017)	(0.019)	(0.034)	(0.043)	(0.016)
School	0.008	0.015	0.032	0.069	0.002
	(0.019)	(0.026)	(0.214)	(0.059)	(0.027)
South	-0.019	0.504**	0.111	0.790***	0.320**
	(0.251)	(0.209)	(0.162)	(0.282)	(0.135)
Initial agricultural share	-0.038***	-0.002	-0.039***	-0.023	0.008
	(0.009)	(0.013)	(0.013)	(0.051)	(0.028)
R^2	0.686	0.715	0.458	0.530	0.523
Adj. R^2	0.622	0.657	0.347	0.434	0.428
Schwarz criterion	1.509	1.259	1.440	2.177	1.116
Durbin-Watson	1.864	2.102	1.498	1.567	1.865
Jarque-Bera normality (p-value)	0.673	0.928	0.000	0.415	0.099
ARCH (p-value)	0.311	0.161	0.457	0.951	0.371
Ramsey RESET (p-value)	0.221	0.161	0.615	0.021	0.176
N	48	48	48	48	49

Notes: White (1980) heteroskedasticity-consistent standard errors are in parentheses. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level, and *** denotes significance at the 1 percent level. See Appendix 1 for the precise definitions and sources of the variables above.

Table 4
Postwar Regressions
1918-1919 Influenza and Pneumonia Death Rates
Dependent Variable: Average Annual Growth in
Real Personal Per Capita State Income

	1950s	1960s	1970s	1980s	1990s
Constant	4.982***	9.492***	-1.856	7.436	-2.143
	(1.727)	(3.131)	(2.862)	(7.138)	(2.093)
Flu (1918-1919)	0.029	0.001	0.064	0.076	-0.031
	(0.057)	(0.043)	(0.038)	(0.083)	(0.049)
Initial income	-0.0003***	-0.00015	0.00001	-0.00003	0.000006
	(0.00009)	(0.00010)	(0.00007)	(0.00010)	(0.00004)
Climate	0.0002	0.00002	0.00004	-0.0005	0.00002
	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0001)
Density	-0.001	0.0004	-0.0005	0.001*	-0.0003
	(0.0008)	(0.0003)	(0.0003)	(0.0006)	(0.0004)
Foreign	0.050**	0.034	-0.072*	0.040	-0.450**
	(0.024)	(0.023)	(0.038)	(0.053)	(0.0176)
School	-0.017	-0.070	0.046	-0.055	0.049**
	(0.019)	(0.042)	(0.032)	(0.085)	(0.236)
South	-0.207	0.292	0.330	0.544	-0.052
	(0.397)	(0.316)	(0.403)	(0.451)	(0.191)
Initial agricultural share	-0.050*	0.024	-0.032	-0.142	-0.058
	(0.027)	(0.033)	(0.023)	(0.156)	(0.081)
R^2	0.619	0.757	0.650	0.555	0.352
Adj. R^2	0.474	0.665	0.517	0.385	0.106
Schwarz criterion	1.608	1.186	1.000	2.303	1.189
Durbin-Watson	1.731	1.621	1.303	2.294	3.025
Jarque-Bera normality (p-value)	0.588	0.175	0.692	0.596	0.091
ARCH (p-value)	0.785	0.353	0.307	0.764	0.982
Ramsey RESET (p-value)	0.001	0.363	0.076	0.013	0.804
N	30	30	30	30	30

Notes: White (1980) heteroskedasticity-consistent standard errors are in parentheses. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level, and *** denotes significance at the 1 percent level. See Appendix 1 for the precise definitions and sources of the variables above.

Table 5
Business Failure Rate Regressions

	Constant	Flu	Constant	Flu-Prime
Dependent variable:				
Business failure rate	0.359	0.009	0.026	0.044**
1919	(0.236)	(0.029)	(0.136)	(0.019)
Business failure rate	0.491*	0.004	0.107	0.046**
1920	(0.273)	(0.033)	(0.162)	(0.022)
Business failure rate	0.729*	0.039	0.560*	0.071
1921	(0.429)	(0.052)	(0.331)	(0.045)
Business failure rate	1.579**	0.052	0.693	0.162**
1919-1921	(0.833)	(0.102)	(0.545)	(0.075)

Notes: * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level, and *** denotes significance at the 1 percent level. See Appendix 1 for the precise definitions and sources of the variables above. All regressions contain 30 observations.

Table 6
Dependent Variable: Average Annual Growth in
Real Personal Per Capita State Income from 1919-21 to 1930

	(1)	(2)	(3)	(4)
Constant	2.414	4.193*		
	(2.739)	(2.387)		
Flu-prime age	0.258**		0.221***	
	(0.095)		(0.073)	
Initial income	-0.0023	-0.003***		
	(0.0014)	(0.0008)		
Climate	-0.0004	-0.0003		
	(0.0004)	(0.0002)		
Density	-0.005**	-0.004**	-0.004***	-0.0023
	(0.002)	(0.002)	(0.001)	(0.0014)
Foreign	0.036	0.044		
	(0.029)	(0.029)		
School	-0.006	-0.007		
	(0.035)	(0.031)		
South	0.082	-0.243		
	(0.723)	(0.582)		
Initial agricultural share	-0.108***	-0.111***	-0.093***	-0.069***
	(0.024)	(0.0215)	(0.011)	(0.010)
Business failure rate (1919-1921)	0.154	0.492**	0.103	0.565***
	(0.228)	(0.208)	(0.289)	(0.152)
R^2	0.735	0.573	0.669	0.454
Adj. R^2	0.615	0.486	0.630	0.429
Schwarz criterion	3.273	3.290	2.811	3.054
Durbin-Watson	1.397	1.490	1.168	1.430
Jarque-Bera normality (p-value)	0.700	0.581	0.608	0.557
ARCH (p-value)	0.178	0.508	0.593	0.926
Ramsey RESET (p-value)	0.929	0.052	0.742	0.436
N	30	30	30	30

Notes: White (1980) heteroskedasticity-consistent standard errors are in parentheses. * denotes significance at the 10 percent level, ** denotes significance at the 5 percent level, and *** denotes significance at the 1 percent level. See Appendix 1 for the precise definitions and sources of the variables above.