Disease and Development: The Effect of Life Expectancy on Economic Growth*

Daron Acemoglu
MIT

Simon Johnson
MIT and IMF

July 2005

Abstract

What is the effect of increasing life expectancy on economic growth? To answer this question, we exploit *the international epidemiological transition*, the wave of international health innovations and improvements that began in the 1940s. We obtain estimates of mortality by disease before the 1940s from the League of Nations and national public health sources. Using these data, we construct an instrument for changes in life expectancy, referred to as *predicted mortality*, which is based on the pre-intervention distribution of mortality from various diseases around the world and dates of global interventions. We document that predicted mortality has a large and robust effect on changes in life expectancy (at birth) starting in 1940, but no effect on changes in life expectancy before the interventions. The instrumented changes in life expectancy have a large effect on population; a 1% increase in life expectancy leads to an increase in population of about 1.5-2%, but a smaller effect on total GDP both initially and over a 40 year horizon. Consequently, there is no evidence that the large exogenous increase in life expectancy has led to a significant increase in economic growth. These results shed doubt on claims that unfavorable health conditions are the root cause of poverty for some nations.

Keywords: disease environment, economic development, economic growth, health, international epidemiological transition, life expectancy, mortality.
JEL Numbers: I10, O40, J11.

Still Preliminary. Please Do Not Circulate Without Permission.

*We thank David Weil and seminar participants at Princeton, Harvard and the NBER summer institute for useful comments. We are also grateful to the staff of the National Library of Medicine and MIT's Retrospective Collection for their assistance.

1 Introduction

Improving health around the world today is an important social objective, which has obvious direct payoffs in terms of longer and better lives for millions.[1] There is also a growing consensus that improving health can have equally large indirect payoffs through accelerating economic growth. This point is forcefully argued by, among others, Bloom and Sachs (1998), Gallup and Sachs (2000), World Health Organization (2001), Alleyne and Cohen (2002), and Bloom and Canning (2005). For example, Bloom and Sachs (1998) argue that wiping out malaria in sub-Saharan Africa could increase that continent's growth rate by as much as 2% a year, and a recent report by the World Health Organization (2001) states unequivocally: "in today's world, poor health has particularly pernicious effects on economic development in sub-Saharan Africa, South Asia, and pockets of high disease and intense poverty elsewhere" (p. 24) and "...extending the coverage of crucial health services... to the world's poor could save millions of lives each year, reduce poverty, spur economic development and promote global security" (p. i).

However, the evidence supporting this recent consensus is not yet conclusive. Although cross-country regression studies show a strong correlation between measures of health (for example, life expectancy or infant mortality) and both the level of economic development and recent economic growth, these studies have not established a causal effect of health and disease environments on economic growth. Since countries suffering from low life expectancy and low health are also disadvantaged in other ways (and often this is the reason for their poor health outcomes), such macro studies may be capturing the negative effect of these other, often omitted, disadvantages. While a range of micro studies demonstrate the importance of health for individual productivity, as discussed below, these studies do not resolve the question of whether health differences are at the root of the large income differences we observe today and whether improvements in health will increase economic growth substantially.

There are at least two reasons why it is important to know whether health has pernicious effects on economic growth. First, understanding whether diseases have a large effect on economic growth is important for the study of the fundamental causes of economic growth (e.g., whether some areas are destined to poverty or whether poor economic performance is a result of their dysfunctional institutions) and for making informed policy decisions to fight poverty.[2] Second, if improving health does not boost growth, then motivating the fight against

[1] See Becker, Phillipson and Soares (2005) and Deaton (2001, 2005) for recent analysis.

[2] In addition to the papers mentioned above, see Gallup, Sachs, and Mellinger (1999) Sachs (2001), McArthur and Sachs (2001), Masters and McMillan (2001), Hall and Jones (1999), Acemoglu, Johnson and Robinson (2001, 2002, and 2003), and Easterly and Levine (2003).

disease in less-developed countries today in terms of large economic growth benefits could backfire when this growth does not materialize.

This paper investigates the effect of health on economic growth. We focus on life expectancy at birth as a general measure of health conditions. Our strategy is to exploit the large convergence in life expectancy (at birth) driven by international health interventions, more effective public-health measures, and the introduction of new chemicals and drugs starting in the 1940s.[3] This episode, which we refer to as the *international epidemiological transition*, led to an unprecedented improvement in life expectancy in a large number of countries.[4] Figure 1 shows this by plotting life expectancy in countries that were initially (circa 1940) poor, middle income and rich.[5] It illustrates that while in the 1930s life expectancy was low in many poor and middle-income countries, this transition brought their levels of life expectancy close to those prevailing in richer parts of the world.[6] As a consequence of these developments, health conditions in many parts of the less-developed world today, though still in dire need of improvement, are significantly better than the corresponding health conditions were in the West at the same stage of development.[7]

The international epidemiological transition provides us with an empirical strategy that can potentially isolate exogenous changes in health conditions. The effect of various health interventions on a country's life expectancy was related to the extent to which that country's population was initially (circa 1940) affected by various specific diseases, for example, tuberculosis, malaria, and pneumonia.

The early data on mortality by disease are available from various international sources, although they have not been widely used in the economics literature for about 50 years.

[3] There were some effective medical and public health innovations prior to 1940. But the positive effects from these innovations were concentrated in richer countries and, in fact, contributed to the large gap in health conditions between rich and poor countries prior to 1940.

[4] The term epidemiological transition was developed by demographers and refers to the entire global process of falling mortality rates after about 1850. We focus here on the rapid decline in mortality in poorer countries after 1940, most of which was driven by the fast spread of new technologies and practices. The seminal works on this episode include Stolnitz (1955), Omran (1971), and Preston (1975).

[5] The rich countries are those with income per capita in 1940 above the level of Argentina (the richest Latin American country at that time, according to Maddison's data, in our base sample). These are, in ascending order, Belgium, Netherlands, Sweden, Denmark, Canada, Germany, Australia, New Zealand, Switerland, the United Kingdom and the United States. The poor countries here are those with income per capita below that of Portugal, which was the poorest European nation (in our base sample). These are, in ascending order: China, Bangladesh, India, Pakistan, Myanmar, Thailand, El Salvador, Honduras, Indonesia, Brazil, Sri Lanka, Malaysia, Nicaragua, Korea, Ecuador, and the Philippines.

[6] Because of data limitations, throughout the paper we focus on countries outside sub-Saharan Africa.

[7] For example, life expectancy at birth in India in 1999 was 60, while life expectancy in Britain in 1820, when income per capita was approximately the same level as in India today, was 40 (Maddison, 2001, p.30). From Maddison (2001, p. 264), income per capita in Britain in 1820 was $1707, while it stands at $1746 in India in 1998 (all figures in 1990 international dollars).

These data allow us to create an instrument for changes in life expectancy based on the pre-intervention distribution of mortality from various diseases around the world and the dates of global intervention (e.g., discovery and mass production of penicillin and streptomycin, or the discovery and widespread use of DDT against mosquito vectors). Therefore, the only source of variation in this instrument, which we refer to as *predicted mortality*, comes from the interaction of baseline cross-country disease prevalence with global intervention dates for specific diseases. We document that this instrument has a large and robust effect on changes in life expectancy (at birth) starting in 1940, but has no effect on changes in life expectancy prior to this date (i.e., before the interventions).

The instrumented changes in life expectancy have a fairly large effect on population; a 1% increase in life expectancy is related to an approximately 1.5-2% increase in population. The magnitude of this estimate suggests that the decline in birth rates was insufficient to compensate for increased life expectancy, a result which we directly confirm by looking at the relationship between life expectancy and total births.

Our main result is that the increase in life expectancy caused by the international epidemiological transition had only a small effect on total GDP. Initially, in fact, there was little change in total GDP in response to the increase in life expectancy and population. After 30 or so years, there was more of an increase in GDP, but this was smaller than the rise in population. Consequently, growth rates for GDP per capita (and GDP per working age population) declined slightly following these large increases in life expectancy. Similarly, we find no evidence of an increase in human capital investments associated with improvements in life expectancy.[8]

Comparing Figure 2, which shows the evolution of income per capita among initially poor, middle-income and rich countries, to Figure 1 gives a glimpse of these patterns. In contrast to the convergence in life expectancy in Figure 1, there is no convergence in income per capita. Our instrumental-variables (IV) regression analysis is essentially a way of further investigating this relationship by focusing on the exogenous component of changes in life expectancy across countries.

The most natural interpretation of these results comes from neoclassical growth theory. The first-order effect of increased life expectancy is to increase population, which initially reduces capital to labor and land to labor ratios, thus depressing income per capita. This initial decline is later compensated by higher output as more people enter the labor force. This compensation can be complete and may even exceed the initial level of income per capita if there are significant

[8] These results also support the view that poor health conditions in some countries are more likely a consequence than a cause of their poverty (see Pritchett and Summers, 1998).

direct productivity benefits from longer life expectancy. Yet, the compensation may also be incomplete if the benefits from higher life expectancy are limited and if scarce factors such as capital or land are important for production. A smaller initial effect on GDP than the longer-run effect is also consistent with the neoclassical growth model when the accumulation of capital is slow.

The importance of scarce factors, especially capital, in the above discussion also suggests that we should expect less negative (or more positive) effects on income per capita in economies that have higher investment rates. We investigate this by estimating models that allow for interactions between life expectancy and initial GDP per capita or initial investment rates (for which the data are weaker), and find some support in favor of this hypothesis.

It is important to emphasize that our findings in no way imply that improved health has not been a great benefit to people in less-developed nations during the postwar era. On the contrary, the results suggest that *it is possible to reduce mortality and improve health conditions dramatically without significant long-run costs in terms of income per capita.* Adopting the approach of Becker, Phillipson and Soares (2005) therefore suggests that these interventions have considerably improved "adjusted income" in these countries. What these interventions have not done, and in fact were not intended to do, is to immediately (or even eventually) enable these countries to produce more output or more output per capita.

Furthermore, our results, though suggestive, may not directly apply to the present date because of the different nature of diseases now prevalent in poor countries, in particular, because of HIV/AIDS. Many of the diseases brought under greater control during the international epidemiological transition were primarily killers of children.[9] In contrast, arguably the most major health problem in the poorest parts of the world today, HIV/AIDS, affects those at the peak of their labor productivity. Preventing HIV/AIDS could conceivably have different effects from those we estimate here (though see Young, 2005).

Finally, it is important to compare our results to those implied by the micro estimates in the literature. The micro development literature has established beyond reasonable doubt that improved health leads to improved individual economic outcomes.[10] These estimates are

[9] Some of these diseases killed infants (under age 1) but many of them had a greater impact on older children (e.g., endemic malaria typically has highest fatality rates for children between ages 1 and 5). While there has been a great deal of convergence in life expectancy at birth between rich and poor countries after 1940, convergence in infant mortality rates has been limited—these have fallen almost everywhere, but still remain very high in many poorer countries today.

The age profile of deaths from tuberculosis pre-1940 was closer to that of AIDS today— with a heavy burden on young adults—but the direct impact on countries' human capital may not have been the same.

[10] See Strauss and Thomas (1998) for an excellent survey of the research until the late 1990s. For some of the more recent research, see Schultz (2002), Bleakley (2004), Miguel and Kremer (2004), and Behrman and Rosenzweig (2004).

difficult to compare with our results, however, since there remains a great deal of uncertainty about the precise size of the relevant effects.[11] Moreover, micro estimates do not directly translate into aggregate quantities because of general equilibrium effects. For example, if there are diminishing returns to effective units of labor (for example because physical capital and land do not adjust perfectly), individual effects will exaggerate the aggregate productivity benefits of increased health. This may be an important concern since our results hint at the presence of such diminishing returns. Moreover, healthier individuals might earn more because they are more successful in their competition with less healthy individuals in the labor market, but when all individuals become healthier, the effects might be much more limited.[12]

In addition to papers mentioned above, our work is related to the literature on the demographic transition both in the West and in the rest of the world, including the seminal work of McKeown (1976) and studies by Preston (1975, 1980), Caldwell (1986), Kelley (1988), Fogel (2004), and Deaton (2003, 2004). A recent paper by Young (2005) evaluates the effect of the HIV/AIDS epidemic in Africa using the neoclassical growth model. He shows that the decline in population resulting from HIV/AIDS may actually increase income per capita despite the significant disruptions and human suffering caused by the disease. See also Arndt and Lewis (2000) and Bell, Devarajan, and Gersbach (2003) on the economic consequences of HIV/AIDS.

The rest of the paper is organized as follows. In the next section, we present a simple model to illustrate the factors that determine the effect of increased life expectancy on economic growth. Section 3 describes the health interventions and the data on disease mortality rates and life expectancy that we constructed from a variety of primary sources. Section 4 presents our estimating framework and the ordinary least square (OLS) relationship between life expectancy and a range of outcomes. Section 5 discusses the construction of our instrument and shows the first-stage relationships, robustness checks, falsification exercises and other supporting evidence. Section 6 presents the main results. Section 7 presents a number of robustness checks and additional results, and Section 8 concludes. Appendices A and B provide information on data sources, data construction and the diseases used in this study. Appendix C, which provides further details, is available upon request.

[11] A recent paper by David Weil (2005) calibrates the effects of health using a range of micro estimates, and finds that these effects could be quite important in the aggregate (see also Bloom and Canning, 2005). Weil's baseline estimate uses the return to the age of menarche from Knaul's (2000) work on Mexico as a general indicator of "overall return to health". Behrman and Rosensweig (2004) obtain smaller estimates using their own results on returns to birthweight differences in monozygotic twins.

[12] See Persico, Postlewaite and Silverman (2005) for evidence suggesting that the major effect of height works through a "competitive advantage" in adolescence.

2 Motivating Theory

To frame the empirical analysis, it is useful to start by outlining the medium-run and long-run implications of increased life expectancy in a simple Solow-type neoclassical growth model. Suppose that economy i has the following constant returns to scale aggregate production function

$$Y_{it} = (A_{it}H_{it})^{\alpha} K_{it}^{\beta} L_{it}^{1-\alpha-\beta}, \tag{1}$$

where $\alpha + \beta \leq 1$, and K_{it} denotes capital, L_{it} denotes the supply of land, and H_{it} is the effective units of labor given by

$$H_{it} = h_{it}N_{it},$$

with N_{it} denoting total population.

To simplify the discussion, let us assume that all agents supply their labor inelastically, so N_{it} is also the supply of labor. Furthermore, we assume that land is inelastically supplied and is constant over time, and without loss of any generality, we normalize it to 1 for all countries, so $L_{it} = L_i = 1$. Throughout, our focus is on cross-country differences, so we ignore technological progress, but we allow for constant technology differences across countries, which could be a function of health differences, so

$$A_{it} = A_i.$$

Capital depreciates at the rate δ and the savings rate of each country is constant at s_i so that

$$K_{it+1} = s_i Y_{it} + (1-\delta) K_{it}.$$

To determine the steady state, suppose that there exists $\bar{t} < \infty$ such that for all $t \geq \bar{t}$,

$$h_{it} = h_i \text{ and } N_{it} = N_i.$$

This implies that there exists a steady state, with $K_{it} = K_i$, satisfying

$$K_i = \frac{s_i}{\delta} Y_i.$$

Substituting into (1), we obtain a simple relationship between income per capita, the savings rate, human capital, technology and population:

$$\frac{Y_i}{N_i} = \left(\frac{s_i}{\delta}\right)^{\frac{\beta}{1-\beta}} (A_i h_i)^{\frac{\alpha}{1-\beta}} N_i^{-\frac{1-\alpha-\beta}{1-\beta}}.$$

Taking logs on both sides, we obtain:

$$\begin{aligned} y_i &\equiv \log\left(\frac{Y_i}{N_i}\right) \\ &= \frac{\alpha}{1-\beta}\log A_i + \frac{\alpha}{1-\beta}\log h_i + \frac{\beta}{1-\beta}\log s_i - \frac{\beta}{1-\beta}\log\delta - \frac{1-\alpha-\beta}{1-\beta}\log N_i. \end{aligned} \tag{2}$$

This equation shows that income per capita is affected positively by technology, human capital and the savings rate, and negatively by population.

For industrialized economies where land plays a small role in production (because only a small fraction of output is produced in agriculture), we can reasonably presume $1-\alpha-\beta \simeq 0$ and population drops out of equation (2). Nevertheless, for many less-developed countries, we may have $1-\alpha-\beta > 0$ and the direct effect of an increase in population may be to reduce income per capita even in the steady state (i.e., even once the capital stock has adjusted to the increase in population).[13]

The overall effect of an increase in life expectancy (and health) may go beyond the direct effect on the increase in population. Better health and longer life spans, for example, may increase productivity through a variety of channels. Let life expectancy be denoted by X_{it} in country i at time t. To capture the beneficial effects of these variables on productivity emphasized in the literature, assume the following iso-elastic relationship:

$$A_{it} = \bar{A}_i X_{it}^{\gamma} \text{ and } h_{it} = \bar{h}_i X_{it}^{\eta}, \tag{3}$$

where $\bar{A}_i$ and $\bar{h}_i$ are some baseline differences across countries.[14] Naturally, greater life expectancy will also lead to greater population (both directly and also potentially indirectly through the birth rate), so we also assume

$$N_{it} = \bar{N}_i X_{it}^{\lambda}. \tag{4}$$

To focus on long run (steady-state) relationships, let us again assume that $X_{it} = X_i$ (at least for $t \geq \bar{t}$ for some $\bar{t} < \infty$), so that there exists a steady state relationship:

$$\begin{aligned} y_i &= \frac{\alpha}{1-\beta}\log \bar{A}_i + \frac{\alpha}{1-\beta}\log \bar{h}_i + \frac{\beta}{1-\beta}\log s_i - \frac{\beta}{1-\beta}\log \delta \\ &\quad -\frac{1-\alpha-\beta}{1-\beta}\log \bar{N}_i + \frac{1}{1-\beta}\left(\alpha\left(\gamma+\eta\right) - \left(1-\alpha-\beta\right)\lambda\right) x_i \end{aligned} \tag{5}$$

where $x_i \equiv \log X_i$ is log life expectancy and recall that $y_i \equiv \log\left(Y_i/N_i\right)$.

An increase in life expectancy will therefore lead to a significant increase in long-run income per capita when there are limited diminishing returns (i.e., $1-\alpha-\beta$ is small) and when life expectancy creates a substantial externality on technology (high γ) and encourages significant

[13] See Galor and Weil (2000), Hansen and Prescott (2002) and Galor (2005) for models in which at different stages of development the relationship between population and income may change because of a change in the composition of output or technology.

[14] On the potential effects of life expectancy and health on productivity, see Bloom and Sachs (1998) and on its effects on human capital see, among others, Kalemli-Ozcan, Ryder and Weil (2000), Kalemli-Ozcan (2002) or Soares (2005).

increases in human capital (high η). On the contrary, when γ and η are small and $1-\alpha-\beta$ is large, an increase in life expectancy can in fact reduce income per capita even in the steady state.

Equation (5) applies to the "long run" once the capital stock has adjusted to the increase in population. It is also interesting to look at what happens to output in the "medium run" where the capital stock is constant (or before it has fully adjusted). This medium-run scenario would be relevant particularly to countries which have low savings rates and can only attract limited foreign capital. To illustrate the issues, let us take an extreme form where the capital stock is fixed at $\bar{K}_i$. In this case

$$\frac{Y_i}{N_i} = \bar{K}_i^{\beta} \left(A_i h_i\right)^{\alpha} N_i^{-(1-\alpha-\beta)}$$

or substituting for (3) and (4), we have:

$$\begin{aligned} y_i \equiv\ & \beta \log \bar{K}_i + \alpha \log \bar{A}_i + \alpha \log \bar{h}_i + \\ & - (1-\alpha-\beta) \log \bar{N}_i + (\alpha(\gamma+\eta) - (1-\alpha)\lambda)\, x_i. \end{aligned} \tag{6}$$

Comparing this equation to equation (5), we see that the medium-run effect of an increase in life expectancy is more negative (or less positive). This is intuitive: before the capital stock adjusts, an increase in population will tend to reduce income per capita.

Our empirical strategy below will be to estimate equations similar to (5) and (6), and compare the estimates to the parameters in these equations.

It is also evident that how quickly an economy approaches the long-run equilibrium depends on its savings and investment rate. Therefore, this framework also suggests that we should investigate the impact of the interaction between life expectancy and the savings/investment rate on the evolution of income per capita.

3 Background and Data

3.1 International Epidemiological Transition

Early improvements in public health began in Western Europe and the United States from the mid-nineteenth century. Initially progress was through empirically observing what worked, but soon came major breakthroughs connected with the *germ theory of disease*. By 1900, tropical medicine had also made impressive progress, most notably with Ronald Ross's demonstration that mosquitoes transmitted malaria and with practical advances against yellow fever in the Caribbean.

However, through 1940 most of the progress in improving mortality was confined to relatively rich countries, with some more limited effects in Southern and Eastern Europe. In most of the Americas, Africa, and Asia, there were very limited improvements.[15] In part, this was because there were few effective drugs against major killers, so most of the measures were relatively expensive public works (e.g., drain swamps). Colonial authorities showed little enthusiasm for such expenditure.

The situation changed dramatically from around 1940 mainly because of four factors. First, there was a wave of global drug innovation. Many of these products offered cures effective against major killers in developing countries. The most important was the discovery and subsequent mass production of penicillin, which provided an effective treatment against a range of bacterial infections (National Academy of Sciences 1970, Easterlin, 1999). A wave of antibiotic development quickly followed, most notably with the discovery of streptomycin, which was effective against tuberculosis.[16] Between 1940 and 1950, most of the major bacterial killers became treatable and, in most cases, curable.[17] Also important was the development of new vaccines, for example, against yellow fever.[18]

The second reason for the dramatic improvement in health was the discovery of DDT, which allowed a major breakthrough in attempts to control one of the major killers of children in less-developed regions of the world, malaria.[19] Aggressive use of inexpensive DDT led to the rapid eradication of malaria in Taiwan, much of the Caribbean, the Balkans, parts of northern

[15] This is not to deny that there was progress before 1940 (partly accounting for the variation in 1940 mortality). During the 1920s and 1930s, there were measures to reduce mortality from smallpox and cholera in Indonesia, smallpox and plague in the Philippines, malaria in India, malaria, respiratory and diarrheal diseases in the British Guyana (see, for example, Preston 1980).

[16] Fleming isolated penicillin in the 1930s but could not produce it in any significant quantity; Florey and Chain made the breakthroughs essential for using pencillin as a drug and they shared the Nobel prize with Fleming in 1945. The first large-scale use of penicillin was in 1943, by Allied armies in North Africa. This was followed quickly by the mass production of penicillin; Andrew Moyer's patent in 1948 is often regarded as the decisive breakthrough.Waksman discovered streptomycin in 1944 and received the Nobel Prize in 1952.

[17] Diseases that could now be treated, without serious side effects for most people, included pneumonia, dysentery, cholera, and venereal diseases. Antibiotics also reduced deaths indirectly caused by (and attributed to) viruses, such as influenza, as these often kill through weakening the immune system and allowing secondary bacterial infections to develop.

[18] The yellow fever vaccine was invented by Max Theiler in 1930 and became widely available in the 1940s. Theiler was awarded a Nobel Prize in 1951. The big wave of vaccine invention followed in the 1950s and 1960s (e.g., against small pox and measles), but antibiotics already provided usually effective treatment against those diseases.

[19] Dichlorodiphenyl trichloroethylene (DDT) was the first chlorinated hydrocarbon insecticide. Desowitz, for example, writes: "There was nothing quite like [DDT] before and has been nothing quite like it since. Here was a chemical that could be sprayed on the walls of a house and for up to six months later any insect that alighted or rested on that wall would die. It was virtually without toxicity to humans. And, for the icing on the chemical cake, it was dirt-cheap to manufacture" (1991, pp. 62-63). DDT was actually first synthesized in 1874, but its discovery is attributed to by Paul H. Müller who received a patent for the insecticide in 1940, and was subsequently awarded a Nobel Prize in 1948.

Africa, northern Australia, large parts of South Pacific, and all but eradicated malaria in Sri Lanka and India (Davis 1956).

The third pillar of the improvements in public health was the establishment of the World Health Organization after World War II, which greatly facilitated the spread of medical and public health technology to poorer countries.[20] From the 1950s, the WHO pushed public health and immunization drives (e.g., against smallpox). In conjunction with the WHO, the US military also played an important role in developing treatments for diseases like cholera and spreading the use of DDT and penicillin.

The fourth factor was a change in international values. As Samuel Preston (1975) emphasizes, after the 1930s, "Universal values assured that health breakthroughs in any country would spread rapidly to all others *where the means for implementation existed*" (p.243).

The consequence of the combination of these four factors was a dramatic improvement in life expectancy in much of the world, especially in the lesser developed parts of the globe, starting in the 1940s. Most of the key changes were available in almost all countries by 1950. Consequently, by the late 1940s and early 1950s, there were significant improvements in health conditions and life expectancy in Central America, South Asia and parts of Eastern and Southern Europe compared to richer countries.[21]

[20] It is notable that Brazil and China, both poor countries at the time, took the initiative in pushing for the formation of the WHO (WHO 1998). A central goal of the organization was to diffuse medical practices and technology to poorer countries. Between the world wars, the League of Nations was largely responsible for international disease interventions, and worked with other European organizations, for example, against typhus in Eastern Europe; see also Office International d'Hygiene Publique (1933). However, in contrast with the WHO, the League of Nations showed less interest in and had only limited resources for combating diseases in less-developed countries, limiting itself to monitoring epidemics that might spread to the West. See League of Nations (1931) for detail on the functioning of the League of Nations Health Organization.

[21] Kingsley Davis (1956) was perhaps the first to write about this in the economics literature, with the aptly titled article, "The Amazing Decline of Mortality in Underdeveloped Areas". He stressed that after that date "these areas do not need to become economically developed to reduce their death rates drastically" (p. 305). The most dramatic drop in death rate was Ceylon—the crude death rate, according to Davis (p. 307), dropped 34 per cent between 1946 and 1947. Similarly large declines during the 1940s were reported for Puerto Rico, Formosa, and Jamaica. Comparing across 18 underdeveloped areas, Davis (p. 307) calculated that the largest fall was between 1945 and 1950 (24.2%), but there was also a large fall of 14% from 1950 to his latest available data (presumably around 1955). In contrast, for the same regions the fall from 1935 to 1940 was 8.3% and the fall from 1940 to 1945 was 5.6%. This pattern had no precedent in richer countries.

Demographers also discussed these developments in detail. Stolnitz (1955) noted: "the middle of this century has marked a revolutionary turning point in the life chances of the world's impoverished nations" (p. 47), while Preston (1975, p.237) argued that: "Factors exogenous to a country's current level of income probably account for 75-90 percent of the growth in life expectancy for the world as a whole between the 1930s and the 1960s" (though later he reduced these estimates). See also Preston (1996).

3.2 Coding Diseases

Central to our empirical strategy is to construct cross-country prevalence and mortality rates for various diseases before the 1940s. For this purpose, we have collected comparable data on the 15 of the most important infectious diseases across a wide range of countries. In all cases, the primary data source is national health statistics, as compiled by the League of Nations (until 1940) and the World Health Organization and the United Nations (after 1945). We have tried several different ways of constructing these data, all of which produce similar results.

We confirm these quantitative assessments of geographic disease incidence with qualitative evidence in the maps and discussion of Cliff, Haggett, and Smallman-Raynor (2004) and the maps of disease incidence published by the American Geographical Society (1951a, b, c, and d) immediately after World War II. Appendix A provides details on sources and construction. Further details are contained in Appendix C. Information on the etiology and epidemiology of each disease is obtained from the comprehensive recent surveys in Kiple (1993). To the extent possible, we have also checked our data against those reported in Preston and Nelson (1974).

The other building block for our approach is intervention dates for each specific disease. These have been obtained from WHO Epidemiological Reports, as well as National Academy of Sciences (1970), Preston (1976), Kiple (1993), Easterlin (1999) and Hoff and Smith (2000).

The 15 diseases we focus on are tuberculosis, malaria, pneumonia, influenza, cholera, typhoid, smallpox, whooping cough, measles, diphtheria, scarlet fever, yellow fever, plague, typhus fever, and dysentery. The most important killers in this list are tuberculosis, malaria and pneumonia, which we discuss here. Information about the remaining diseases is summarized in Appendix B.

Tuberculosis was probably the largest single cause of death around the world in 1940. It is primarily caused by *Mycobacterium tuberculosis*, transmitted through the air. Vaccination had been available from the 1920s, but the breakthrough cure was the 1944 invention of streptomycin. The drug spread quickly and has remained important. In our baseline instrument, we code the intervention for tuberculosis as occuring in the 1940s.

Malaria is caused by four types of parasites, transmitted by the bite of an infected female *Anopheles* mosquito. Control of mosquito vectors had been underway since the late nineteenth century, but became much more effective with the invention of DDT. The use of DDT became widespread in the late 1940s (particularly following a successful demonstration in Greece) and was intensified following the 1955-57 WHO decision to campaign systematically to eradicate malaria (see Bradley 1992 and World Health Organization 2004). In our baseline instrument, the intervention against malaria occurs in the 1940s, but in our alternative instrument it occurs in the 1950s.

Pneumonia is caused by a variety of infectious agents and toxins, including various bacterial and viral pathogens. Frequently, it appears as a secondary bacterial infection that causes death. The primary causes are often tuberculosis, influenza, and more recently AIDS. Antibiotics, for example penicillin, proved highly effective against bacterial pneumonia in the 1940s (although by now resistant strains have developed). Also from the 1940s, there were partially effective vaccines against pneumonia. In our baseline instrument, the intervention against pneumonia takes place in the 1940s.

3.3 Life Expectancy, Population and GDP Data

Other key variables for our investigations include life expectancy at birth, total births, infant mortality, which are all obtained from historical U.N. data (various issues of the *Demographic Yearbook*) and League of Nations reports.[22] Details on the construction of these data are provided in Appendix A.

Since we need data on population in GDP before World War II, we use the data compiled by Maddison (2001). For postwar demographics data we also use UN data sources (see Appendix A).

Our base sample consists of 59 countries, from Western Europe, Oceania, the Americas, Asia, and North Africa. Eastern European and Communist bloc countries are excluded from the base sample, since communist policies may have had differential effects on these countries' population and economic growth in the postwar period. We show robustness results including these countries as well. Africa is excluded throughout because of data availability.[23]

We focus on the period 1940 to 1980 as our base sample. We look at pre–1940 changes as a falsification exercise. Post-1980 is excluded because the emergence of AIDS appears to have led to a divergence in life expectancy between some poor countries and the richer nations.[24] Nevertheless, we provide robustness checks by extending our sample through 2000 (particularly as this allows us to look at longer potential lags in the impact of health on economic outcomes).

Table 1 provides some basic descriptive statistics on the key variables. The first column is for the whole world, while the second column refers to our base sample. A comparison of

[22] All of these data are, at some level, rough estimates. For example, life expectancy is calculated by combining data on age-specific death rates at a point in time, but often approximations are made using standard life tables. Preston (1975) previously used a version of the pre-war data, for 1930 (from *Demographic Yearbooks* in the 1960s).

[23] There are some early data for South Africa, which are only for the white population. We have also found some some scattered reports of mortality in North Africa and sub-Saharan Africa, but these are far from systematic.

[24] Malaria also reappeared in the 1970s and 1980s because of relaxation of international efforts, the international ban on the use of DDT, and the development of insecticide resistant mosquitoes and drug-resistant strains of malaria. Tuberculosis has to some extent returned as a secondary infection associated with AIDS.

these two columns indicates that, despite the absence of sub-Saharan Africa, averages of life expectancy, population, GDP and GDP per capita are similar between the whole world and our sample. The next three columns show numbers separately for the three groups of countries used in Figures 1 and 2 – rich, middle-income and poor countries. These columns show the same patterns as Figures 1 and 2: there is a large convergence in life expectancy among the three groups of countries between 1940 in 1980, but no convergence in GDP per capita. The three columns also give information on predicted mortality, which will be our instrument for life expectancy.

4 Estimating Framework and OLS Estimates

4.1 Estimating Framework

Our empirical approach is to estimate equations similar to equations (5) and (6) above. We interpret these equations as providing the conditional expectation function for our variables of interest. Thus, adding an error term, our estimating equation becomes

$$y_{it+k} = \pi x_{it} + \zeta_i + \mu_t + \mathbf{Z}'_{it}\boldsymbol{\beta} + \varepsilon_{it+k} \tag{7}$$

where y is log income per capita, ζ_i is a fixed effect capturing the technology term and other time-invariant omitted effects, μ_t captures time-varying factors common across all countries, $\mathbf{Z}$ is a vector of other controls, and x is log life expectancy as above. The coefficient π is the parameter of interest. Including a full set of country fixed effects, the ζ_i's, is important, since many country-specific factors will simultaneously affect health and economic outcomes; fixed effects at least remove the time-invariant components of these factors.[25]

Notice also that in equation (7) the left-hand side variable has timing potentially different from the right-hand side variables. This allows us to investigate potential differences between medium-run and long-run effects. In particular, for $k > 0$, this equation would estimate the effect of life expectancy differences at time t on future (date $t+k$) income per capita differences.

[25] Many authors estimate growth regressions of the following form:

$$g_{it} = \tilde{\alpha} y_{it-1} + \pi x_{it-1} + \mathbf{Z}'_{it}\boldsymbol{\beta} + \varepsilon_{it}$$

where y_{it-1} is log income per capita, g_{it} is growth between $t-1$ and t, and x_{it-1} (log) life expectancy at birth or some other measure of health. Since $g_{it} \simeq \Delta y_{it}$, this is equivalent to

$$y_{it} = (1+\tilde{\alpha}) y_{it-1} + \pi x_{it-1} + \mathbf{Z}'_{it}\boldsymbol{\beta} + \varepsilon_{it}$$

This way of writing highlights that the growth regressions are very similar to the levels regressions like (7) or (9), which we estimate. But because such regressions do not include country fixed effects, given the correlation of x_{it-1} with other determinants of income per capita, they are more likely to lead to biased estimates.

Before investigating the effect of life expectancy on income per capita, we look at its effects on population, total births, and total income. The equations for these outcome variables are identical to (7), with the only difference being the dependent variable.

The most serious challenge in estimating the causal effect of life expectancy on income per capita or population is potential omitted variable bias and reverse causality. In particular, in equation (7), typically the population covariance term $\text{Cov}(x_{it}, \varepsilon_{it+k})$ is not equal to 0, because even conditional on fixed effects, health could be endogenous to economics.

As mentioned in the Introduction, our strategy is to exploit the potentially-exogenous source of variation in life expectancy because of global interventions. More specifically, our first-stage relationship is

$$x_{it} = \psi M_{it}^I + \tilde{\zeta}_i + \tilde{\mu}_t + \mathbf{Z}_{it}'\tilde{\boldsymbol{\beta}} + u_{it} \tag{8}$$

where M_{it}^I is *predicted mortality* driven by "exogenous" factors. The identifying assumption (exclusion restriction) is $\text{Cov}(M_{it}^I, \varepsilon_{it+k}) = 0$. Naturally, the plausibility of this exclusion restriction depends on M_{it}^I, which will be described in the next section. Before doing this, we present some basic OLS estimates.

Our main estimating equation, (7), does not allow for mean-reverting dynamics in the outcome variable (for example in income per capita). A more general model would be:

$$y_{it+k} = \rho y_{it-1} + \pi x_{it} + \zeta_i + \mu_t + \mathbf{Z}_{it}'\boldsymbol{\beta} + \varepsilon_{it+k}^m. \tag{9}$$

Though conceptually attractive, this equation is considerably harder to estimate because of the simultaneous presence of fixed effects and a lagged dependent variable (see, for example, Wooldridge, 2002). This motivates our initial focus on (7). Even if the true model is (9), instrumental-variables estimate of (7) will lead to consistent estimates of π as long as $\text{Cov}(M_{it}^I, \varepsilon_{it+k}^m) = 0$. We also report estimates of equation (9) in Section 7.

4.2 OLS Estimates

Tables 2 and 3 report OLS regressions of the main variables of interest. These results are useful as comparison to the instrumental variables (IV) estimates that will be reported in the rest of the paper. All regressions in these tables and throughout the paper include a full set of year dummies and country fixed effects, so all estimates exploit only the within-country variation.

Table 2 focuses on log population (Panels A and B) and log number of births (Panels C and D) on log life expectancy. Throughout the paper, we report results in pairs: first, we estimate versions of equation (7) on our baseline panel, which consists of observations at 10 year intervals between the indicated dates (1940-1980, 1930-1980, etc.). Second, we estimate "long-difference" models, essentially the same equation using only two data points at the beginning

and the end of the sample. The first approach uses all the available data, while the second approach exploits only the longer-run changes. This may be useful both because it may be less vulnerable to problems caused by autocorrelation in the error term (a potential problem which we also address more directly later), and also because it enables us to be more agnostic on how quickly life expectancy should affect the outcome variables. Also for comparison with other OLS models, we report regressions using the standard postwar period 1960 to 2000.

Focusing on the population results, a number of features are notable. First, the 1960-2000 sample gives very similar results to our baseline sample of 1940-1980. For example, for the panel between 1960 and 2000, the estimate of the effect of log life expectancy on log population is between 1.46 and 1.69 (standard errors 0.29 and 0.43 respectively), whereas the estimate for our base sample of 1940-1980 is 1.11 (standard error = 0.21). Second, when we exclude the richest countries from the sample in column 4, this makes little difference. Now the estimate is 1.03 (standard error = 0 .28). Third, in columns 5-10, we look at the effect of life expectancy on future levels of population. In terms of equation (7), this corresponds to the case where $k > 0$. This enables us to investigate the effect of life expectancy on future changes in population. These results are broadly similar to the contemporaneous ones. In all cases, a 1 percent increase in life expectancy is associated with approximately a 1-1.34 percent increase in population. The estimates using the long-differences in Panel B are slightly larger (and slightly less precise), but broadly similar.

To interpret the effect of (log) life expectancy on (log) population, it is useful to consider a simple mechanical model. Suppose each individual faces a Poisson death rate of $1/a$. Life expectancy is then equal to a. Assume also that the flow rate of births is $B(a)$ (with constant birth rate corresponding to $B(a)$ proportional to a). Then steady state population is given by

$$\ln L = \ln a + \ln B(a).$$

This implies that we should expect an elasticity of 1 when the total number of births remain constant in response to an increase in life expectancy. Naturally if there is no change in the birth rate, there will be an increase in the total number of births, and the elasticity we obtain suggests that the birth rate did not declined enough to reduce or keep constant the number of births. This is confirmed in Panels C and D of Table 2, which show an overall increase in the total number of births in response to change in life expectancy.

Table 3 presents results that are parallel to those in Table 2, but now the dependent variables are log GDP (Panels A and B) and log GDP per capita (panels C and D). Again, all regressions have a full set of country and time fixed effects, and we show both panel results and long differences.

Panels A and B in Table 3 indicate a positive relationship between log life expectancy and log GDP. For example, the results in columns 1-5 indicate an effect of life expectancy on GDP with an elasticity of approximately 0.7-1.7.[26]

Columns 5-10 again look at leads. With the exception of column 6, which corresponds to 20-year lead for all countries, the estimates are similar to those in columns 1-4. Overall, the results in Table 3 suggest the presence of a positive and typically significant effect of life expectancy on total GDP. Nevertheless, as pointed out above, these results do not correspond to the causal effect of life expectancy on total output, and might reflect the fact that life expectancy increases precisely when countries are adopting a range of other measures that increase income, or alternatively, as emphasized by demographers, it may be the increase in income per capita that is increasing life expectancy.

While Panels A and B show a positive relationship between life expectancy and total income, the rest of Table 3 suggests that positive effect on population size outweighs the increase in GDP per capita; the net effect on GDP per capita, though typically not significant, is generally negative. There is no evidence of a positive effect of life expectancy on GDP per capita in Table 3. Nevertheless, since these estimates are not necessarily causal, the true effect of life expectancy on income per capita might actually be larger or smaller than those shown in Table 3. The rest of the paper investigates this question.

5 Predicted Mortality and First-Stages

5.1 The Predicted Mortality Instrument

Prior to the international epidemiological transition, there was considerable variation in the prevalence of diseases across the world. For example, while malaria was endemic in parts of South Asia and Central America, in 1940 it was relatively rare in much of Western Europe and in the Southern Cone of Latin America. We may therefore expect variation in the effects from global interventions on life expectancy in different countries depending on the baseline distribution of diseases. For example, DDT should reduce malarial infections and mortality, and increase life expectancy in Central America and South Asia relative to Western Europe or the Southern Cone of Latin America.

[26] Interestingly, the correlation between life expectancy and income per capita in the period 1960-2000 appears to be larger (about twice) compared with that during our base sample period; 1.70 versus 0.82). This is consistent with the fact that a large part of the variation in life expectancy during our base sample period is exogenous, driven by the international epidemiological transition, so the upward bias in the OLS estimate resulting from common shocks to income per capita and health will have less effect during this period.

Motivated by this reasoning, our instrument, *predicted mortality,* is constructed as

$$M_{it}^{I} = \sum_{d \in \mathcal{D}} \left((1 - \Delta_{dt}) M_{di,30} + \Delta_{dt} M_{dFt} \right), \tag{10}$$

where M_{dit} denotes mortality in country i from disease d at time t, Δ_{dt} is a dummy for intervention for disease d at time t (it is equal to 1 for all dates after the intervention), and $\mathcal{D}$ includes the 15 diseases listed above. It is measured as the number of deaths per 100 individuals per annum (for ease of exposition, we transform it from the original numbers which are per 100,000). $M_{di,30}$ refers to the pre-intervention mortality from this disease in the same units, while M_{dFt} is the mortality rate from disease d at the *health frontier* of the world at time t. In our baseline instrument, we take M_{dFt} to be equal to zero. As a robustness checks, we also calculate an alternative measure of predicted mortality using the average mortality rate from disease d at time t among the richest countries, but since these rates are very close to zero, this alternative measure is very similar to our baseline predicted mortality series.

Equation (10) therefore creates a predicted mortality series, M_{it}^{I}, which uses the baseline mortality rate from the 15 diseases in our list in the country in question until there is a global intervention, and after the global intervention, the mortality rate from each disease goes down to the frontier mortality rate.

This expression makes it clear that the only source of variation in predicted mortality comes from the interaction of baseline distribution of diseases with global interventions (in particular, note that M_{di30} applies until the time of global intervention). Whether a country has successfully eradicated a disease or has been quick at adopting international technologies will have no effect on M_{it}^{I}. The dummy Δ_{dt} turns on for all countries at the same time. This makes our exclusion restriction, $\text{Cov}(M_{it}^{I}, \varepsilon_{it+k}) = 0$, with ε_{it+k} as the error term in the second stage equation (for population or GDP etc.), plausible. Since variations in M_{it}^{I} are unrelated to any actions or economic events in the country, there is no obvious reason for it to be correlated with economic or population shocks in the country in question. Nevertheless, exclusion restrictions are, by definition, not testable, so we will try to also support our exclusion restriction by conducting a range of falsification exercises below.

5.2 Alternative Instruments

We also construct a number of alternative instruments to investigate the robustness of our results. Our first strategy is to construct an alternative version of M_{it}^{I} using only the three big killers, which also have the advantage of having clearly marked interventions: malaria (DDT, from the 1940s), tuberculosis and pneumonia (both antibiotics, from the 1940s). We leave influenza out of this "big 3" classification as our sources do not break down the deaths from

viral influenza, so the timing of the key intervention is less clear. This alternative instrument is constructed in the same way as it equation (10), with the only difference that the set $\mathcal{D}$ only includes these big three killers. We check all of our results with this alternative instrument.

Second, we construct an alternative instrument using different timings of interventions when there is reasonable doubt about the exact dates as discussed in Appendix B. Third, as noted in the previous subsection, we create an alternative instrument setting M_{dFt} equal to average mortality among the richest countries. Both of these alternatives give very similar results to our baseline instrument.

Finally, M_{it}^{I} does not use any information about the differential timing of interventions in different countries. As noted above, this makes it easier for us to defend our exclusion restriction. If interventions were country-specific, there would be a concern that countries that are adopting a range of growth-enhancing policies may also be more likely to adopt the global health interventions more rapidly. It is nonetheless informative to look at the variation resulting from different timings of interventions, and we construct an alternative instrument exploiting country-specific interventions as:

$$M_{it}^{I} = \sum_{d \in \mathcal{D}} \left((1 - \Delta_{idt}) M_{di30} + \Delta_{idt} M_{dFt} \right),$$

where the only difference is that the intervention dummy Δ_{idt} is now country specific. We check the robustness of our results using this alternative instrument.

5.3 Zeroth-Stage Estimates

Our approach is predicated on the notion that global interventions reduce mortality from various diseases. Therefore, before documenting the first-stage relationship between our predicted mortality measure and log life expectancy, we would like to show the effect of various global interventions on mortality from specific diseases. In this exercise, in addition to the 15 diseases above, we also use deaths from cancers and malignant tumors as control diseases, since these were not affected by the global interventions.

Figure 3 shows the effect of the global interventions on worldwide deaths from tuberculosis and pneumonia, using deaths from cancers and malignant tumors for comparison (all numbers are unweighted averages).[27] In this figure and in Table 4, mortality rates are expressed per 100,000 (instead of per 100). For both infectious diseases, there are large declines in mortality

[27] For malaria, average mortality rates are substantially lower than tuberculosis and pneumonia, partly because large areas of the world were not affected by malaria. Unweighted average mortality rates from malaria were roughly constant before 1940 and declined starting in the 1940s and continuing to do so through the 1950s. By 1960, there was little malaria-related mortality even places such as India, which had previously had a serious problem with malaria.

(relative to cancer and tumors) between 1940 and 1950, which is precisely the time of global intervention for these diseases. For example, for tuberculosis, average mortality rates are approximately constant between 1930 and 1940 and show a big decline between 1940 and 1950, followed by a much smaller decline after 1950. For pneumonia, we do not have pre-1940 data, so the figure only shows a large decline between 1940 in 1950, followed by stabilization after 1950.

Though suggestive, Figure 3 does not use information about changes in death rates at the country level. To investigate this issue further, Panel A of Table 4 estimates the following "zeroth stage regression":

$$M_{idt} = \theta \Delta_{dt} + \mu_t + \pi_d + \delta_i + v_{it}. \tag{11}$$

The dependent variable here is mortality in country i from disease d at time t, and the regression equation includes a full set of time, disease, and country dummies. The coefficient of interest is θ, which measures whether there is a decline in mortality from a specific disease associated with an intervention. If there is, we would estimate that $\theta < 0$.

Table 4 reports estimates of equation (11). In all cases the estimate of θ is negative and significant. For example, in column 1, θ is estimated to be -46.04 (standard error = 9.40), which indicates an average of a reduction of 46 per 100,000 deaths due to the interventions. In column 2, when we add lagged intervention, the coefficient on the intervention dummy is largely unchanged (-43.33). More challenging for us is the specification in column 3, which includes contemporaneous and lead interventions. Thus this specification investigates whether it is the interventions or pre-existing trends that are responsible for the declines in mortality. Reassuringly, the estimate of the coefficient on contemporaneous intervention, θ, is largely unchanged, -46.04 (standard error = 9.40), while lead intervention has the opposite sign.

Columns 4-7 investigate whether one of the main diseases is responsible for the results in columns 1-3 by excluding tuberculosis, pneumonia, malaria and influenza one at a time. Without tuberculosis or pneumonia, the coefficient estimates are somewhat smaller, but still highly significant (-33.93 and -36.31, standard errors 8.66 and 8.99, respectively). Without malaria or influenza, the coefficient estimates are very similar to the baseline.

In Panel B, we look at each disease separately. The estimates in this case show how effective various interventions have been in reducing mortality from each specific disease. For example, the coefficient of -108.51 for tuberculosis in column 4 and -137.92 for pneumonia in column 5 show the large declines in tuberculosis and pneumonia mortality resulting from the introduction of antibiotics. The estimate of -19.97 in column 6 shows a significant decline in malaria mortality, but also confirms that mortality from malaria was less important for our entire sample than mortality from tuberculosis or pneumonia. The declines in mortality from

the other diseases are even smaller, but, with the exception of influenza, always statistically significant.[28]

5.4 First-Stage Estimates

We next turn to the first-stage relationship between life expectancy and predicted mortality. While the zeroth stage regression in equation (11) is at the disease-country-time level, our first-stage relationship is at the country-time level, since the left-hand side variable is life expectancy (at birth).

Figure 4 shows the first-stage relationship visually. The horizontal axis is the change in predicted mortality between 1940 and 1980, while the vertical axis is the change in log life expectancy during the same time period. We focus on the 1940-1980 period, since 1940 represents a pre-intervention year and 1980 is the end of the sample for most of our specifications. A strong negative relationship is clearly visible in Figure 4. Predicted mortality declined by a large amount in India, the Philippines, Indonesia, and parts of Central America, while remaining largely constant in parts of Western Europe, Uruguay, Argentina, Korea, Australia, and New Zealand. Life expectancy, in turn, increases by a large amount in the first group of countries, and much less in the second period

Figure 5 shows that the same relationship holds when we remove the richest countries from the sample. It thus indicates that the first-stage relationship is not simply a result of a contrast between rich and poor countries.

Table 5 shows the first-stage relationship in regression form by estimating equation (8). Country and year dummies are again included, and this set of specifications do not include any covariates. The top panel uses our entire data starting from either 1940 or 1930, while the bottom panel reports the long difference specifications.

The first column is our baseline specification. It shows an estimate of ψ equal to -0.35 with a standard error of 0.06, which is significant at less than 1%. This estimate implies that an improvement of 0.43 (per 100 or 430 per 100,000 p.a., which is the mean improvement between 1940 and 1950 in our base sample) in predicted mortality leads to approximately a 15 percent increase in life expectancy (mean life expectancy in our sample in 1940 was 49.29, so this is an increase of about 7.1 years, while the mean improvement in life expectancy between 1940 and 1950 is 5.27 years).

With long differences, the coefficient estimate is -0.47, which is somewhat larger, but also slightly less precisely estimated (standard error = 0.09). Results are also similar when we use the slightly longer sample 1930-1980 or the slightly shorter one, 1940-1970. Column 4 shows

[28] There is also not a significant coefficient if we run measles by itself.

similar results when we include Eastern Europe, and column 5 shows a similar estimate when the rich countries are excluded (in all cases the coefficient is about -0.28, and is significant at less than 1%). The corresponding specifications with long differences show more variability in the estimate, but are generally similar.

Column 6 shows that limiting the sample to a balanced panel makes little difference. The estimate of ψ is now -0.33 (standard error= 0.06).

Columns 7-10 investigate robustness to alternative instruments. Columns 7 and 8 use the instrument constructed only from information on tuberculosis, malaria and pneumonia. Both for the base sample and for the sample excluding the richest countries, the results are very similar with this alternative instrument. Column 9 uses the instrument with alternative timing, with little effect on the estimate. Finally, column 10 uses information on country have specific interventions, with again very similar estimates to those in column 1.

Overall, the results in Table 5 show a large and relatively robust effect of the predicted mortality instrument on life expectancy. Tables 6 and 7 investigate the robustness of this finding, while Table 8 looks at pre-existing trends.

5.5 Robustness: Importance of Disease Composition

Table 6 investigates the importance of disease composition to see whether a specific disease is responsible for the first-stage relationships shown in Figures 4 and 5 and in Table 5. As in earlier tables, there are two panels, the bottom panel corresponding to long differences. Throughout this table, the sample period starts in 1940, as in our base case. For comparison, column 1 repeats the base result from Table 5, which uses data on deaths from all 15 diseases to construct predicted mortality.

Columns 2, 3 and 4 present results dropping data on the three main killers from our predicted mortality measure: tuberculosis, malaria and pneumonia respectively. Dropping tuberculosis actually strengthens the first stage estimate slightly—in the long difference specification this is now -0.62, with a standard error of 0.1. Dropping malaria and pneumonia strengthen our first stage only marginally. None of the other diseases has a significant impact on the first stage coefficient. The bottom panel, which looks at long-difference specifications, also confirms this result. We conclude from these results that the first-stage relationship does not reflect the impact of a single disease.

5.6 Robustness: Mean Reversion and Timing

The specifications in Tables 5 and 6 do not allow for mean reversion in life expectancy, and also assume that it is the contemporaneous predicted mortality that affects life expectancy. These

assumptions raise a number of important questions. First, mean reversion may significantly change the relationship between predicted mortality and life expectancy. Second, in more general specifications we may find that it is lags or leads of predicted mortality that affect life expectancy. In particular, if it is the leads of (future changes in) predicted mortality that affect life expectancy, this would shed considerable doubt on our interpretation of the first-stage relationship. Table 7 investigates these issues.

The top panel is for the entire sample, while the bottom panel looks at a sample that excludes the richest countries. Column 1 replicates our baseline specification. Column 2 reports OLS estimates from the following model:

$$x_{it} = \nu x_{it-1} + \psi M_{it}^I + \delta_i' + \mu_t' + u_{it}, \tag{12}$$

which allows lagged log life expectancy to affect current log life expectancy. The regression finds evidence for mean reversion. The coefficient ν in the top panel is estimated to be 0.46 (standard error = 0.09). Nevertheless, the negative relationship between predicted mortality and life expectancy remains. The parameter of interest, ψ, is now estimated at -0.23, which is about 50% smaller than the baseline estimate, but the standard error is also smaller, 0.06. The results for the sample that excludes the initially rich countries in Panel B are similar. For example, in column 2 the estimate of ν is 2.37 (standard error = 0 .11), while the estimate of the impact of predicted mortality on life expectancy, ψ, is -0.20 (standard error = 0.08).

Because we have a relatively short panel, OLS estimation of (12) will lead to inconsistent estimates, however. To deal with this problem, we follow the method of Anderson and Hsiao (1992) in column 3. This involves first-differencing (12), to obtain:

$$\Delta x_{it} = \nu \Delta x_{it-1} + \psi \Delta M_{it}^I + \Delta \mu_t' + \Delta u_{it},$$

where the fixed country effects are removed as a result of differencing. Although this equation cannot be estimated consistently by OLS either, in the absence of serial correlation in the original residual, u_{it} (i.e., no second order serial correlation in Δu_{it}), x_{it-2} is uncorrelated with Δu_{it}, so can be used as instrument for Δx_{it-1} to obtain consistent estimates. Similarly M_{it-1}^I is used as an instrument for M_{it}^I. This estimation procedure leads to very similar results to the OLS estimation. The estimate of ψ in the full sample is -0.30 (standard error = 0 .10), while in the sample excluding the richest countries, it is -0.38 (standard error = 0.11).

Although the instrumental variable estimator of Anderson and Hsiao (1982) leads to consistent estimates, it is not efficient, since, under the assumption of no further serial correlation in u_{it}, not only x_{it-2}, but all further lags of x_{it} are uncorrelated with Δu_{it}, and can also be used as additional instruments. Arellano and Bond (1991) develop a Generalized Method-of-Moments (GMM) estimator using all of these moment conditions. When all these moment

conditions are valid, this GMM estimator is more efficient than the Anderson and Hsiao's (1982) estimator. GMM estimation, which we use in column 4, leads to similar but more precisely estimated coefficients. The estimate of ψ in the full sample is now -0.21 (standard error = 0.05) and in the sample excluding the richest countries, it is -0.17 (standard error = 0.07). Tests for second-order autocorrelation in residuals, reported at the bottom, shows that there is no evidence of additional serial correlation in the residuals. However, the Hansen J-test shows that the overidentification restrictions are rejected, presumably, because different lags of life expectancy lead to a different estimates of the mean reversion coefficient. This rejection is not a major concern for our empirical strategy since the exact magnitude of the mean reversion coefficient, ν, is not of direct interest to us, and the models in (8) and (12) will be the first stage and all we need is that M^I_{it-1} should have no direct effect on the second-stage outcomes.

Columns 5-8 investigate the effect of lagged and lead mortality. In column 5, contemporaneous and lagged mortality are included together. While they are both significant, contemporaneous predicted mortality is larger. The fact that lagged mortality also affects life expectancy is not surprising since many of the interventions were implemented slowly over time.

The more important challenge for our approach is the inclusion of lead predicted mortality. Since the global interventions did not start before 1940, lead mortality should have no effect on life expectancy. Column 6 investigates this by including contemporaneous and lead mortality together. Reassuringly, lead mortality has a very small and highly insignificant coefficient while the estimate of the effect of contemporaneous predicted mortality is the same as our baseline estimate. Column 7 repeats this regression by also including lagged life expectancy on the right hand side. Once again, lead mortality has no effect on life expectancy. These results suggest that, consistent with our hypothesis, it was indeed the global interventions of the 1940s onwards that led to the increase in life expectancy in countries previously affected by these diseases. In other words, they show no evidence of pre-existing trends in life expectancy that are being picked up by our predicted mortality instrument. The issue of pre-existing trends is further investigated in Table 8.

Finally, column 9 investigates whether controlling for the effect of income per capita affects the relationship between predicted mortality and life expectancy, and column 10 reports on the balanced panel. In both cases the results are very similar to our baseline estimates.

5.7 Robustness: Pre-Existing Trends and Falsification

Table 7 already showed that life expectancy responds to contemporaneous changes in predicted mortality and does not respond to future changes. This suggests that our first stage is unlikely to be driven by some pre-existing trends. Nevertheless, the exercise in Table 7 uses only data

from the post-1930 period. An alternative falsification exercise on pre-existing trends is to look at changes in life expectancy during the pre-period, 1900-1940, and see whether they correlate with future (post-1940) changes in predicted mortality. This is done in Figures 6 through 9 and Table 8. This exercise shows no evidence of significant pre-war declines in mortality in countries that would later experience big declines in predicted mortality because of the international epidemiological transition.

Figure 6 shows the change in log life expectancy 1900-1940 against the change in predicted mortality 1940-1980. There is no evidence of a negative relationship similar to those in Figures 4 and 5. In fact, there is a slight positive slope (though column 1 of Table 8 shows that this relationship is not significant). Figure 7 shows the same relationship without the richest countries, and there is now a somewhat stronger positive relationship (though column 2 of Table 8 again shows that this is not significant). Both figures give no indication that there was a pre-existing trend that can explain our first-stage results.

Figures 8 and 9 look at changes in log life expectancy between 1930 and 1940, just before the global interventions. These figures also show no evidence of a negative relationship either for the whole sample (Figure 8) or for subsample excluding the richest countries (Figure 9). Our measure of predicted mortality explains changes in life expectancy after 1940 but not before 1940.

Table 8 extends our examination of potential pre-existing trends to the outcome measures. Specifically, we have enough data to look for a potential relationship between our measure of predicted mortality and changes in log population, log GDP, and log GDP per capita between 1900 and 1940.[29] Columns 1 and 2 show a positive but insignificant relationship between change in predicted mortality 1940-1980 and the change in life expectancy 1900-1940, which was already seen in Figures 6 and 7. Columns 3 and 4 show that there is no pre-existing trend in log population between 1900-1940 either for the entire sample or for the sample excluding the richest countries. Columns 5-8 show similar results for log GDP and log GDP per capita.

These results therefore indicate that there were no pre-existing trends in life expectancy or in our key outcome variables prior to the international epidemiological transition. This makes it possible for us to use predicted mortality as an instrument to investigate the effect of life expectancy on a range of economic outcomes.

Finally, in Table 8 we investigate the reduced-form relationships between predicted mortality and some of our outcome variables. Recall that for us life expectancy at birth is a measure or proxy for overall health of the population. Hence, the reduced-form relationships between

[29] We do not have enough data to do this for total births. Also data limitations mean our sample size is smaller for this exercise than for our main regressions.

predicted mortality and the outcome variables are, in some sense, as informative as the 2SLS estimates of the effect of life expectancy on these variables in Tables 9-16. Panel B of Table 8 shows these reduced-form relationships. As already documented, there is a significant negative relationship between life expectancy and predicted mortality in the period 1940-80. In addition, there is a significant negative relationship between predicted mortality and population during the same period, which indicates the increase in population in previously high mortality areas resulting from the international epidemiological transition. The other rows show that there is no relationship between total GDP and the decline in predicted mortality, and a positive relationship between mortality in GDP per capita. This suggests that declines in mortality have been associated with lower GDP per capita (since total GDP did not increase much and population grows substantially). The 2SLS estimates presented in the next section confirm these reduced-form relationships.

6 Main Results

We now present our main results, which are the 2SLS (two-stage least square) estimates of the effect of log life expectancy on six outcome variables. These outcome variables are: log population, log total births, log GDP, log GDP per capita, log GDP per working age population, and years of schooling.

For each outcome we use two estimation strategies. The first is a full panel with decadal observations between 1940 and 1980, while the second looks only at the long difference using data from 1940 and 1980. The tables have a parallel structure (except when data availability makes this impossible). In addition, in each case, we look both for contemporaneous effects and for effects after 10, 20, 30 and 40 years. This is useful to distinguish the medium-run effects from long-run effects.

6.1 Population

Figure 10 shows the reduced form relationship between change in log population, 1940-80, and the change in predicted mortality over the same period. This is useful both to have visual representation of the relationship and also because the slope of this relationship divided by the slope of the first-state relationship in Figure 4 gives the 2SLS estimate. The figure shows a strong negative relationship, which was already seen in Panel B of Table 8—countries with a larger decline in predicted mortality experienced a larger increase in log population, i.e., more population growth. Given the negative relationship between predicted mortality and life expectancy in Figure 4, this translates into a positive effect of life expectancy on population.

This is confirmed in Table 9, which reports 2SLS results from regressing log population on log life expectancy in either a panel specification (Panel A) or in long differences (Panel B). The first stages underlying these regressions are reported in Table 5 and are not repeated here to save space.

In column 1 we look at contemporaneous effects during 1940-80 and find a coefficient on log life expectancy of 1.30, with a standard error of 0.38 (which is thus statistically significant at less than 1%). This estimate is comparable to the OLS estimates in Table 2.

The coefficient increases to 1.47 when we look at 1930-80 (column 2) and is even larger when we include Eastern Europe (column 3). When we exclude the richest countries in column 4, the coefficient estimate is again similar, 1.42 (standard error = 0.71). This suggests that the results are not driven by a comparison of the initially poor to the initially rich countries.[30]

Column 5 uses the alternative instrument constructed only from the three main killers, and shows a very similar result to that of our baseline specification.

Columns 6-9 investigate the longer-term effects of life expectancy in population growth by looking at the lead specifications. The coefficients are on the whole very similar to the baseline estimate (slightly higher for 10 and 20 year leads and slightly smaller for the 40 year lead). This suggests that changes in life expectancy led to relatively permanent increases in population growth.

Panel B shows the same results with the long difference specifications. The estimates are somewhat larger, between 1.6 and 1.9 instead of 1.3 and 1.5 as in the panel specification.

Overall, there is a large, relatively precise, and robust effect of life expectancy on population. The elasticity is estimated consistently to lie between 1 and 2, which is similar to the OLS estimates.

6.2 Births

Figure 11 shows the reduced form relationship between the change in log births and the change in predicted mortality over 1940-1980. There is a strong negative correlation—countries that experienced a bigger fall in predicted mortality had a larger increase in log births.

Table 10 presents 2SLS estimates of log life expectancy on log total births. Consistent with the magnitude of the response of population to life expectancy, Table 10 indicates that the increase in life expectancy has been associated with an increase in the total number of births. In column 1 Panel A, the estimate is 2.33 (standard error = 0.77). The estimates are similar

[30] If instead of excluding the richest countries, we exclude all of the OECD (and do not include Eastern Europe), the effect of predicted mortality on life expectancy remains significant, but the instrumented effect of life expectancy on population becomes insignificant.

in the long difference specifications, when we include Eastern Europe, when we exclude the richest countries or when we use the alternative instrument constructed from only death rates due to tuberculosis, pneumonia and malaria.

Looking at the leads shows an interesting pattern whereby the effects become smaller at future dates. This confirms that there was a delayed decline in birth rates in response to the increase in life expectancy, and it fits with the evidence reviewed, for example, in Kelley (1988).

6.3 GDP

The main focus of this paper is the effect of life expectancy on GDP and GDP per capita. Figure 12 shows the reduced form relationship between change in log GDP and change in predicted mortality during 1940-1980. Consistent with the pattern in Panel B of Table 8, there is a slight (but not statistically significant) downward slope, which indicates that countries with larger declines in predicted mortality experienced higher GDP growth, though this effect is not very large.

Table 11 presents the related 2SLS regression evidence. In column 1, the estimate of the key parameter is 0.17, which is very small and insignificant (standard error = 0.57). The pattern is similar when we look at different sample periods, when we include Eastern Europe, exclude the richest countries, or use the alternative instrument. In none of the cases is there a significant effect of the increase in life expectancy on total GDP.

The pattern in columns 6-9 is interesting, however. These estimates show that at the longer horizons there is a somewhat more positive effect of life expectancy on GDP (though still not significant). For example, with the 10-year lead the coefficient is now 0.72 (standard error = 0.45) and with the 20-year lead it is 0.84 (standard error = 0.52). The effect starts declining after the 30-year lead. The estimates using the long differences are similar and somewhat larger (though considerably less precise).

We interpret these estimates as suggesting that the increase in life expectancy and the associated increase in population had a relatively small effect on total GDP at first, with a somewhat larger effect over time. The over-time increase in the impact of life expectancy on GDP is likely a result of a combination of the larger population reaching working age and capital inputs and other factors of production adjusting to the increase in population. Nevertheless, it has to be emphasized that even after 40 years, the increase in GDP is small relative to the increase in population.

6.4 GDP Per Capita and Per Working Age Population

The discussion of the response of total GDP already reveals that the effect of the increase in life expectancy on GDP per capita (or GDP per working age population) is going to be negative. This is shown in Figure 13, which depicts a strong positive reduced-form relationship between the change in log GDP per capita and the change in predicted mortality during 1940-1980. Evidently, countries with larger declines in predicted mortality also experienced smaller growth in GDP per capita. Clearly, this is the result of the larger increase in population than in GDP in these countries, which was already shown in Figures 11 and 12 and in Panel B of Table 8.

Table 12 confirms this pattern by presenting the 2SLS estimates of the effect of log life expectancy on GDP per capita. There is a significant negative effect of life expectancy on GDP per capita in columns 1 and 2 of Panel A. The effect is weaker and not as significant in Panel B. In either case, the coefficient estimate, corresponding to π in equation (7), is around -1. The results in columns 3 through 5 hover around statistical significance.

Similar to the results for total GDP in Table 11, the lead results indicate that GDP and GDP per capita are increasing in countries experiencing increases in life expectancy over the following 40 years. Nevertheless, even after 40 years, the effect of life expectancy on GDP per capita is negative (though far from significant).

One concern with these results is that the increase in population is largely at young ages, so GDP per capita may be low precisely because the denominator has increased, while the working age population has not. To investigate the importance of this issue, Table 13 looks at GDP per working age population.[31] The results in Table 13 show that the effect of life expectancy on GDP per working age population are also slightly negative.

Overall, our 2SLS estimates show no evidence that the large increase in life expectancy in parts of the globe starting in the 1940s led to a significant increase in GDP per capita. Instead, the increase in life expectancy was associated with a significant increase in population and a somewhat smaller increase in total GDP. As noted in the Introduction, this is not bad news except for those who believe that improvements in life expectancy can kickstart rapid economic growth. Instead, these results suggest that with international health interventions it is possible to improve life expectancy and health conditions dramatically without significant costs in terms of income per capita.

These results are also consistent with the neoclassical growth model. To elaborate on this point, let us return to the model discussed in Section 2. Let us think of the contemporaneous effects as the "medium run" impact with the capital stock held constant. These

[31] Note, however, that estimates of the age distribution of the population and hence of the working age population for this time period are often rough.

correspond to a coefficient of $\pi = (\alpha(\gamma+\eta) - (1-\alpha)\lambda)$ in terms of equation (6). Recall that λ here is the response of population to changes in life expectancy, so according to the estimates in Table 9, we can think of $\lambda \approx 1.5$. The coefficient α corresponds to the share of labor. Since the countries in question here include many low income countries where land is an important factor of production, we may think of $\alpha \approx 1/3$ (with capital and land as having also one third shares, i.e., $\beta \approx 1/3$ also). This would imply that our estimate of $\pi = (\alpha(\gamma+\eta) - (1-\alpha)\lambda) \approx -1$ is consistent with $\gamma + \eta \approx 0$.[32] Therefore, these results suggest that the benefits of higher life expectancy in terms of direct productivity gains and human capital gains are relatively small. This is also confirmed when we look at the longer-run effects. In this case the estimate of π is approximately -0.60. In terms of the model, the long run response is $\pi = (\alpha(\gamma+\eta) - (1-\alpha-\beta)\lambda)/(1-\beta)$, and again taking $\beta \approx 1/3$ suggests that $\gamma + \eta$ are either zero, or positive but small. Consequently, our estimates suggest that increasing life expectancy, which is an excellent goal of social policy in and of itself, is unlikely to be a magic bullet for stimulating economic growth.

6.5 Years of Schooling

We next look directly at whether increasing life expectancy led to higher human capital. Figure 14 shows the change in average years of schooling, 1960-1990, plotted against the change in predicted mortality, 1940-1980. Lack of data prevents us from using earlier data on schooling, but this may not be a severe limitation as there is presumably an important lagged effect here, i.e., a child who survived due to a decline in mortality rates might not complete schooling for 10 or so years after the mortality decline. In any case, Figure 14 shows that years of schooling is positively related to predicted mortality—countries with larger declines in predicted mortality later experienced less (rather than more) increase in schooling.

Table 14 estimates the corresponding 2SLS regressions. With 10-year or 20-year leads, there is no effect of life expectancy on schooling either in the OLS or in the IV (columns 1-6). With 30-year leads, there is a positive OLS estimate, but the IV estimates are again insignificant (either positive or negative depending on the sample as shown in columns 8 and 9).

Overall, there seems to be no evidence that the increase in life expectancy has been associated with substantially greater investment in human capital, which is consistent with the finding in the previous subsection. The most likely reason why the increase in life expectancy did not translate into greater education during this episode is that the affected countries likely

[32] If we take λ as 2, $\gamma + \eta$ could be as large as 1. But in turn, if α were higher, the implied values of $\gamma + \eta$ would be correspondingly lower.

faced bottlenecks in their education systems, making it impossible for them to increase the education of the much larger cohorts of children that survived as a result of the international epidemiological transition.

7 Robustness and Further Results

The results in the previous section suggest that the increase in life expectancy led to a substantial increase in population, but not to economic growth. In this section, we investigate the robustness of these results further.

7.1 Mean Reversion in the Second Stage

As noted above, our estimates so far ignore mean reversion in the second stage. This implies that the exclusion restriction for our 2SLS estimation was $\text{Cov}(M^I_{it}, \varepsilon_{it+k}) = 0$. It is possible that this exclusion restriction might hold conditional on past levels of income per capita (or other outcome variables), i.e., $\text{Cov}(M^I_{it}, \varepsilon_{it+k} \mid y_{it-1}) = 0$, while at the same time we may have $\text{Cov}(M^I_{it}, \varepsilon_{it+k}) \neq 0$. Although we have no a priori reason to expect this particular configuration, it is informative to check whether allowing for mean reversion in the second stage affects our results. We investigate this issue for our main variable of interest, income per capita, in Table 15.

Our first strategy is to estimate (9) directly with GMM again using the Arellano-Bond procedure. In this estimation, further lags of GDP per capita are used as instruments for lagged GDP per capita, while predicted mortality, M^I_{it}, is used as the exogenous (external) instrument for life expectancy. Columns 1 and 2 of Panel A in Table 15 report results from this estimation strategy. In column 1, only predicted mortality is used as an instrument for life expectancy, where as in column 2 both predicted mortality and lagged life expectancy are used as instruments for life expectancy (which is in line with model (12)). The estimates of the effect of log life expectancy on income per capita are negative and insignificant, though much smaller than our baseline estimates in Table 12 (e.g., -0.31 versus -1.09). Given the additional exclusion restrictions that are necessary for consistent estimation of equation (9), we have more confidence in the results from estimating equation (8) (i.e., Table 12).

As a second strategy, instead of directly estimating equation (9), we estimated a transformed model that removes the effect of mean reversion in income per capita. This strategy is both more transparent and does not require the full moment conditions of the GMM strategy used in columns 1 and 2, and is thus our preferred approach.

Suppose we know the mean-reversion parameter ρ in equation (9), then we can construct,

$$\tilde{y}_{it}^{\rho} = y_{it} - \rho y_{it-1} \text{ and } \tilde{x}_{it}^{\rho} = x_{it} - \rho x_{it-1},$$

and run an instrumental-variables regression of $\tilde{y}_{it}^{\rho}$ on $\tilde{x}_{it}^{\rho}$ to identify the coefficient of interest π. Although we do not know ρ, we can implement a two-stage version of this procedure by first estimating $\hat{\rho}$.[33] Columns 3-10 of Table 15 apply this procedure using a range of values for ρ that covers (and exceeds) this range. Panel B shows the first stages for these transformed variables. The results hover between -1.38 and -1.58, and show that irrespective of the value of ρ, there is a similar relationship between life expectancy and GDP per capita are similar to but somewhat more negative than the baseline estimates in Table 12.

Overall, there is no evidence that allowing for mean reversion changes the conclusions of the previous section.

7.2 Interaction Results

As discussed in Section 2, we may expect different results of log life expectancy on GDP per capita depending on the savings and investment rate. We investigate this issue in Table 16. Since reliable cross-country data on the savings rate are difficult to obtain and generally not reliable, we start by looking at initial (1940) GDP per capita and investment rates from the 1940s (or immediately after).

Although income differences in 1940 likely had various causes, we expect them to be correlated with savings and investment rates. Our empirical strategy is therefore to include an interaction between log life expectancy and initial GDP per capita or investment as a percent of GDP. This interaction term is instrumented by the interaction between predicted mortality and initial GDP per capita (or investment). All variables are demeaned, so that the main effects are evaluated at the sample mean. First-stages are not shown to save space.

The results of this exercise are presented in Table 16. Panel A reports results for log population, Panel B is for total GDP, and Panel C is for GDP per capita.

Panel A shows that the effect of log life expectancy on population is the same irrespective of initial GDP or the investment rate; the interactions between log life expectancy and these baseline characteristics are insignificant both in contemporaneous and lead specifications.

The picture is different in Panel B. The interaction terms are typically positive, large and sometimes significant. For example, the interaction between log life expectancy and initial

[33] More specifically, in regressions of log income per capita on its lag and county and time fixed effects as those in Table 15, the estimates of ρ are between 0.4 and 0.65 depending on estimation strategy (e.g., OLS or Arellano and Bond's (1991) GMM) and on whether or not log life expectancy is included..

GDP per capita in the 10-year lead specification of column 2 is significant at the 5% level and in the 20-year lead specification of column 3, it is significant at the 10% level. With investment, the contemporaneous interaction is significant at 10%. In both cases, the interaction terms decline both in size and insignificance as we look at further leads.

Panel C shows similar results for GDP per capita. The interaction effects are, on the whole, similar to those for total GDP, since there are no interaction results for population.

Overall, consistent with our theoretical expectations, there is some evidence that countries with high investment rates (measured directly or proxied by high initial income per capita) appear to have been able to increase their GDP more rapidly in response to increases in life expectancy and population. Moreover, consistent with equation (5) in Section 2, these investment rate differences appear to have had no effect on the long-run relationship between log GDP (or log GDP per capita) and log life expectancy. Nevertheless, the results of this exercise have to be interpreted with caution, since data quality and sometimes large standard errors limit the extent to which we can pin down the exact timing of changes in GDP.

8 Conclusion

A newly-emerging consensus in academic and policy circles holds that disease environment and health conditions lie at the root of large income differences across countries today, and argues that improving health will not only improve lives but will by itself spur rapid economic growth.

This paper investigated these claims by estimating the effect of life expectancy at birth on economic growth. The innovation in our approach is to exploit the international epidemiological transition, which led to potentially exogenous differential changes in mortality from a number of major diseases across the world. As a result of new chemicals, drugs, and other international interventions, mortality from tuberculosis, pneumonia, malaria, and various other diseases declined sharply in many parts of the world, while other countries that were largely unaffected by these diseases did not experience similar improvements in health. Exploiting these differential changes in predicted mortality as an instrument for life expectancy, we estimate the effect of life expectancy on a range of economic variables, most importantly population and GDP.

Our results indicate that the increase in life expectancy led to a significant increase in population; birth rates did not decline sufficiently to compensate for the increase in life expectancy. We find a small initial positive effect of life expectancy on total GDP, and this effect grows slightly over the next 40 years. Overall, the increases in life expectancy (and the associated increases in population) appear to have reduced income per capita at first, with this negative

effect slowly wearing off over the next 40 years. There is no evidence that the increase in life expectancy led to faster economic growth.

This evidence sheds considerable doubt on the view that health has a first-order impact on economic growth. But it is also good news for the efforts to improve health in the poorer parts of the world. Our results indicate that global health interventions can lead to substantial improvements in life expectancy and health without significant negative effects on income per capita, and suggest that efforts to combat poor health conditions in less developed nations can be very effective. They caution, however, against claims that such efforts are likely to accelerate economic growth or should be supported because of their positive growth implications.

In conclusion, it is also important to emphasize the limitations of our results. The most important limitation is that since our approach exploits the international epidemiological transition in and around the 1940s, the results may not be directly applicable to other episodes. This is for at least two reasons. First, the international epidemiological transition was a unique event and perhaps similar changes in life expectancy today will not lead to an increase in population and the result on GDP may be more positive. In particular, today there could be a greater fertility response to mortality declines—birth rates are already lower almost everywhere than they were in the 1930s. Second, the diseases that take many lives in the poorer parts of the world today are not the same as those 60 years ago; most notably HIV/AIDS is a major killer today but was unknown in 1940. Most of the diseases we focus on had the greatest impact on children (with the exception of tuberculosis), while HIV/AIDS affects individuals at the peak of their labor productivity and could have a larger negative impact on growth. Investigating the specific effects of the HIV/AIDS epidemic on economic outcomes, as in Young (2005), is an important area for future research.

9 Appendix A: Data Sources and Construction

Our main aggregate economic data come from Maddison (2003); specifically, we use data on population and GDP from his CD-Rom. Working age population is from the UN *Demographic Yearbook*, various issues, taking the population between 15 to 60 as working age. Life expectancy in 1940 and earlier are from the UN *Demographic Yearbook* (1947, 1948-49, 1950, and particularly the retrospective section of 1967).[34] Life expectancy from 1950 was downloaded from the UN's on-line demographic database.[35] Details by country are in Appendix C.

For cause of death, we use the Abridged List of the 1938 revision of the International Classification of Disease. This list is comprehensive and has 44 categories. We omit any diseases that are not infectious or could be degenerative, e.g., "diseases of the heart" (Abridged List No. 24). We also omit residual categories, such as "other infectious or parasitic diseases" (Abridged List No. 14), as we do not know which specific global interventions may have been effective against these diseases. We omit syphilis (Abridged List No. 9) because, while penicillin provided an effective cure, there are issues of behavior and risk taking that are likely distinct from our focus here. We also omit diseases that were never major causes of death, even though they may have had serious effects on health (e.g., acute poliomyelitis). There are 15 infectious diseases for which we can obtain comparable cross-country data on deaths per 100,000 in 1940 (or 1939).

The classification of death rates by cause changed in 1948, and some of our data for 1950 and after are available only according to the Abbreviated List, 1948 Revision of the International Classification of Disease.[36] Most of our 15 diseases can be tracked through this reclassification, but dysentery/diarrhoea-related diseases cannot.

The basis for our cause of death data is the *Summary of International Vital Statistics, 1937-1944,* published by the Federal Security Agency (1947) of the US government immediately after World War II. This source pulled together comparable comprehensive data on cause of death around 1940, as well as longer time series on the more important diseases (i.e., death rates by country), but it did not use all the available data (Federal Security Agency 1947, p.2). We use their original sources, which are national health statistics collected, cleaned and republished between the wars by the League of Nations Health Organization (see Federal Security Agency, 1947, pp.1-3).

[34] Some of these data were previously used by Preston (1975).

[35] These data are in five year intervals, so we use 1950-55 for 1950 and 1960-65 for 1960, etc.

[36] For example, UN *Demographic Yearbook* (1954) reports cause of death in and around 1950 for some countries using the 1938 classification and for others using the 1948 classification.

The League of Nations Health Organization helped establish comparable international health statistics for a large number of countries, but never published a comprehensive retrospective of the data. Their first relevant publication was issue number 7 of the *Annual Epidemiological Report*, published in October 1923, but only from 1929 (covering the year 1927) did this publication include death rates from specified causes (League of Nations Health Organization, 1929).[37] We use the rates for 1930 from League of Nations Health Organization (1933).[38] For 1940 we use World Health Organization (1951), which provided data for 1939-46, based on the League of Nations' work.[39] Data for 1950 and 1960 are from the UN *Demographic Yearbook* for 1954, 1962 and 1966.

By reviewing every available interwar issue of the League of Nations' *Weekly Epidemiological Record*, we confirm that our data do not miss major epidemics.[40] We also confirm that our numbers are consistent with contemporary qualitative assessments, in particular in the League of Nations and WHO's annual reports. For more detail see Appendix C.

Predicted mortality in 1940 is calculated by adding deaths per 100,000 from the 15 component diseases (for ease of exposition, we then convert to per 100 of population).[41] As discussed in the text, we assume this 1940 value for each country-disease pair holds until there is an "intervention."

Years of schooling are from the Barro-Lee dataset, downloadable from the NBER website. Our investment data are based on Maddison (1992), but we fill gaps with data for the early 1950s from Kuznets (1960). We assume the investment rate in China in 1940 was the same as in Korea.

[37] Early issues of this publication are also refered to as *Statistics of Notifiable Diseases*. The first six issues focused on Eastern Europe, particularly typhus and malaria in Russia.

[38] We use the report published in 1933 as it has less missing data for 1930; presumably some data became available with a lag or was corrected. While the League of Nations Health Organization made an explicit effort to expand its country coverage, it never published a comprehensive retrospective assessment, and reports for one year would only include all the available data for the previous four years.

[39] In addition, for malaria, we use data from the Leauge of Nations' Malaria Commission (League of Nations Health Organization, 1932), as well as information on location of malaria in the 1940s from American Geographical Society (1951a).

[40] For example, for the distribution of cholera in 1938, see *Weekly Epidemiological Record*, March 3rd, 1938. For the distribution of small pox in 1930, see *Weekly Epidemiological Record*, August 21st, 1930; for 1938, see *Weekly Epidemiological Record*, March 3rd, 1938; for the early 1940s see *Weekly Epidemiological Record*, January 3rd, 1946. For the pre-war distribution of diphtheria, with a focus on Europe, see *Weekly Epidemiological Record*, December 21st, 1939. For the distribution of plague in 1938, see *Weekly Epidemiological Record*, March 3rd, 1938. For more detail on the pre-1940 distribution of typhus, see *Weekly Epidemiological Record*, September 14th, 1939. For the endemic yellow fever zone in 1951, see the *Supplement to the Weekly Epidemiological Record*, 25 September 1952.

[41] Preston (1980) points out that data on precise cause of death should be handled with care; for example, it is notoriously difficult to determine how many deaths are due directly and indirectly to malaria. However, as our analysis is about changes in total predicted mortality from infectious disease and because most of the global interventions were clustered in the late 1940s and early 1950s, this issue is less of a concern here.

10 Appendix B: List and Details On Diseases

The main text reviewed the etiology of and global "interventions" against the three diseases in our data responsible for the most deaths: malaria, pneumonia, and tuberculosis. Here we provide detail on the remaining 12 infectious diseases, in rough descending order of their contribution to global deaths around 1940.[42] The relevant global interventions are (a) new drugs for treatment that became available globally (particularly antibiotics where relevant), (b) new preventive measures that became available globally (particularly vaccines and chemicals that were effective against insects) and, (c) specific WHO campaigns against diseases.

Influenza is caused by various strains of the influenza virus, including type A (the most dangerous), type B, and type C. Transmission is through coughing, sneezing, or directly through mucous membranes. Associated deaths are often due to various secondary bacterial infections. The primary control mechanism is vaccination, but the introduction of antibiotics from the 1940s reduced deaths from secondary bacterial infections. There has been no global campaign to eradicate influenza, but WHO efforts to control and track the disease started in the 1950s. For an assessment of measures taken against influenza during 1921-50, see Deutschman (1953). In our baseline instrument we take the intervention date as the 1940s (antibiotics) and in our alternative instrument we take the 1950s (WHO action).

Cholera is caused by the bacterium *Vibrio cholerae*, and is transmitted by drinking contaminated water or eating contaminated food. Public works to properly treat or dispose of sewage have been effective against the disease since the mid-nineteenth century. Some antibiotics reduce the symptoms, but oral rehydration or intraveneous fluids are needed to replace minerals and fluids lost due to diarrhoea. Major steps to improve the effectiveness of oral rehydration were taken during the 1950s; in part these innovations were supported by the US military. For our baseline instrument we take the intervention date as the 1950s (rehydration therapy) and in our alternative instrument we take the 1940s (antibiotics).

Typhoid is caused by the bacterium *Salmonella typhi* and is transmitted through feces, either directly or by flies. It can be treated effectively with antibiotics (available since the 1940s). We take the 1940s as the intervention date for both our baseline and alternative instruments (based on antibiotics).

Smallpox was caused by the viruses *Variola major* (the more deadly) and *Variola minor*. The disease was highly contagious, with the virus spreading through contact or through the air. Since 1798 the primary treatment has been vaccination. The WHO passed a resolution declaring the need to eradicate the disease in 1958 and the invention of the jet injector with

[42] The main sources for this section are Kiple (1993), Hoff and Smith (2000), and the Centers for Disease Control and Prevention website.

foot pedal in 1962 made it possible to easily vaccinate people in places without electricity. In 1979, smallpox was declared entirely eradicated. In our baseline instrument we take the 1950s as our intervention date and in our alternative instrument we take the 1960s.

Shigella dysentery is caused by the bacterium *Shigella dysenteriae* type 1 or by the protozoan *Endamoeba histolytica* and is transmitted in the same fashion as typhoid. While we do not have fully comparable international data on dysentery, there are data on deaths from diarrhea among infants under the age of 2; we convert these into per 100 population equivalent and add to our predicted mortality estimates. The disease is controlled with public health measures, antibiotics, and rehydration therapy. We take the 1940s as our intervention date for both our baseline instrument (based on antibiotics) and the 1950s for our alternative instrument (based on rehydration therapy).

Whooping cough is caused by the bacteria *Bordetella pertussis*. It can be treated with antibiotics and prevented by vaccination (one component of the DTP vaccine.) The vaccine became available in the 1920s. We take the 1940s as our intervention date for both our baseline and alternative instruments (based on the effectiveness of antibiotics).

Measles (rubeola) is caused by a virus of the *Rubivirus* genus; it spreads through airborne droplets from an infected person.[43] Prevention is through vaccination, which became available in 1963; this is also effective if administered within three days of exposure to the disease. Currently the largest vaccine-preventable killer of children, it may be targeted for global eradication. We take the 1960s as our intervention date for both our baseline and alternative instruments.

Diphtheria is caused by the bacterium *Corynebacterium diphtheriae* when it has been infected by certain bacteriophages (parasites that only infect bacteria). Transmission is through the air or by touch. It can be treated with antitoxins and antibiotics. An antitoxin has been available since the 1890s and immunization spread after its introduction in the early 1920s (usually provided today in the DTP, diphtheria-tetanus-pertussis, vaccine for infants). Treatment became more effective with the introduction of antibiotics in the 1940s. We take the 1940s as our intervention date for our baseline and alternative instruments (based on antibiotics).

Scarlet fever is caused by the *Streptococcus* bacteria; it often develops in strep throat patients and is similarly spread by droplets from an infected person (e.g., coughing or sneezing). It generally can be treated with antibiotics, including penicillin. We take the 1940s as our intervention date for our baseline and alternative instruments (based on antibiotics).

Yellow fever is caused by the yellow fever virus, and transmitted by the bite of an infected

[43] This is a different disease, caused by a different virus, than German measles (rubella). Vaccines for both are included in the MMR vaccine (measles-mumps-rubella).

Aedes aegepti mosquito. It is controlled by vaccination and public health measures against the mosquito vector. The vector was definitively identified by Walter Reed, head of the U.S. Army Yellow Fever Commission, in 1900-1901. The first vaccine was developed by Max Theiler in the 1937; he was awarded a Nobel Prize in 1951. We take the 1940s as the intervention date for our baseline instrument and the 1930s for our alternative instrument.

Plague is caused by the bacterium *Yersinia pestis* and is transmitted from infected animals to humans through the bite of an infected flea. The disease is controlled through antibiotics, especially streptomycin, and the elimination of rodent population near human habitations. Some protection from vaccination has been available since the end of the nineteenth century. The WHO attempts to help deal with outbreaks. We take the 1940s as the intervention date for both our baseline and alternative instruments (based on antibiotics).

Typhus is caused by any microbe of the genus *Rickettsia*, and is transmitted by insects (lice, fleas, mites, and ticks). Antibiotics are usually an effective treatment. Public health measures include good hygiene and sanitation. We take the 1940s as the intervention date for both our baseline and alternative instruments (based on antibiotics).

11 References

Acemoglu, Daron, Simon Johnson and James A. Robinson (2001) "The Colonial Origins of Comparative Development: An Empirical Investigation," *The American Economic Review*, December, Volume 91, Number 5, 1369-1401.

Acemoglu, Daron, Simon Johnson and James A. Robinson (2002) "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics* vol 117, November, 1231-1294.

Acemoglu, Daron, Simon Johnson and James A. Robinson (2003) "Disease and Development in Historical Perspective," *Journal of the European Economic Association,* April, vol. 1, no.2-3, 397-405.

Alleyne, George A. O. and Daniel Cohen (2002) "The Report of Working Group I of the Commission on Macroeconomics and Health," WHO Commission on Macroeconomics and Health, April.

American Geographical Society (1951a) *World Distribution of Malaria Vectors*, New York.

American Geographical Society (1951b) *Distribution of Cholera, 1816-1950*, New York.

American Geographical Society (1951c) *World Distribution of Plague*, New York.

American Geographical Society (1951d) *World Distribution of Rickettial Diseases: Louse-Borne and Flea-Borne Typhus*, New York.

Anderson, Theodore W. and Cheng Hsiao (1982) "Formulation and Estimation of Dynamic Models using Panel Data," *Journal of Econometrics*, 18, 67-82.

Arellano, Manuel and Stephen R. Bond (1991) "Some Specification Tests for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 58, 277-298.

Arndt, Channing and Jeffrey D. Lewis (2000) "The Macro Implications of HIV/AIDS in South Africa: A Preliminary Assessment," *South African Journal of Economics*, LXVIII, 856-887.

Becker, Gary S., Tomas J. Philipson, and Rodrigo R. Soares (2005) "The Quantity and Quality of Life and the Evolution of World Inequality", *The American Economic Review*, March, 277-291.

Behrman, Jere R. and Mark R. Rosenzweig (2004) "The Returns to Birthweight," *Review of Economics and Statistics*, 86:2, May, 586-601.

Bell, Clive, Shantanyanan Devarajan, and Hans Gersbach (2003) "The Long-Run Economic Costs of AIDS: Theory and an Application to South Africa," working paper, June.

Bleakley, Hoyt (2004) "Disease and Development: Evidence from Hookworm Eradication in the American South," working paper, August.

Bloom, David E. and David Canning (2005) "Health and Economic Growth: Reconciling the Micro and Macro Evidence," working paper, February

Bloom, David E. and Jeffrey D. Sachs (1998) "Geography, Demography, and Economic Growth in Africa," *Brookings Papers on Economic Activity*, 1998:2, 207-295.

Bradley, D.J. (1992) "Malaria: Old Infections, Changing Epidemiology," *Health Transition Review*, vol. 2, supplementary issue, 137-152.

Caldwell, J. C. (1986) "Routes to Low Mortality in Poor Countries," *Population and Development Review*, 12, 2:171–220.

Cliff, Andrew, Peter Haggett, and Matthew Smallman-Raynor (2004) *World Atlas of Epidemic Diseases*, Arnold Publishers/Oxford University Press.

Davis, Kingsley (1956) "The Amazing Decline of Mortality in Underdeveloped Areas," *The American Economic Review*, 46, 2, May, 305-318.

Deaton, Angus (2003) "Health, Inequality, and Economic Development," *Journal of Economic Literature*, March, v.41, issue 1, 113-158.

Deaton, Angus (2004) "Health in an Age of Globalization," *Brookings Trade Forum*, edited by Susan Collins and Carol Graham, The Brookings Institution, Washington DC.

Deutschman Z. (1953) "Trend of Influenza Mortality During the Period 1920-51," *Bulletin of World Health Organization*, 8, 633-645.

Desowitz, Robert S. (1991) *The Malaria Capers : More Tales of Parasites and People, Research and Reality*, W.W. Norton & Company.

Easterlin, Richard (1999) "How Beneficent is the Market? A Look at the Modern History of Mortality," *European Review of Economic History*, 3 (3), 257-294.

Easterly, William and Ross Levine (2003) "Tropics, germs, and crops: how endowments influence economic development," *Journal of Monetary Economics*, 50, 3-39.

Federal Security Agency (1947) *Summary of International Vital Statistics, 1937-1944*, U.S. Public Health Service, National Office of Vital Statistics.

Fogel, Robert William (2004) *The Escape from Hunger and Premature Death, 1700-2100*, Cambridge University Press.

Gallup, John Luke and Jeffrey D. Sachs (2000) "The Economic Burden of Malaria," CID Working Paper No.52, Center for International Development, Harvard University, July.

Gallup, John Luke and Jeffrey D. Sachs with Andrew D. Mellinger (1999) "Geography and Economic Development," CID Working Paper No. 1, Center for International Development, Harvard University, March.

Galor, Oded (2005) "From Stagnation to Growth: Unified Growth Theory," *Handbook of Economic Growth* (forthcoming).

Galor, Oded and Weil, David (2000) "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *The American Economic Review*, 90(4), 806-828.

Hall, Robert E. and Charles I. Jones (1999) "Why do Some Countries Produce so Much More Output per Worker than Others?" *Quarterly Journal of Economics*, Vol. 114, February, 83-116.

Hansen, Gary D. and Edward C. Prescott (2002) "Malthus to Solow," *The American Economic Review*, 92(4), 1205-1217.

Hoff, Brent and Carter Smith III (2000) *Mapping Epidemics: A Historical Atlas of Disease*, Franklin Watts, New York.

Kalemli-Ozcan, Sebnem (2002) "Does the Mortality Decline Promote Economic Growth?," *Journal of Economic Growth*, volume 7(4), 311-349.

Kalemli-Ozcan, Sebnem, Harl E. Ryder, and David Weil (2000) "Mortality Decline, Human Capital Investment, and Economic Growth," *Journal of Development Economics*, Vol. 62, 1-23.

Kelley, Allen C. (1988) "Economic Consequences of Population Change in the Third

World," *Journal of Economic Literature*, Vol. 26, No. 4, December, 1685-1728.

Kiple, Kenneth F. (1993) *The Cambridge World History of Human Disease,* Cambridge University Press.

Knaul, Felicia Marie (2000) "Health, Nutrition, and Wages: Age at Menarche and Earnings in Mexico," in William D. Savedoff and T. Paul Schultz, eds., *Wealth from Health: Linking Social Investments to Earnings in Latin America*, Washington: Inter-American Development Bank.

Kuznets, Simon (1960) "Quantitative Aspects of the Economic Growth of Nations, V. Capital Formation Proportions: International Comparisons for Recent Years," *Economic Development and Cultural Change,* Vol. VIII, No. 4, Part II, July.

League of Nations Health Organization (1929), *Statistics of Notifiable Diseases for the Year 1927*, Information Section, April, Geneva.

League of Nations Health Organization (1932) *Enquiry into the Quinine Requirements of Malarial Countries and the World Prevalence of Malaria*, C.H./Malaria/185, Series of League of Nations Publications, III, Health, December.

League of Nations Health Organization (1933) *Statistics for Notifiable Diseases for the Year 1931,* Annual Report No.16, Geneva.

Maddison, Angus (1992) "A Long-Run Perspective on Saving," *Scandinavian Journal of Economics*, 94(2), 181-196.

Maddison, Angus (2001) *The World Economy: A Millennial Perspective*, OECD, Development Centre Studies, 2001.

Maddison, Angus (2003) *The World Economy: Historical Statistics*, OECD, Development Centre Studies, 2003.

McArthur, John W. and Jeffrey D. Sachs (2001) "Institutions and Geography: Comment on Acemoglu, Johnson, and Robinson (2000)," NBER Working Paper No. 8114, February.

McEvedy, Colin and Richard Jones (1978) *Atlas of World Population History*, Facts on File, New York.

McKeown, Thomas (1976) *The Modern Rise of Population*, Academic Press, New York.

Masters, William A. and Margaret S. McMillan (2001) "Climate and Scale in Economic Growth," *Journal of Economic Growth*, Volume 6, Number 3, September, 167-186.

Miguel, Edward and Michael Kremer (2004) "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities," *Econometrica*, Volume 72, Issue 1, 159-217.

National Academy of Sciences (1970) *The Life Sciences*, Committee on Research in

Life Sciences of the Committee on Science and Public Policy, Washington DC.

Office International d'Hygiene Publique (1933) *Vingt-Cinq Ans d'Activite de L'Office International d'Hygiene Publique*, Paris.

Omran, Abdel R. (1971) "The Epidemiologic Transition," *The Milbank Memorial Fund Quarterly*, 1971, 49(4), 509-538.

Persico, Nicola, Andrew Postlewaite, and Dan Silverman (2004) "The Effect of Adolescent Experience on Labor Market Outcomes: The Case of Height," *Journal of Political Economy*, 112, 1019–1053

Preston, Samuel H. (1975) "The Changing Relation Between Mortality and Level of Economic Development," *Population Studies*, 29(2), 231-248.

Preston, Samuel H. (1980) "Causes and Consequences of Mortality Declines in Less Developed Countries during the Twentieth Century," in *Population and Economic Change in Developing Countries*, edited by Richard A. Easterlin, The University of Chicago Press, Chicago and London.

Preston, Samuel H. (1996) "American Longevity: Past, Present, and Future," Syracuse University, Center for Policy Research, Policy Brief, No. 7/1996.

Preston, Samuel H. and Verne E. Nelson (1974) "Structure and Change in Causes of Death: An International Summary," *Population Studies*, 28, 1, March, 19-51.

Pritchett, Lant and Lawrence H. Summers (1998) "Wealthier Is Healthier" *Journal of Human Resources*, Vol. 31, pp. 841-868.

Sachs, Jeffrey D. (2003) "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income," NBER Working Paper 9490, February.

Schultz, T. Paul (1993) "Mortality Decline in the Low-Income World: Causes and Consequences," *The American Economic Review*, 83, 2, May, 337-342.

Schultz, T. Paul (2002) "Wage Gains Associated with Height as a Form of Health Human Capital," *The American Economic Review*, 92(2), May, 349-353.

Schultz, T. Paul and Aysit Tansel (1997) "Wage and Labor Supply Effects of Illness in Cote d'Ivoire and Ghana: Instrumental Variable Estimates for Days Disabled" *Journal of Development Economics*, volume 53, pp. 251-286.

Soares, Rodrigo R. (2005) "Mortality Reductions, Educational Attainment, and Fertility Choice," *The American Economic Review*, July, forthcoming.

Stolnitz, George J. (1955) "A Century of International Mortality Trends: I," *Population Studies*, July, 24-55.

Stolnitz, George J. (1965) "Recent Mortality Trends in Latin America, Asia and Africa: Review and Re-interpretation," *Population Studies*, 19, 2, November, 117-138.

Strauss, John and Duncan Thomas (1998) "Health, Nutrition, and Economic Development," *Journal of Economic Literature*, June, Vol. XXXVI, pp.766-817.

Swaroop S. and R. Pollitzer (1955) "Cholera Studies," *Bulletin of the World Health Organization*, 12, 311-358.

Thomas, Duncan and John Strauss (1997) "Health and Wages: Evidence on Men and Women in Urban Brazil" *Journal of Econometrics*, Vol. 77, pp. 159-185.

Weil, David N. (2005) "Accounting for the Effect of Health on Growth," working paper, May.

World Health Organization (1951) *Annual Epidemiological and Vital Statistics, 1939-46*, Geneva.

World Health Organization (1998) "Milestones on the Way Towards the World Health Organization," at www.who.int/archives/who50/en/milestones.htm.

World Health Organization (2001) *Macroeconomics and Health: Investing in Health for Economic Development*, on the web at http://www3.who.int/whosis/cmh.

World Health Organization (2004) *Fifty Years of the World Health Organization in the Western Pacific Region*, Report of the Regional Director to the Regional Committee for the Western Pacific.

Young, Alwyn (2005) "The Gift of the Dying: The Tragedy of AIDS and the Welfare of Future African Generations," *Quarterly Journal of Economics*, Volume CXX, Issue 2, May, 423-466.

Table 1
Descriptive Statistics

	Whole World	Base Sample	Initially Rich Countries	Initially Middle Income Countries	Initially Poor Countries
Life expectancy in 1900	30.90	37.04	49.36	36.92	28.77
	(8.83)	(10.45)	(3.67)	(8.13)	(5.42)
Life expectancy in 1940	47.50	49.29	65.34	49.82	40.67
	(11.29)	(12.45)	(2.30)	(9.08)	(7.86)
Life expectancy in 1980	61.11	66.20	74.31	69.66	61.77
	(11.05)	(7.56)	(1.13)	(4.58)	(7.32)
Predicted mortality in 1940	n.a.	0.48	0.17	0.48	0.53
		(0.28)	(0.05)	(0.22)	(0.32)
Log population in 1940	8.94	9.07	9.35	8.82	9.15
	(1.55)	(1.55)	(1.34)	(1.41)	(1.79)
Log population in 1980	8.89	9.71	9.76	9.44	10.00
	(1.62)	(1.31)	(1.29)	(1.26)	(1.75)
Log GDP in 1940	9.78	9.89	11.08	9.75	9.19
	(1.68)	(1.61)	(1.40)	(1.49)	(1.71)
Log GDP in 1980	10.00	11.34	12.47	11.42	10.89
	(1.98)	(1.40)	(1.33)	(1.36)	(1.52)
Log GDP per capita in 1940	7.65	7.73	8.64	7.84	6.95
	(0.69)	(0.71)	(0.15)	(0.34)	(0.33)
Log GDP per capita in 1980	7.99	8.54	9.62	8.89	7.79
	(1.08)	(0.90)	(0.13)	(0.45)	(0.74)

Mean values of variables; standard deviation in parentheses. Base sample is 59 countries. Initially rich countries had log GDP per capita over 8.4 in 1940; middle income had log GDP per capita between 7.37 and 8.4 in 1940; and low income countries had log GDP per capita below 7.37 in 1940. Predicted mortality is per 100 per annum. "n.a." denotes not available.

Table 2
Life Expectancy, Population, and Births: OLS Estimates

Dependent variable indicated for each panel separately

	All Countries	*Base Sample*		*Low & Middle Income Countries Only*	*All Countries*		*Base Sample*			
	No leads	No leads	No leads	No leads	10 year lead	20 year lead	10 year lead	20 year lead	10 year lead	20 year lead
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
				Panel A: Dependent variable is log population						
	Panel, 1960-2000	Panel, 1960-2000	Panel, 1940-1980	Panel, 1940-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1940-1980	Panel, 1940-1980
Log Life Expectancy	1.46	1.69	1.11	1.03	1.72	1.61	1.34	0.97	1.21	1.11
	(0.29)	(0.43)	(0.21)	(0.28)	(0.26)	(0.34)	(0.46)	(0.46)	(0.22)	(0.21)
Number of observations	600	294	284	267	480	360	235	176	284	284
Number of countries	120	59	59	48	120	120	59	59	59	59
				Panel B: Dependent variable is log population						
	Just 1960 and 2000	Just 1960 and 2000	Just 1940 and 1980	Just 1940 and 1980	Just 1960 and 1990	Just 1960 and 1980	Just 1960 and 1990	Just 1960 and 1980	Just 1940 and 1980	Just 1940 and 1980
Log Life Expectancy	1.60	1.74	1.41	1.40	1.92	1.70	1.42	1.58	1.48	1.37
	(0.42)	(0.57)	(0.32)	(0.52)	(0.35)	(0.41)	(0.57)	(0.20)	(0.29)	(0.27)
Number of observations	240	118	96	74	240	240	118	116	96	96
Number of countries	120	59	48	37	120	120	59	58	48	48
				Panel C: Dependent variable is log number of births						
	Panel, 1960-1990	Panel, 1960-1990	Panel, 1940-1980	Panel, 1940-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1930-1970	Panel, 1930-1970
Log Life Expectancy		2.02	1.54	1.54			1.39	0.30	1.24	0.85
		(0.46)	(0.32)	(0.20)			(0.49)	(0.57)	(0.24)	(0.24)
Number of observations		188	233	233			141	94	234	187
Number of countries		47	47	47			47	47	47	47
				Panel D: Dependent variable is log number of births						
	Just 1960 and 1990	Just 1960 and 1990	Just 1940 and 1980	Just 1940 and 1980	Just 1960 and 1980	Just 1960 and 1970	Just 1960 and 1980	Just 1960 and 1970	Just 1940 and 1980	Just 1940 and 1970
Life Expectancy		2.00	1.47	1.25	1.42	0.38	1.37	0.30	1.22	0.96
		(0.42)	(0.44)	(0.62)	(0.57)	(0.60)	(0.59)	(0.57)	(0.34)	(0.33)
Number of observations		94	92	70	100	100	94	94	92	92
Number of countries		47	46	35	50	50	47	47	46	48

OLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panels A and C are unbalanced panels with one observation per decade. Panels B and D are long-difference specifications with observations for only the beginning and end dates. Dependent variable is log population in Panels A and B and log total births in Panels C and D. Independent variable in all regressions is log life expectancy at birth. In columns 1-4, the dependent variable and independent variable are for the same time period; in columns 5-10, the dependent variable is for t+10 or t+20 as indicated, while the independent variable is for time t. "All countries" are those for which we have data on the dependent and independent variables. Base sample is countries for which we have disease data. Assignment of countries to low and middle income categories is based on 1940 income per capita; see text and Appendix A for details and definitions.

Table 3
Life Expectancy, GDP and GDP per capita: OLS Estimates

	Dependent variable indicated for each panel separately									
	All Countries	*Base Sample*		*Low & Middle Income Countries Only*	*All Countries*		*Base Sample*			
	No leads	No leads	No leads	No leads	10 year lead	20 year lead	10 year lead	20 year lead	10 year lead	20 year lead
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Panel A: Dependent variable is log GDP									
	Panel, 1960-2000	Panel, 1960-2000	Panel, 1940-1980	Panel, 1940-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1940-1980	Panel, 1940-1980
Log Life Expectancy	1.35	1.70	0.82	0.74	1.09	0.29	1.37	0.97	0.83	0.97
	(0.49)	(0.45)	(0.34)	(0.41)	(0.44)	(0.62)	(0.37)	(0.52)	(0.27)	(0.33)
Number of observations	600	294	284	229	480	360	235	176	284	284
Number of countries	120	59	59	48	120	120	59	59	59	59
	Panel B: Dependent variable is log GDP									
	Just 1960 and 2000	Just 1960 and 2000	Just 1940 and 1980	Just 1940 and 1980	Just 1960 and 1990	Just 1960 and 1980	Just 1960 and 1990	Just 1960 and 1980	Just 1940 and 1980	Just 1940 and 1980
Log Life Expectancy	1.17	1.55	0.93	0.76	1.07	0.39	1.61	1.11	1.02	1.20
	(0.80)	(0.49)	(0.58)	(0.79)	(0.59)	(0.76)	(0.48)	(1.02)	(0.51)	(0.61)
Number of observations	240	118	96	74	240	240	118	116	96	96
Number of countries	120	59	48	37	120	120	59	58	48	48
	Panel C: Dependent variable is log GDP per capita									
	Panel, 1960-1990	Panel, 1960-1990	Panel, 1940-1980	Panel, 1940-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1960-1990	Panel, 1960-1980	Panel, 1940-1980	Panel, 1940-1980
Log Life Expectancy	-0.10	0.003	-0.27	-0.20	-0.63	-1.31	0.03	-0.001	-0.38	-0.14
	(0.48)	(0.46)	(0.26)	(0.36)	(0.51)	(0.69)	(0.50)	(0.75)	(0.25)	(0.37)
Number of observations	600	294	284	229	480	360	235	176	284	284
Number of countries	120	59	59	48	120	120	59	59	59	59
	Panel D: Dependent variable is log GDP per capita									
	Just 1960 and 2000	Just 1960 and 2000	Just 1940 and 1980	Just 1940 and 1980	Just 1960 and 1990	Just 1960 and 1980	Just 1960 and 1990	Just 1960 and 1980	Just 1940 and 1980	Just 1940 and 1980
Log Life Expectancy	-0.42	-0.19	-0.46	0.20	-0.84	-1.31	0.18	-0.48	-0.46	-0.17
	(0.82)	(0.76)	(0.40)	(0.66)	(0.70)	(0.85)	(0.82)	(1.18)	(0.53)	(0.64)
Number of observations	240	118	96	56	240	240	118	116	96	96
Number of countries	120	59	48	28	120	120	59	58	48	48

OLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panels A and C are unbalanced panels with one observation per decade. Panels B and D are long-difference specifications with observations for only the beginning and end dates. Dependent variable is log total GDP in Panels A and B and log GDP per capita in Panels C and D. Independent variable in all regressions is log life expectancy at birth. In columns 1-4, the dependent variable and independent variable are for the same time period; in columns 5-10, the dependent variable is for t+10 or t+20 as indicated, while the independent variable is for time t. "All countries" are those for which we have data on the dependent and independent variables. Base sample is countries for which we have disease data. Assignment of countries to low and middle income categories is based on 1940 income per capita; see text and Appendix A for details and definitions.

Table 4

The Effect of Interventions on Disease Mortality (zeroth stage)

	Dependent Variable is mortality per 100,000 from disease i in country j at period t						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A	*Base Sample*	*Base Sample*	*Base Sample*	*Without TB*	*Without pneumonia*	*Without malaria*	*Without influenza*
	Panel, 1930-1960						
Intervention	-46.04	-43.33	-46.04	-33.93	-36.31	-48.57	-48.62
	(9.40)	(10.36)	(9.40)	(8.66)	(8.99)	(9.23)	(9.69)
Lagged Intervention		-4.59					
		(8.05)					
Lead Intervention			20.57				
			(9.47)				
R-squared	0.52	0.47	0.47	0.49	0.48	0.48	0.48
Number of observations	1479	1479	1479	1327	1364	1361	1328
Panel B	*Just scarlet fever*	*Just typhoid*	*Just diphtheria*	*Just TB*	*Just pneumonia*	*Just malaria*	*Just influenza*
Intervention	-0.25	-8.84	-2.47	-108.51	-137.92	-19.97	-14.95
	(0.10)	(3.01)	(0.92)	(22.91)	(26.96)	(9.67)	(11.37)
R-squared	0.56	0.71	0.63	0.72	0.82	0.58	0.61
Number of observations	140	148	147	152	115	118	151
Number of countries	49	49	49	49	49	49	49

OLS regressions with a full set of disease, year, and country fixed effects. Robust standard errors, adjusted for clustering by country-disease pair, in parentheses. Unbalanced panels with data for 1930, 1940, 1950 and 1960. Data are stacked; dependent variable is deaths per 100,000 from disease i in country j at year t. Base sample is 10 infectious diseases plus cancer. Independent variables: dummy for intervention (e.g., for malaria equals 1 for 1950 and 1960, zero otherwise), dummy for lead intervention (e.g., for malaria equals 1 for 1940, 1950 and 1960), dummy for lagged intervention (e.g., for malaria equals 1 for 1950 and 1960).

Table 5

First Stage Estimates: Predicted Mortality and Life Expectancy

	Dependent Variable is log life expectancy									
	Baseline predicted mortality						TB, malaria, and pneumonia mortality only		Alternative timing	specific interventions
	Base Sample			*Including Eastern Europe*	*Low and Middle Income Countries Only*	*Balanced Panel Sample*	*Base Sample*	*Low and Middle Income Countries Only*	*Base Sample*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A	Panel, 1940-1980	Panel, 1930-1980	Panel, 1940-1970	Panel, 1940-1980	Panel, 1940-1980	Balanced Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980
Predicted Mortality	-0.35	-0.36	-0.31	-0.28	-0.28	-0.33	-0.39	-0.30	-0.35	-0.36
	(0.06)	(0.06)	(0.06)	(0.11)	(0.09)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)
R-squared	0.93	0.93	0.93	0.92	0.93	0.94	0.93	0.93	0.93	0.93
Number of observations	284	333	226	314	229	230	284	229	284	284
Number of countries	59	59	59	65	48	46	59	48	59	59
Panel B	Just 1940 and 1980	Just 1930 and 1980	Just 1940 and 1970	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980
Predicted Mortality	-0.47	-0.47	-0.40	-0.47	-0.35	-0.46	-0.52	-0.39	-0.47	-0.47
	(0.09)	(0.06)	(0.08)	(0.09)	(0.13)	(0.09)	(0.10)	(0.10)	(0.09)	(0.09)
R-squared	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.94	0.95	0.95
Number of observations	96	96	96	108	74	92	92	74	96	96
Number of countries	48	48	48	54	37	46	46	37	48	48

OLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panel A is unbalanced panel with one observation per decade. Panel B is long-difference specifications with observations for only the beginning and end dates. Dependent variable in both panels is log life expectancy at birth. Independent variable in columns 1-6 is baseline predicted mortality; in columns 7-8, predicted mortality constructed from tuberculosis, pneumonia, and malaria deaths only; in column 9, predicted mortality has alternative timing, and in column 10 predicted mortality has country-specific interventions. See text and Appendix A for the construction of the predicted mortality instrument, definitions and data sources. Eastern Europe is countries that became part of the Soviet bloc after 1945. Assignment of countries to low and middle income categories is based on 1940 income per capita. Balanced panel is countries with no missing data between 1940 and 1980.

Table 6

First Stage Estimates: Importance of Disease Composition

Dependent Variable is log life expectancy

Diseases used to calculate predicted mortality are indicated in each column

Base Sample

	All Diseases	*Without Tuberculosis*	*Without Malaria*	*Without Pneumonia*	*Without Influenza*	*Without Typhoid*	*Without Small Pox*	*Without Typhus*	*Without Cholera*	*Without Measles*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980
Predicted Mortality	-0.35	-0.46	-0.36	-0.42	-0.36	-0.36	-0.35	-0.36	-0.36	-0.36
	(0.06)	(0.07)	(0.08)	(0.13)	(0.06)	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)
R-squared	0.93	0.94	0.93	0.92	0.93	0.93	0.93	0.93	0.93	0.93
Number of observations	284	284	284	284	284	284	284	284	284	284
Number of countries	59	59	59	59	59	59	59	59	59	59
Panel B	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980
Predicted Mortality	-0.47	-0.62	-0.49	-0.51	-0.48	-0.47	-0.47	-0.47	-0.48	-0.48
	(0.09)	(0.10)	(0.12)	(0.20)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
R-squared	0.95	0.95	0.94	0.92	0.95	0.95	0.95	0.95	0.95	0.95
Number of observations	96	96	96	96	96	96	96	96	96	96
Number of countries	48	48	48	48	48	48	48	48	48	48

OLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panel A uses data from each 10 years in the indicated period, e.g., 1940-1980 is an unbalanced panel for 1940, 1950, 1960, 1970, and 1980; Panel B uses data from a balanced panel for just the start and end year indicated. Dependent variables: log life expectancy at birth. Independent variable: predicted mortality per 100 per annum. Base sample is countries for which we have disease data. Measure of predicted mortality in column 1 is baseline estimate, based on deaths from 15 infectious diseases. Other columns drop individual diseases from calculation of predicted mortality, as indicated in column heading.

Table 7

First Stage Estimates: Mean Reversion and Robustness

Dependent Variable is log life expectancy

OECD disease level after intervention

Base Sample

	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980 Using second lag of LE as instrument for lagged LE	Panel, 1940-1980 GMM (Arellano-Bond)	Panel, 1940-1980	Panel, 1930-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Balanced Panel, 1940-1980
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Base Sample										
Predicted Mortality	-0.35	-0.23	-0.30	-0.21	-0.22	-0.35	-0.23	-0.31	-0.20	-0.18
	(0.06)	(0.06)	(0.10)	(0.05)	(0.06)	(0.06)	(0.07)	(0.06)	(0.07)	(0.05)
Lagged Log Life Expectancy		0.46	0.40	0.68			0.46		0.49	0.57
		(0.09)	(0.17)	(0.07)			(0.09)		(0.11)	(0.06)
Lagged Predicted Mortality					-0.17				-0.005	
					(0.03)				(0.05)	
Lead Predicted Mortality						-0.017	-0.051			
						(0.027)	(1.210)			
Lagged Log GDP per capita								-0.04	-0.043	
								(0.03)	(0.020)	
p-value of test for 2nd order autocorrelation				0.99						
Hansen J Test (p-value)				0.003						
R-squared	0.93	0.95	0.95		0.94	0.93	0.95	0.93	0.95	0.96
Number of observations	284	284	248	248	284	333	284	284	289	280
Number of countries	59	59	59	59	59	59	59	59	62	56
Panel B: Low and Middle Income Countries										
Predicted Mortality	-0.28	-0.20	-0.38	-0.17	-0.20	-0.28	-0.20	-0.27	-0.17	-0.15
	(0.05)	(0.08)	(0.11)	(0.07)	(0.09)	(0.09)	(0.08)	(0.09)	(0.10)	(0.05)
Lagged Log Life Expectancy		0.37	0.16	0.60			0.37		0.40	0.52
		(0.11)	(0.17)	(0.10)			(0.11)		(0.13)	(0.07)
Lagged Predicted Mortality					-0.11				-0.02	
					(0.04)				(0.05)	
Lead Predicted Mortality						-0.001	-0.09			
						(0.03)	(1.35)			
Lagged Log GDP per capita								-0.04	-0.05	
								(0.04)	(0.03)	
p-value of test for 2nd order autocorrelation				0.79						
Hansen J Test (p-value)				0.018						
R-squared	0.93	0.94	0.95		0.93	0.93	0.94	0.93	0.94	0.96
Number of observations	229	229	196	196	199	267	229	219	219	180
Number of countries	48	48	48	48	42	48	48	48	48	36

OLS (columns 1-2 and 5-10) and 2SLS (columns 3-4) regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. All columns are unbalanced panels with one observation per decade: Panel A uses base sample of countries, Panel B uses only initially low and middle income countries. Dependent variable in both panels is log life expectancy at birth. Independent variables vary by column; lagged values are 10 years earlier and lead predicted mortality is 10 years ahead. Assignment of countries to low and middle income categories is based on 1940 income per capita. In column 3, the second lag of log life expectancy is used as an instrument for lagged log life expectancy. In column 4, we use the GMM of Arellano-Bond, with all available lags of log life expectancy as instruments. Balanced panel is countries with no missing data between 1940 and 1980.

Table 8

Falsification Exercise

	Base Sample	*Low & Mid. Income Countries Only*	*Base Sample*	*Low & Mid. Income Countries Only*	*Base Sample*	*Low & Mid. Income Countries Only*	*Base Sample*	*Low & Mid. Income Countries Only*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: dependent variable is: --	*change in life expectancy from 1900 to 1940*		*change in log population from 1900 to 1940*		*change in log GDP from 1900 to 1940*		*change in log GDP per capita from 1900 to 1940*	
Change in Predicted Mortality from 1940 to 1980	0.04	0.17	-0.06	-0.08	-0.18	-0.27	-0.12	-0.18
	(0.11)	(0.17)	(0.14)	(0.29)	(0.22)	(0.36)	(0.17)	(0.22)
R-squared	0.003	0.03	0.003	0.005	0.01	0.02	0.0095	0.01
Number of countries	29	19	29	19	29	19	29	19
Panel B: dependent variable is: --	*change in life expectancy from 1940 to 1980*		*change in log population from 1940 to 1980*		*change in log GDP from 1940 to 1980*		*change in log GDP per capita from 1940 to 1980*	
Change in Predicted Mortality from 1940 to 1980	-0.47	-0.35	-0.76	-0.65	-0.27	-0.03	0.48	0.59
	(0.06)	(0.09)	(0.15)	(0.21)	(0.25)	(0.32)	(0.17)	(0.23)
R-squared	0.53	0.32	0.31	0.19	0.003	0.0003	0.12	0.12
Number of countries	48	37	49	38	49	38	49	38

OLS regressions. Robust standard errors in parentheses. Both panels regress change in variable indicated from start to end date on change in predicted mortality from 1940 to 1980. Predicted mortality is deaths per 100 population. Panel A uses subset of base sample for which data on all outcome variables are available and for which there is no discontinuity in boundaries of country in Maddison's data during the relevant period.

Table 9

The Effect of Life Expectancy on Log Population: 2SLS Estimates

	Dependent variable is log population								
	Baseline instrument				Just TB, pneumonia and malaria mortality	Baseline instrument			
	Base Sample		*Including Eastern Europe*	*Low and Middle Income Countries Only*	*Base Sample*				
	No leads Panel, 1940-1980	No leads Panel, 1930-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	10 year lead Panel, 1940-1980	20 year lead Panel, 1940-1980	30 year lead Panel, 1940-1970	40 year lead Panel, 1940-1960
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Log Life Expectancy	1.30	1.47	1.53	1.42	1.35	1.46	1.46	1.37	1.10
	(0.38)	(0.39)	(0.43)	(0.71)	(0.35)	(0.38)	(0.36)	(0.38)	(0.38)
Number of observations	284	333	314	229	284	284	284	226	167
Number of countries	59	59	65	48	59	59	59	59	59
	No leads Just 1940 and 1980	No leads Just 1930 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	10 year lead Just 1940 and 1980	20 year lead Just 1940 and 1980	30 year lead Just 1940 and 1980	40 year lead Just 1940 and 1980
Panel B									
Log Life Expectancy	1.64	1.87	1.83	1.84	1.67	1.75	1.69	1.87	1.90
	(0.51)	(0.53)	(0.54)	(1.00)	(0.46)	(0.48)	(0.44)	(0.49)	(0.59)
Number of observations	96	96	108	74	96	96	96	96	96
Number of countries	48	48	54	37	48	48	48	48	48

2SLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panel A is unbalanced panel with one observation per decade. Panel B is long-difference specification with observations for only the beginning and end dates. Dependent variable in both panels is log total population. Independent variable in both panels is log life expectancy at birth. In columns 1-4 and 6-9, log life expectancy is instrumented by predicted mortality (baseline instrument), and in column 5 it is instrumented by predicted mortality constructed from tuberculosis, pneumonia and malaria only. First stages are in Table 5. In columns 1-5, the dependent and independent variables are for the same time period; in columns 6-9, the dependent variable is t+10, t+20 etc., as indicated, while the independent variable is at time t. See text and Appendix A for construction of the predicted mortality instrument, definitions and data sources.

Table 10
The Effect of Life Expectancy on Log Births: 2SLS Estimates

	Dependent variable is log total births								
	Baseline instrument				Just TB, pneumonia and malaria mortality	Baseline instrument			
	Base Sample		*Including Eastern Europe*	*Low and Middle Income Countries Only*	*Base Sample*				
	No leads	No leads	No leads	No leads	No leads	10 year lead	20 year lead	30 year lead	40 year lead
	Panel, 1940-1980	Panel, 1930-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980	Panel, 1940-1980
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Log Life Expectancy	2.33	2.23	2.60	2.85	2.33	1.64	1.78	1.02	0.04
	(0.77)	(0.67)	(0.83)	(1.59)	(0.75)	(0.48)	(0.60)	(0.57)	(0.49)
Number of observations	233	277	261	178	233	234	187	140	93
Number of countries	47	47	53	36	47	47	47	47	47
	No leads	No leads	No leads	No leads	No leads	10 year lead	20 year lead	30 year lead	40 year lead
	Just 1940 and 1980	Just 1930 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1980	Just 1940 and 1970	Just 1940 and 1960	Just 1940 and 1950
Panel B									
Log Life Expectancy	2.49	2.31	2.66	2.75	2.45	1.60	1.48	0.84	0.05
	(0.84)	(0.75)	(0.87)	(1.62)	(0.81)	(0.57)	(0.61)	(0.60)	(0.48)
Number of observations	90	88	98	68	90	90	90	90	90
Number of countries	45	44	49	34	45	45	45	45	45

2SLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panel A is unbalanced panel with one observation per decade. Panel B is long-difference specification with observations for only the beginning and end dates. Dependent variable in both panels is log total number of births. Independent variable in both panels is log life expectancy at birth. In columns 1-4 and 6-9, log life expectancy is instrumented by predicted mortality (baseline instrument), and in column 5 it is instrumented by predicted mortality constructed from tuberculosis, pneumonia and malaria only. First stages are in Table 5. In columns 1-5, the dependent and independent variables are for the same time period; in columns 6-9, the dependent variable is t+10, t+20 etc., as indicated, while the independent variable is at time t. See text and Appendix A for construction of the predicted mortality instrument, definitions and data sources.

Table 11

The Effect of Life Expectancy on Log GDP: 2SLS Estimates

	Dependent variable is log GDP								
	Baseline instrument				Just TB, pneumonia and malaria mortality	Baseline instrument			
	Base Sample		*Including Eastern Europe*	*Low and Middle Income Countries Only*	*Base Sample*				
	No leads Panel, 1940-1980	No leads Panel, 1930-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	10 year lead Panel, 1940-1980	20 year lead Panel, 1940-1980	30 year lead Panel, 1940-1980	40 year lead Panel, 1940-1980
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Log Life Expectancy	0.17	0.33	0.32	0.06	0.43	0.72	0.84	0.87	0.41
	(0.57)	(0.57)	(0.59)	(0.92)	(0.58)	(0.45)	(0.52)	(0.60)	(0.67)
Number of observations	284	333	314	229	284	284	284	226	167
Number of countries	59	59	65	48	59	59	59	59	59
	No leads Just 1940 and 1980	No leads Just 1930 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	10 year lead Just 1940 and 1980	20 year lead Just 1940 and 1980	30 year lead Just 1940 and 1970	40 year lead Just 1940 and 1960
Panel B									
Log Life Expectancy	0.54	0.73	0.70	-0.05	0.87	0.91	1.07	1.05	0.79
	(0.75)	(0.80)	(0.76)	(1.29)	(0.74)	(0.71)	(0.76)	(0.75)	(0.78)
Number of observations	96	96	108	74	96	96	96	96	96
Number of countries	48	48	54	37	48	48	48	48	48

2SLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panel A is unbalanced panel with one observation per decade. Panel B is long-difference specification with observations for only the beginning and end dates. Dependent variable in both panels is log total GDP. Independent variable in both panels is log life expectancy at birth. In columns 1-4 and 6-9, log life expectancy is instrumented by predicted mortality (baseline instrument), and in column 5 it is instrumented by predicted mortality constructed from tuberculosis, pneumonia and malaria only. First stages are in Table 5. In columns 1-5, the dependent and independent variables are for the same time period; in columns 6-9, the dependent variable is t+10, t+20 etc., as indicated, while the independent variable is at time t. See text and Appendix A for construction of the predicted mortality instrument, definitions and data sources.

Table 12

The Effect of Life Expectancy on Log GDP per capita: 2SLS Estimates

	Dependent variable is log GDP per capita								
	Baseline instrument				Just TB, pneumonia and malaria mortality	Baseline instrument			
	Base Sample		*Including Eastern Europe*	*Low and Middle Income Countries Only*	*Base Sample*	*Base Sample*	*Base Sample*	*Base Sample*	*Base Sample*
	No leads Panel, 1940-1980	No leads Panel, 1930-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	10 year lead Panel, 1940-1980	20 year lead Panel, 1940-1980	30 year lead Panel, 1940-1980	40 year lead Panel, 1940-1980
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Log Life Expectancy	-1.09	-1.07	-1.16	-1.27	-0.90	-0.74	-0.62	-0.50	-0.69
	(0.51)	(0.46)	(0.53)	(0.93)	(0.53)	(0.42)	(0.56)	(0.61)	(0.77)
Number of observations	284	333	314	0.94	284	284	284	226	167
Number of countries	59	59	65	48	59	59	59	59	59
	No leads Just 1940 and 1980	No leads Just 1930 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	10 year lead Just 1940 and 1980	20 year lead Just 1940 and 1980	30 year lead Just 1940 and 1970	40 year lead Just 1940 and 1960
Panel B									
Log Life Expectancy	-1.07	-1.05	-1.40	-1.82	-0.79	-0.84	-0.62	-0.44	-0.48
	(0.58)	(0.55)	(0.78)	(1.16)	(0.61)	(0.74)	(0.84)	(0.81)	(0.91)
Number of observations	96	96	114	74	96	96	96	96	96
Number of countries	48	48	57	37	48	48	48	48	48

2SLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panel A is unbalanced panel with one observation per decade. Panel B is long-difference specification with observations for only the beginning and end dates. Dependent variable in both panels is log GDP per capita. Independent variable in both panels is log life expectancy at birth. In columns 1-4 and 6-9, log life expectancy is instrumented by predicted mortality (baseline instrument), and in column 5 it is instrumented by predicted mortality constructed from tuberculosis, pneumonia and malaria only. First stages are in Table 5. In columns 1-5, the dependent and independent variables are for the same time period; in columns 6-9, the dependent variable is t+10, t+20 etc., as indicated, while the independent variable is at time t. See text and Appendix A for construction of the predicted mortality instrument, definitions and data sources.

Table 13

The Effect of Life Expectancy on Log GDP per population of working age: 2SLS Estimates

	Dependent variable is log GDP per capita								
	Baseline instrument				Just TB, pneumonia and malaria mortality	Baseline instrument			
	Base Sample		*Including Eastern Europe*	*Low and Middle Income Countries Only*	*Base Sample*	*Base Sample*	*Base Sample*	*Base Sample*	*Base Sample*
	No leads Panel, 1940-1980	No leads Panel, 1930-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	No leads Panel, 1940-1980	10 year lead Panel, 1940-1980	20 year lead Panel, 1940-1980	30 year lead Panel, 1940-1980	40 year lead Panel, 1940-1980
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A									
Log Life Expectancy	-1.41	-1.36	-1.46	-1.92	-1.22	-0.88	-1.04	-0.90	-1.20
	(0.61)	(0.56)	(0.63)	(1.35)	(0.63)	(0.40)	(0.49)	(0.57)	(0.94)
Number of observations	234	280	264	179	234	234	234	187	167
Number of countries	47	47	53	36	47	47	47	47	59
	No leads Just 1940 and 1980	No leads Just 1930 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	No leads Just 1940 and 1980	10 year lead Just 1940 and 1980	20 year lead Just 1940 and 1980	30 year lead Just 1940 and 1970	40 year lead Just 1940 and 1960
Panel B									
Log Life Expectancy	-1.33	-1.28	-1.36	-2.26	-1.09	-1.16	-1.16	-1.00	-1.12
	(0.57)	(0.60)	(0.57)	(1.25)	(0.56)	(0.66)	(0.76)	(0.78)	(1.07)
Number of observations	92	92	104	70	92	92	92	92	92
Number of countries	46	46	52	35	46	46	46	46	46

2SLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Panel A is unbalanced panel with one observation per decade. Panel B is long-difference specification with observations for only the beginning and end dates. Dependent variable in both panels is log GDP per working age population. Independent variable in both panels is log life expectancy at birth. In columns 1-4 and 6-9, log life expectancy is instrumented by predicted mortality (baseline instrument), and in column 5 it is instrumented by predicted mortality constructed from tuberculosis, pneumonia and malaria only. First stages are in Table 5. In columns 1-5, the dependent and independent variables are for the same time period; in columns 6-9, the dependent variable is t+10, t+20 etc., as indicated, while the independent variable is at time t. See text and Appendix A for construction of the predicted mortality instrument, definitions and data sources.

Table 14

The Effect of Life Expectancy on Years of Schooling: 2SLS Estimates

Dependent variable is years of schooling

	OLS	Baseline instrument	Baseline instrument	OLS	Baseline instrument	Baseline instrument	OLS	Baseline instrument	Baseline instrument
	Base Sample		*Low and Middle Income Countries Only*	*Base Sample*		*Low and Middle Income Countries Only*	*Base Sample*		*Low and Middle Income Countries Only*
	10 year lead	10 year lead	10 year lead	20 year lead	20 year lead	20 year lead	30 year lead	30 year lead	30 year lead
	Panel, 1950-1980	Panel, 1950-1980	Panel, 1950-1980	Panel, 1950-1970	Panel, 1950-1970	Panel, 1950-1970	Panel, 1950-1960	Panel, 1950-1960	Panel, 1950-1960
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Log Life Expectancy	-0.71	-0.15	-0.73	-0.12	0.22	1.10	4.71	1.75	-1.40
	(1.44)	(3.52)	(4.98)	(1.56)	(3.56)	(5.01)	(1.47)	(2.88)	(4.14)
Number of observations	224	224	168	168	168	126	112	112	84
Number of countries	56	56	42	56	56	42	56	56	42

OLS and 2SLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses.Unbalanced panel with one observation per decade. Dependent variable is years of schooling. Independent variable is log life expectancy at birth. In columns 2, 3, 5, 6, 8 and 9, log life expectancy is instrumented by predicted mortality (baseline instrument). First stages are in Table 5. In columns 1-3, the dependent and independent variables are for the same time period; in columns 4-9, the dependent variable is t+10, t+20, and t+30 as indicated, while the independent variable is at time t. See text and Appendix A for construction of the predicted mortality instrument, definitions and data sources.

Table 15
2SLS Estimates: robustness

Dependent variable indicated for each panel separately

	Base Sample								
			ρ=0.4	ρ=0.45	ρ=0.5	ρ=0.55	ρ=0.6	ρ=0.65	ρ=0.7	ρ=0.75
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Dependent variable is--	*Log GDP per capita*		*Transformed log GDP per capita*							
Lagged GDP per capita	0.41	0.52								
	(0.15)	(0.15)								
Log Life Expectancy	-0.31	-0.26								
	(0.36)	(0.35)								
Transformed Log Life Expectancy			-1.38	-1.40	-1.42	-1.45	-1.47	-1.50	-1.54	-1.58
			(0.62)	(0.64)	(0.66)	(0.69)	(0.73)	(0.78)	(0.84)	(0.92)
Panel B: Dependent variable is--	*Log life expectancy*		*Transformed log life expectancy*							
Predicted Mortality	n.a.	n.a.	-0.23	-0.22	-0.20	-0.19	-0.18	-0.17	-0.15	-0.14
			(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)
R-squared			0.85	0.83	0.8	0.75	0.69	0.62	0.53	0.44
Number of observations			273	273	273	273	273	273	273	273
Number of countries			59	59	59	59	59	59	59	59

2SLS regressions with a full set of year fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. All panels have unbalanced panels with one observation per decade. Columns 1 and 2 are estimated using GMM of Arellano-Bond; dependent variable is differenced so there are no country fixed effects. Both columns use all available lags of lagged dependent variable and the contemporaneous value of predicted mortality as instruments. Column 1 uses lagged values of log life expectancy as additional instruments, while column 2 does not. In columns 3 through 10, transformed variables are defined as x(t)-ρx(t-1), where value of ρ is indicated in the column heading; instrument is predicted mortality and there is a full set of country fixed effects.

Table 16

Interactions with Initial Conditions: 2SLS Estimates

	Interaction with Log GDP per capita in 1940					Interaction with investment as share of GDP in 1940s				
	No lead	10 year lead	20 year lead	30 year lead	40 year lead	No lead	10 year lead	20 year lead	30 year lead	40 year lead
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: dependent variable is log population										
Log Life Expectancy	1.57	1.70	1.73	1.65	1.56	1.33	1.47	1.46	1.33	1.04
	(0.35)	(0.38)	(0.37)	(0.41)	(0.48)	(0.31)	(0.32)	(0.32)	(0.34)	(0.34)
Log Life Expectancy	0.34	0.28	-0.14	-0.07	0.15	0.01	-0.01	-0.02	-0.04	-0.04
x Interaction term	(0.61)	(0.61)	(0.59)	(0.58)	(0.61)	(0.09)	(0.10)	(0.09)	(0.10)	(0.10)
Number of observations	293	293	293	245	196	240	240	240	192	144
Number of countries	49	49	49	49	49	48	48	48	48	48
Panel B: dependent variable is log total GDP										
Log Life Expectancy	0.65	1.05	1.48	1.69	1.48	0.53	1.02	1.09	1.05	0.35
	(0.44)	(0.43)	(0.51)	(0.63)	(0.69)	(0.43)	(0.34)	(0.47)	(0.52)	(0.62)
Log Life Expectancy	1.19	1.65	1.94	1.87	1.34	0.21	0.12	0.11	0.02	-0.07
x Interaction term	(0.80)	(0.67)	(1.11)	(1.37)	(1.46)	(0.09)	(0.10)	(0.13)	(0.16)	(0.14)
Number of observations	293	293	293	245	196	240	240	240	192	144
Number of countries	49	49	49	49	49	48	48	48	48	48
Panel C: dependent variable is log GDP per capita										
Log Life Expectancy	-0.89	-0.65	-0.25	0.04	-0.08	-0.79	-0.45	-0.37	-0.28	-0.69
	(0.40)	(0.42)	(0.52)	(0.62)	(0.68)	(0.39)	(0.44)	(0.61)	(0.67)	(0.81)
Log Life Expectancy	0.66	1.29	1.80	1.81	1.19	-0.19	0.14	0.13	0.05	-0.03
x Interaction term	(0.62)	(0.65)	(1.10)	(1.27)	(1.30)	(0.13)	(0.11)	(0.16)	(0.17)	(0.18)
Number of observations	293	293	293	245	196	240	240	240	192	144
Number of countries	49	49	49	49	49	48	48	48	48	48

2SLS regressions with a full set of year and country fixed effects. Robust standard errors, adjusted for clustering by country, in parentheses. Unbalanced panels with one observation per decade. Dependent variable: in Panel A, log total population; in Panel B, log total GDP; in Panel C, log GDP per capita. Independent variable in all panels is log life expectancy at birth and interaction of log life expectancy with, in columns 1-5, log GDP per capita in 1940, and in columns 6-10, investment share of GDP in 1940s. All variables are demeaned so main effects are evaluated at sample mean. In all columns, instruments are predicted mortality (baseline instrument) and interaction of predicted mortality with either log GDP p.c. in 1940 (columns 1-5) or investment share of GDP 1940s (columns 6-10). First stages not reported to save space. In columns 1 and 6, the dependent and independent variables are for the same time period; in columns 2-5 and 7-9, the dependent variable is t+10, t+20 etc., as indicated, while the independent variable is at time t. See text and Appendix A for details and definitions.

Figure 1: Log life expectancy at birth for initially rich, middle-income and poor countries

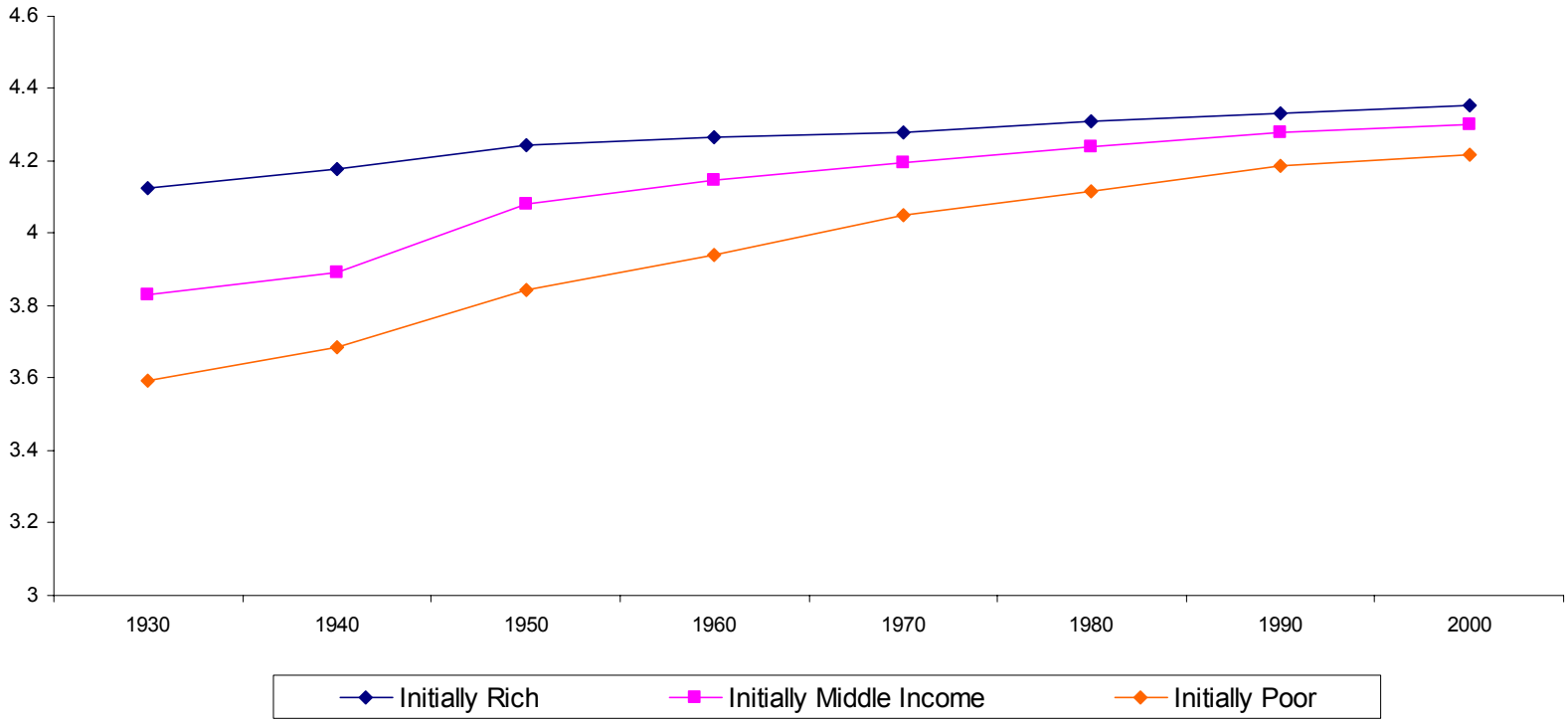

Figure 2: Log GDP per capita for initially rich, middle-income and poor countries

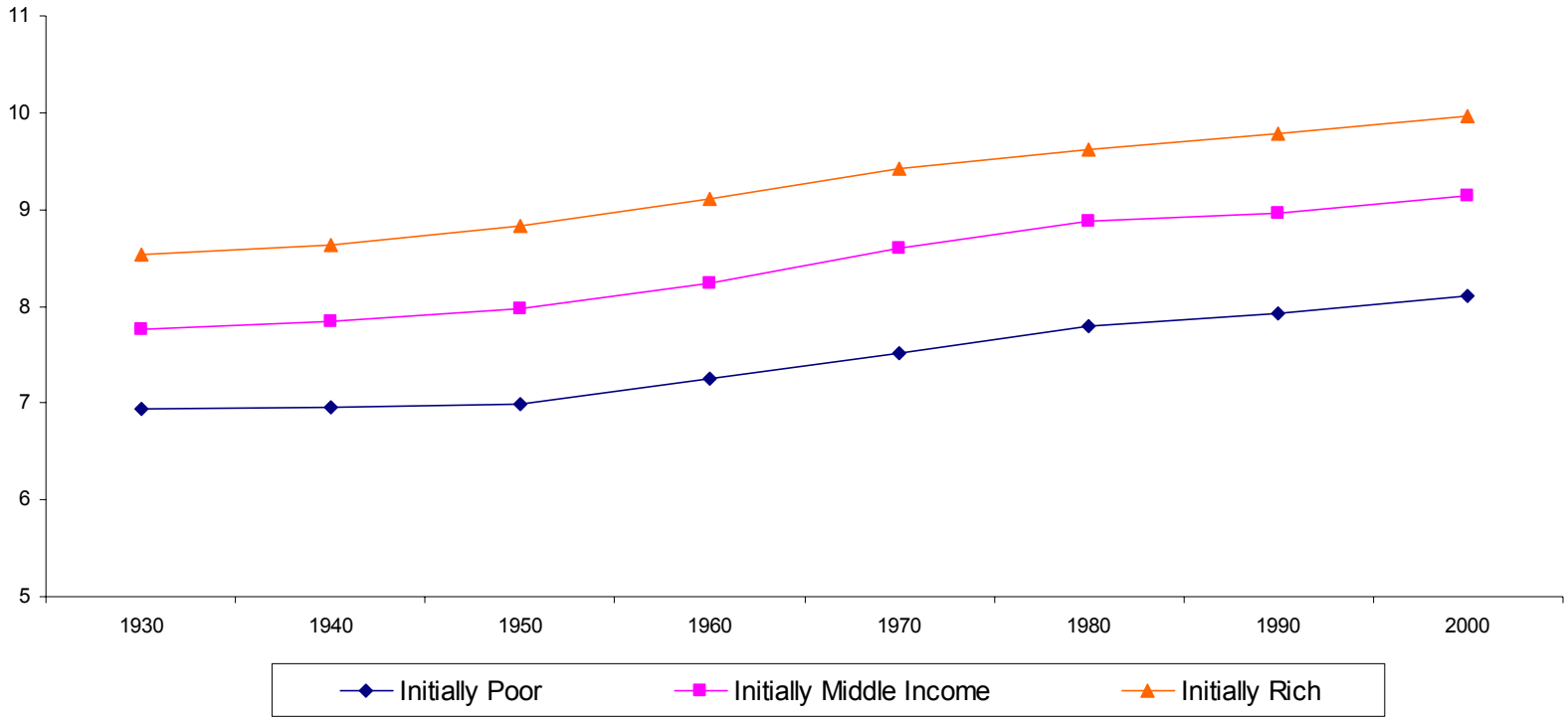

Figure 3: Average mortality rates from cancer, tuberculosis and pneumonia (deaths per 100,000 per annum)

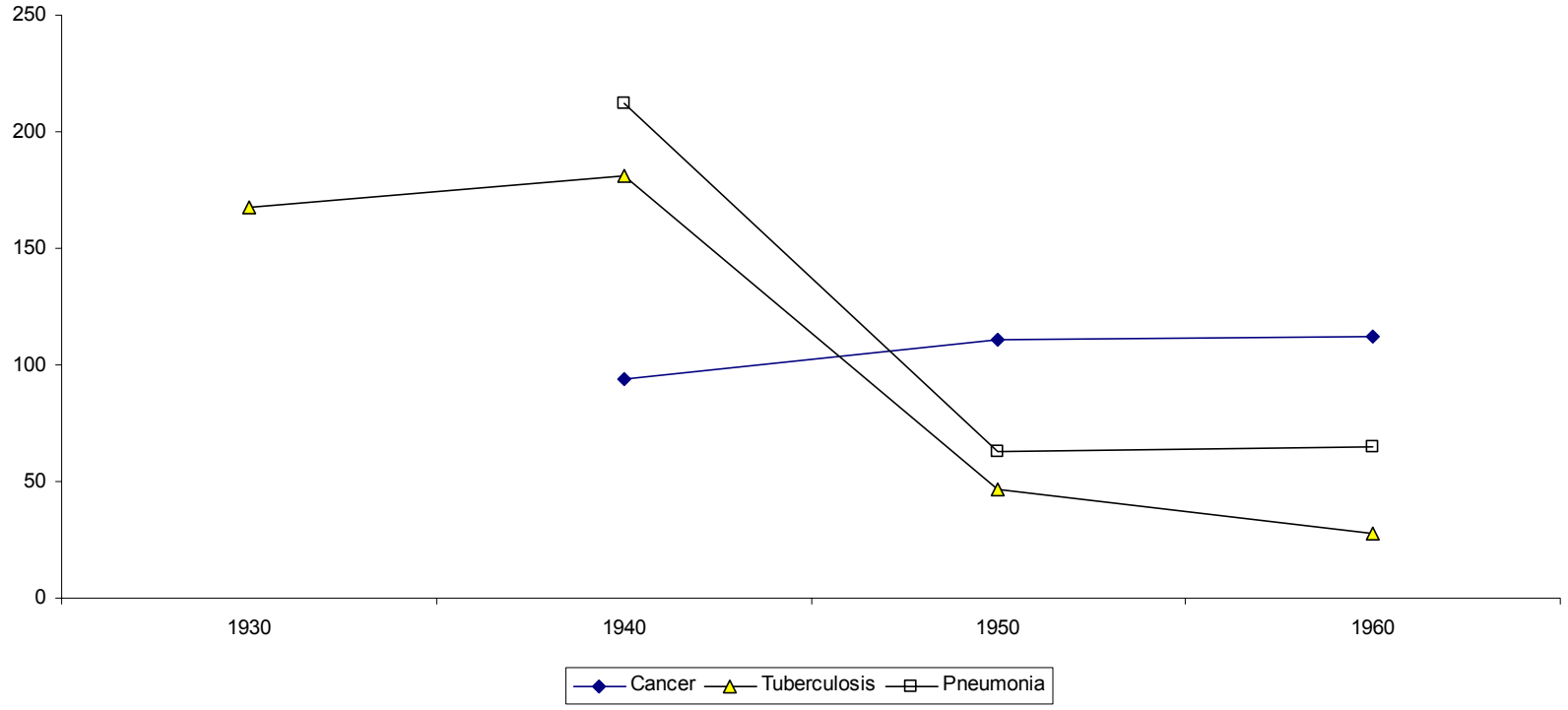

Figure 4: Change in log life expectancy and change in predicted mortality, 1940-80, base sample

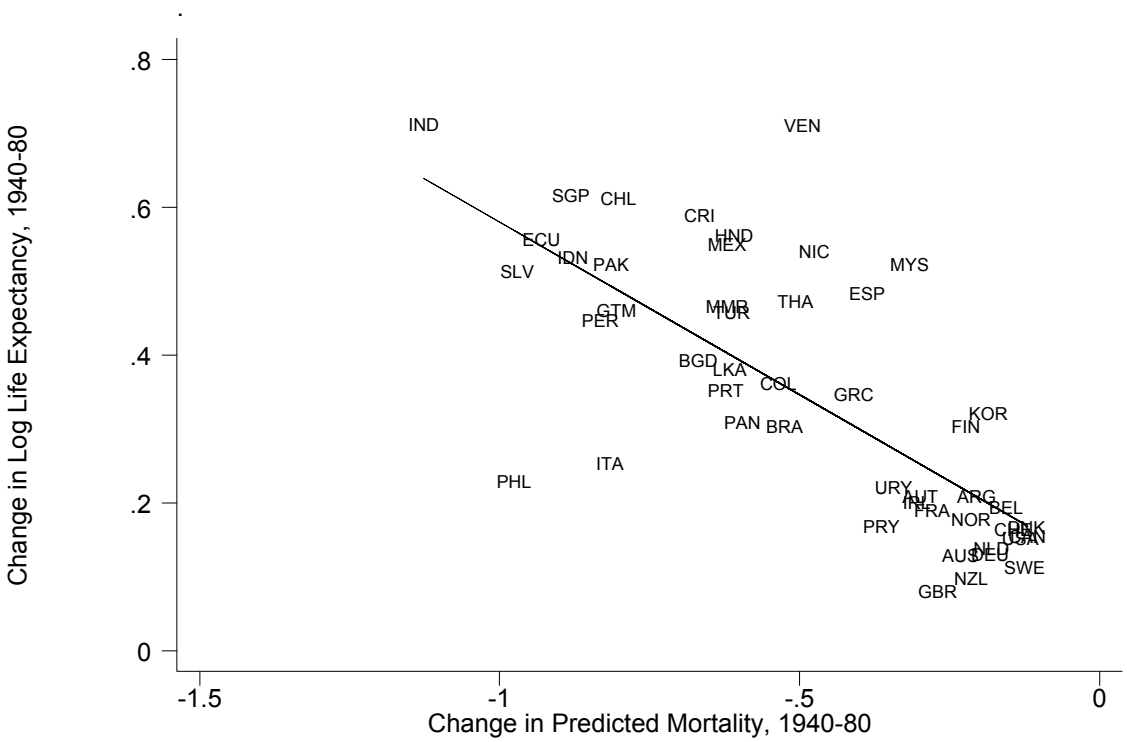

Figure 5: Change in log life expectancy and change in predicted mortality, 1940-80, low and middle-income countries

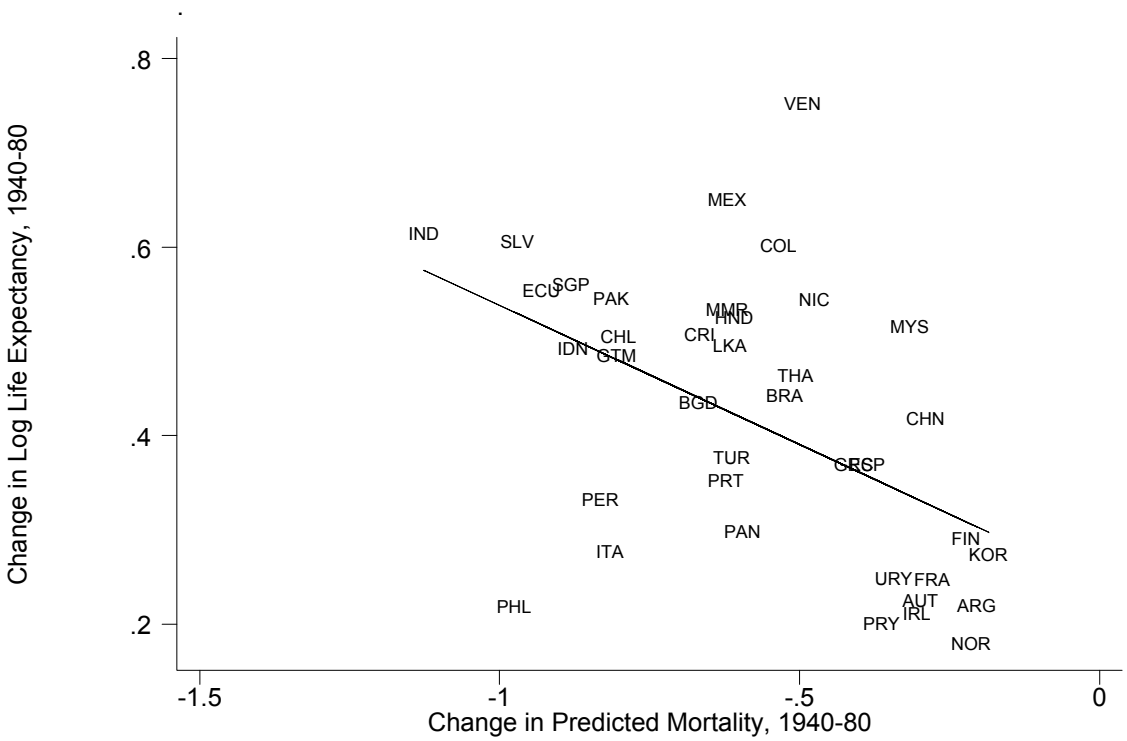

Figure 6: Change in log life expectancy, 1900-40, and change in predicted mortality, 1940-80, base sample

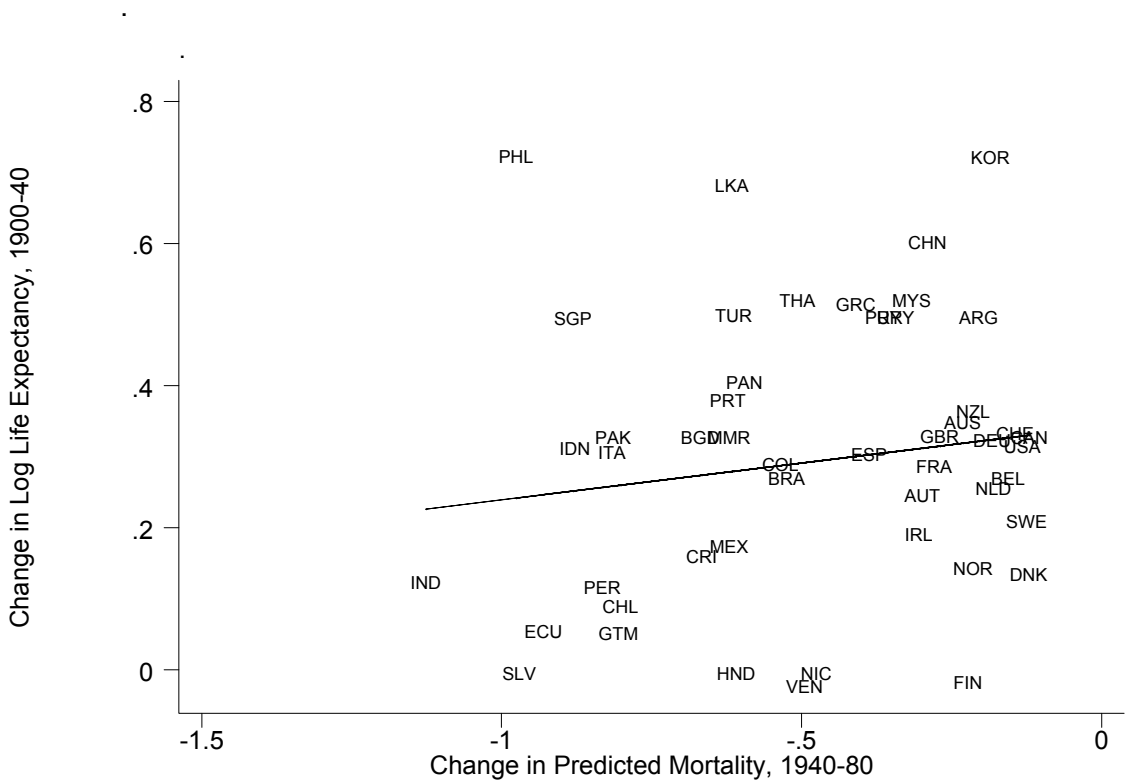

Figure 7: Change in log life expectancy, 1900-40, against change in predicted mortality, 1940-80, low and middle-income countries

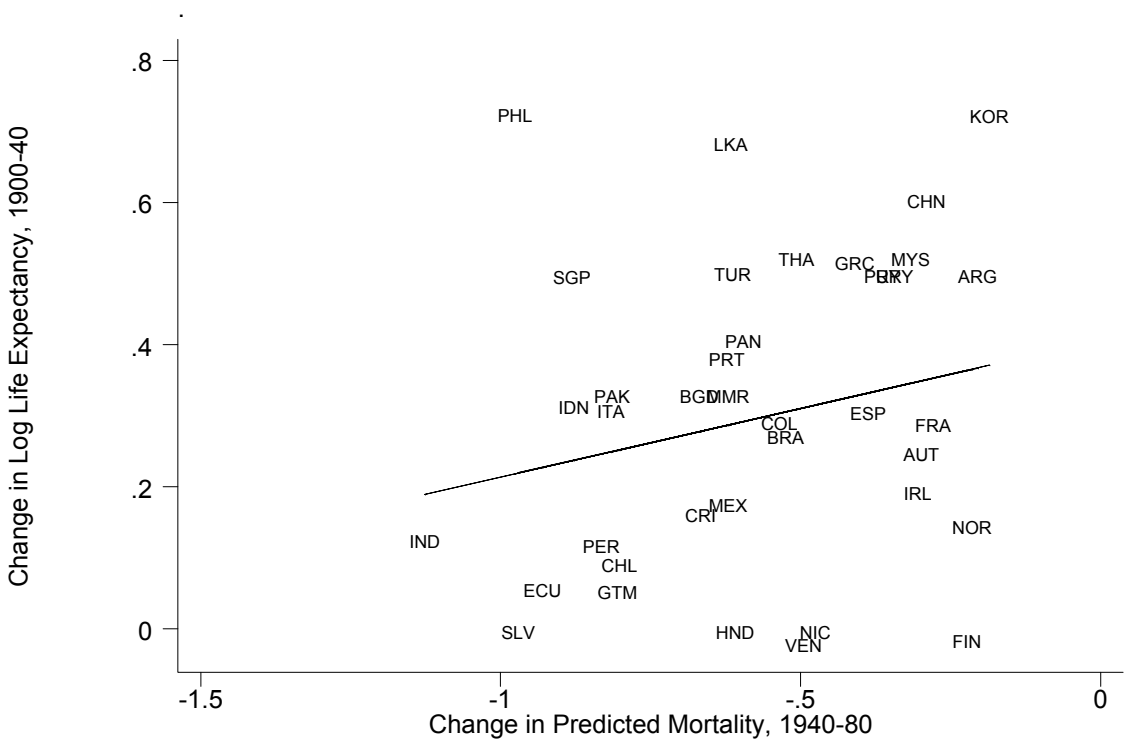

Figure 8: Change in log life expectancy, 1930-40, and change in predicted mortality, 1940-80, base sample

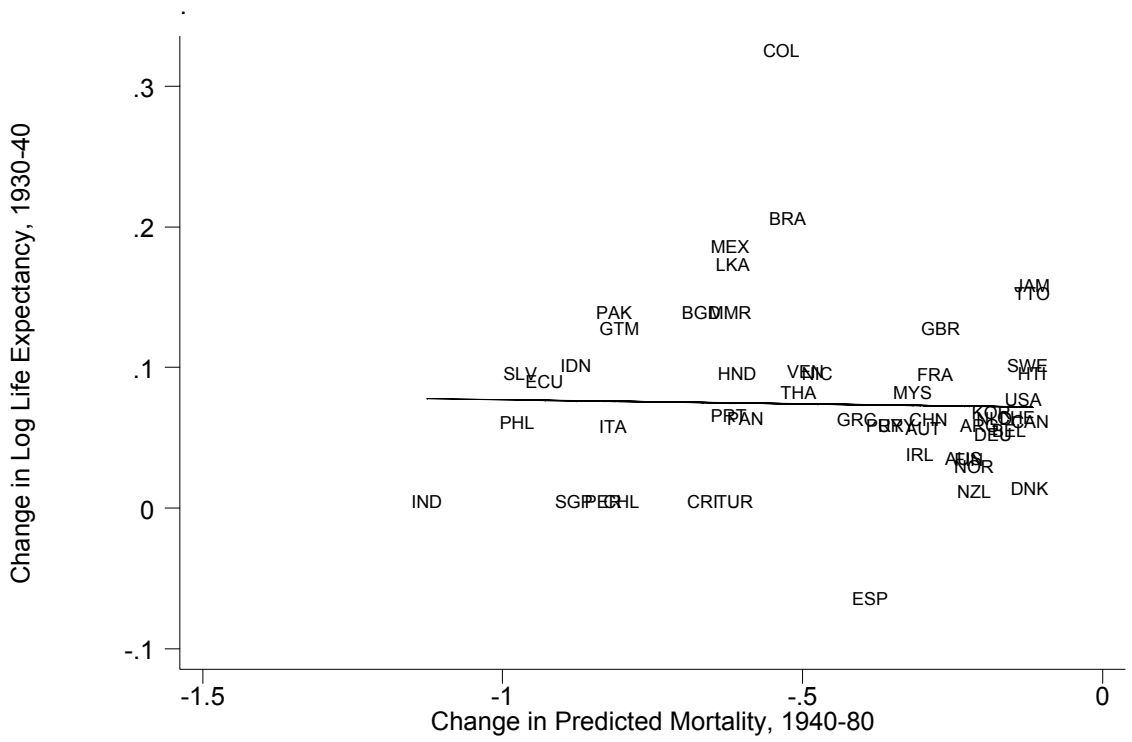

Figure 9**:** Change in log life expectancy, 1930-40, and change in predicted mortality, 1940-80, low and middle-income countries

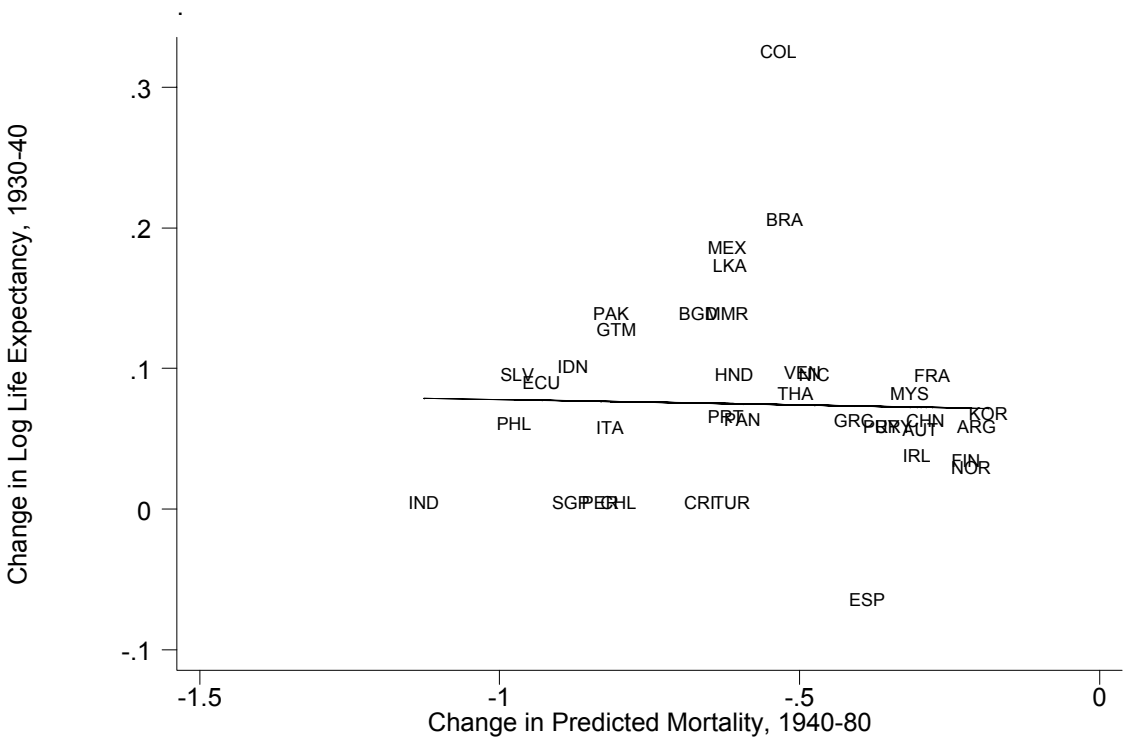

Figure 10: Change in log population and change in predicted mortality, 1940-80, base sample

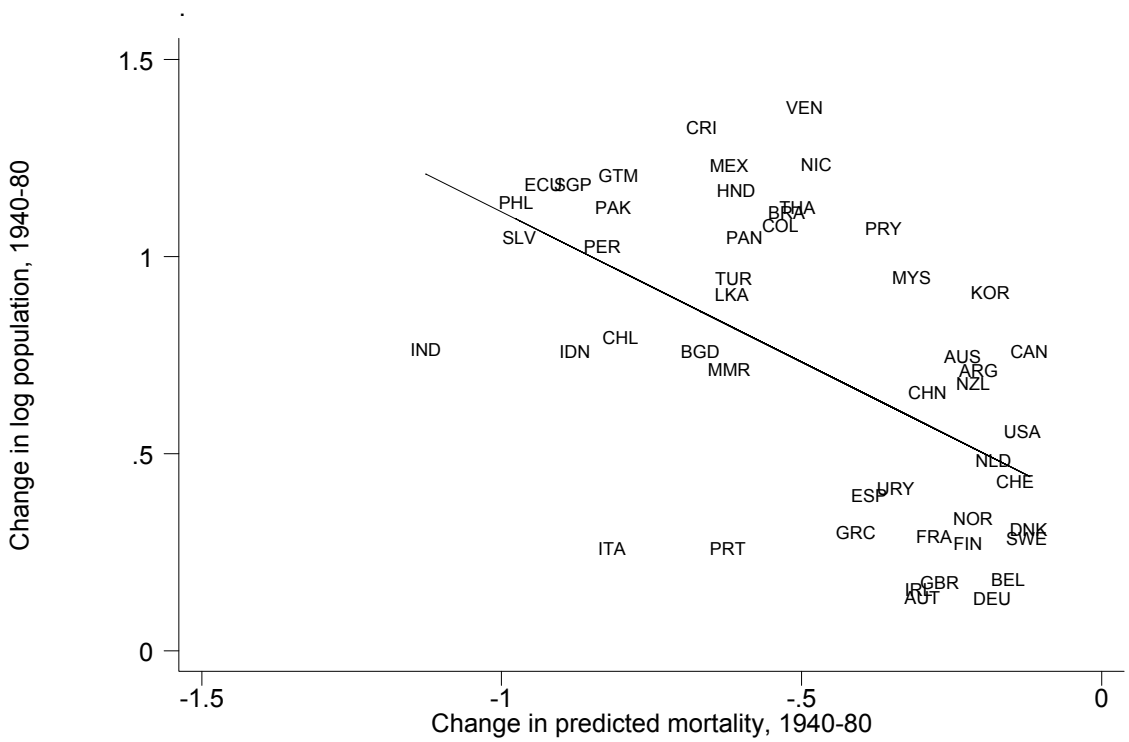

Figure 11: Change in log total births and change in predicted mortality, 1940-80, base sample

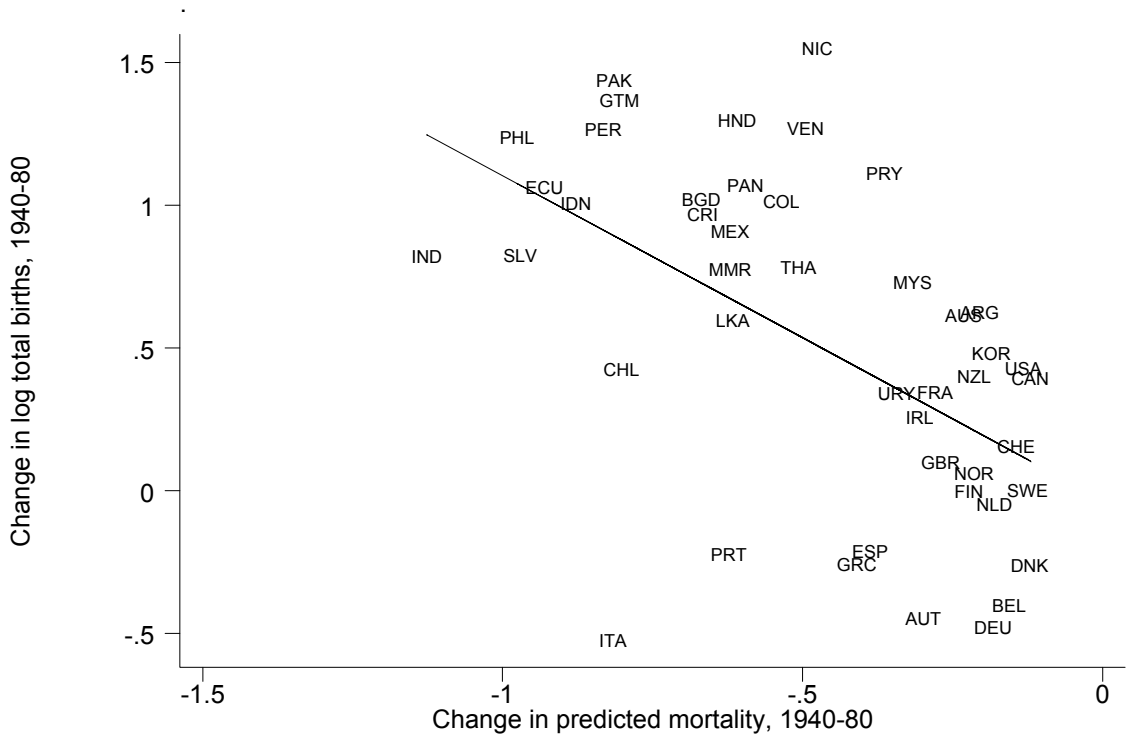

Figure 12: Change in log total GDP and change in predicted mortality, 1940-80, base sample

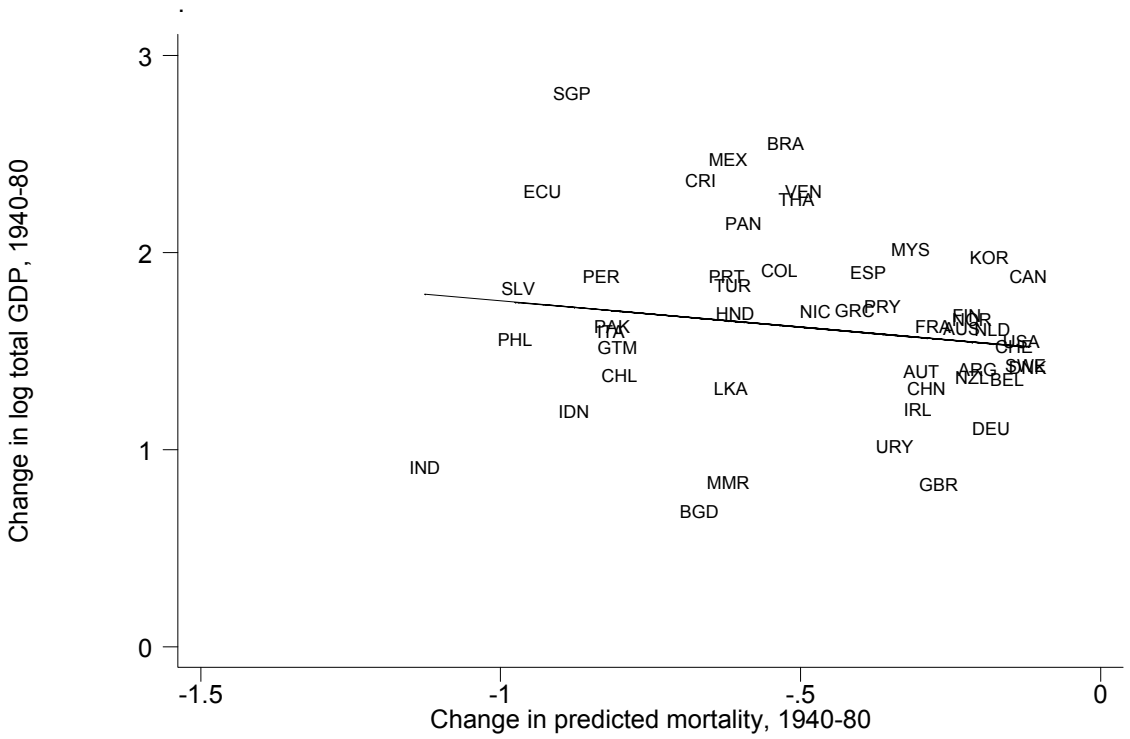

Figure 13: Change in log GDP per capita and change in predicted mortality, 1940-80, base sample

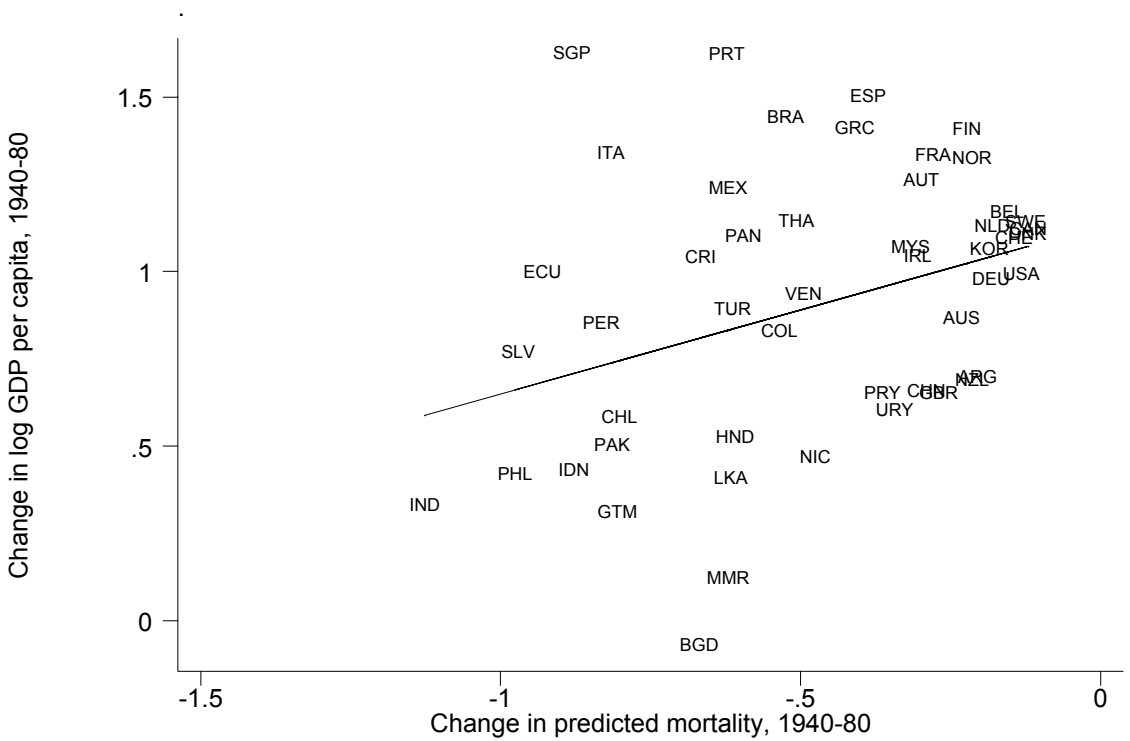

Figure 14: Change in log GDP per working age population and change in predicted mortality, 1940-80, base sample

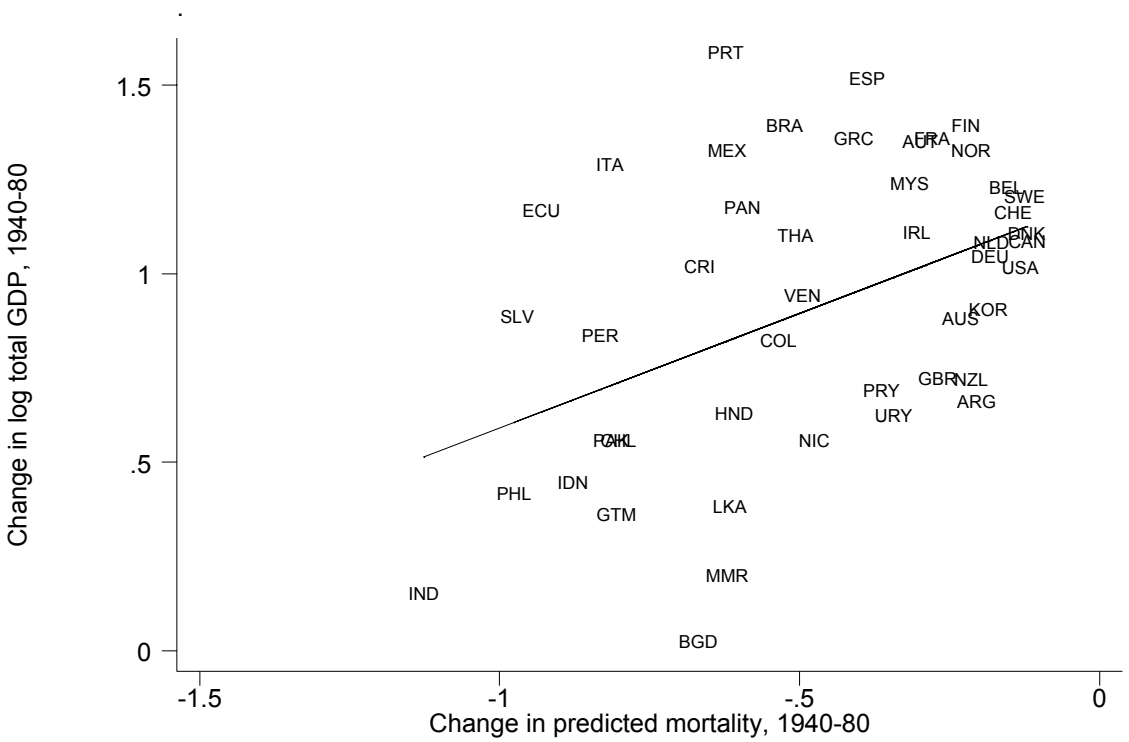

Figure 15: Change in years of schooling, 1960-90, and change in predicted mortality, 1940-80, base sample