

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

Appendices B and C (not for publication)

Daron Acemoglu

Simon Johnson

MIT

MIT

December 18, 2007

Appendix B: List and Further Information 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 details on the remaining 12 infectious diseases, in rough descending order of their contribution to global deaths around 1940 (see Kiple, 1993, Hoff and Smith, 2000, Heymann 2004, and the Centers for Disease Control and Prevention website). 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. It is useful to note that the timing of interventions would not be changed if we would put greater emphasis on sulphur drugs. Sulfonamide drugs were invented in the 1930s, but were often toxic and not available in the most effective doses (see Conybeare, 1948, pp. 65-66). This changed only from 1939, when the drugs became more effective (though Loudon, 2002, puts the useful breakthrough a little earlier).

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 intravenous 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 introduction of antibiotics in the 1940s as the intervention date for both our baseline and alternative instruments.

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 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 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 (which is one component of the DTP vaccine). The vaccine became available in the 1920s. We take the 1940s as our intervention date both for 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.¹ 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 *Aedes aegypti* mosquito. It is controlled by vaccination and public health measures against the

¹This 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).

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 and was widely used in the 1940s. 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 introduction of the antibiotics in the 1940s as the intervention date for both our baseline and alternative instruments.

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. Once again, based on antibiotics, the 1940s are the intervention date for both our baseline and alternative instruments.

Appendix C: Further Information on Sources, Data and Coding

This appendix explains in detail the construction and sources of our life expectancy estimates for 1930 and 1940.

Sources

The underlying data sources are national authorities. We take the data as reported to and republished by the United Nations. However, the UN does not appear to have produced a comprehensive set of data for the pre-1950 period; later publications sometimes revised estimates, but they also dropped earlier data.

At regular intervals, the UN’s Demographic Yearbooks focused on mortality and we concentrate our attention on these issues.² The UN’s own advice on accessing early data is helpful but not sufficient. Specifically, the introduction (p. 1) of the United Nations Demographic Yearbook 1978, Historical Supplement (Special Issue), published in 1979 (but really dealing with data only since 1950), suggests that readers should refer to the Demographic Yearbook 1948 for “many of the same tables showing annual data for the period 1932 to 1947” and should use this in conjunction with the historical supplement.³ This advice is not sufficient, however, because the UN added and revised historical data, as far as we can determine, in its Demographic Yearbooks through 1968.⁴

We also pay close attention to Preston’s (1975) data as this represents what a leading scholar decided could be used alongside the UN’s data (at a time when the UN’s historical data had been revised and extended). Preston’s (1975, Table A-2) estimates are “in the 1930s”

²The U.N. Demographic Yearbook 1948 (United Nations 1949) reports life expectancy at birth on pp. 514-523. The U.N. Demographic Yearbook 1951 (United Nations 1951) reports life expectancy at birth on pp. 526-539. The U.N. Demographic Yearbook 1961 (United Nations 1961) reports life expectancy at birth on pp. 622-641. The U.N. Demographic Yearbook 1967 (United Nations 1968) reports life expectancy at birth on pp. 704-741.

³Life expectancy at birth for both males and females, generally from 1950-55, in 5 year intervals, is on pp. 542-563 of this historical supplement.

⁴The publication date for the UN Demographic Yearbooks does not always match the year covered by the Yearbook, and this can be confusing. For example, UN (1949) is the Demographic Yearbook for 1948 (a one year lag in publication), but UN (1951) is the Demographic Yearbook for 1951, UN (1961) is the Demographic Yearbook for 1961 and UN (1968) is the Demographic Yearbook for 1967.

but the exact date ranges from the late 1920s to the early 1940s; we use his data for either 1930 or 1940, depending on which is closer (and sometimes as part of an interpolation, as explained below). Preston’s most frequently used source is the U.N. Demographic Yearbook 1967 (UN 1968) but he also provides additional data from other country-specific work on life tables. We also note when Preston uses the same data as the U.N., as this is an important indication of external validation. In some cases, Preston uses other data when the U.N. estimates are available.

IVS is our abbreviation for International Vital Statistics, published by the Federal Security Agency (1947); life expectancy data are on pp. 220-225. We use these data to fill gaps where possible and more generally to provide a further external validation of pre-war sources.

We report alternative estimates from the UN (for different years), Preston, and IVS so readers can see the range of available data.

Coding Rules

Our decision rule for how to combine data from alternative sources for our main data series is as follows.

1. We look for male and female life expectancy data from the same time period and calculate an unweighted average (to match our post-1950 data that we take directly from the UN’s on-line database).⁵

2. Preston’s main source is the U.N. Demographic Yearbook for 1967 (UN 1968), but he supplements this with clearly identified and referenced country studies that he regards as comparable (or better than what the UN published). We also start with the UN Demographic Yearbook as our default source, as this is more comprehensive, but prefer Preston when he has data for a year close to 1930 or 1940 (as appropriate). We use IVS primarily as another source of validation and to fill a few gaps. If the U.N. Demographic Yearbooks have different data in different editions, we use the latest Yearbook (we note agreement or disagreement between alternative UN estimates below) through 1967—this appears to be the most recent year for which retrospective pre-1950 were systematically available.

⁵We have looked at the implications of using only male life expectancy. This does not significantly affect our results.

3. When we have data from before and after our reference dates (1930 and 1940), we interpolate by assuming a linear trend (which amounts to an “even-paced changed” with the same annual average change in life expectancy over the period for which we are interpolating). If data are available for a range of years, we take the midpoint for the purpose of calculating the trend. If the midpoint is halfway through a year (e.g., the midpoint of 1930-33 is halfway through 1931), we treat this as half a year. The formula is:

$$X_t = \left(\frac{kX_{t-s} + sX_{t+k}}{k + s} \right), \quad (\text{C1})$$

where X is a life expectancy, t is the reference year (1930 or 1940) and we have data from s years before the reference year and also from k years after the reference year. For example, if the life expectancy estimates are 50 for 1925 and 60 for 1935, we calculate life expectancy in 1930 as $(50+60)/10$.⁶

4. When the available data are only for shortly after our reference years and when we have two datapoints to calculate a reasonable rate of change, we extrapolate backwards. In our main sample, we only do one extrapolation for Portugal (for a period of 18 months). The extrapolation formula is:

$$X_t = \left(\frac{pX_{t+k} - kX_{t+p}}{k - p} \right) \quad (\text{C2})$$

where $p < k$.

5. For European countries affected by World War II in or before 1940, we prefer data from 1938 or 1937 where available; we interpolate these data to 1940. We rely on Urlanis (2003) for an assessment of the demographic impact from the war; where this was small, we prefer data for 1940.

6. If the data are drawn from a country that subsequently divided (e.g., India), we use these data for all the new countries that emerged.

7. In a few cases (El Salvador, Ecuador and Honduras), when data are not available for

⁶In robustness checks, we have constructed alternative data using log-interpolation. Instead of equation (C1), this uses

$$x_t = \left(\frac{kx_{t-s} + sx_{t+k}}{k + s} \right),$$

where x is log life expectancy. The results with log-interpolation are similar.

a country, but we have similar disease conditions, crude death rates and age structure of the population between neighboring countries, we average the life expectancy in neighboring countries with similar disease conditions. If an estimate is missing for a neighbor (typically for 1930), then we do not compute the average for that year but leave it as missing. Estimates relying on information from neighbors are dropped in additional robustness checks.

Life Expectancy by Country in our Base Sample

Most of the pre-1950 data we use are reported to two digits and we use them in this form in our calculations. However, since the post-1950 UN data is reported with one digit, we round all our data to one digit in the appendix table and in the regressions. All life expectancies numbers here are *at birth* unless otherwise stated. Life expectancy at age 20 was constructed from the same UN sources, using the same methods as explained above; where possible we used the same pages as for life expectancy at birth.

The panel base sample refers to noncommunist base sample, which contains 47 countries from 1940 and an additional 12 countries from 1950. “Long differences” refers to the sample that has data for 1940 and 1980; this comprises 47 countries. To include a country in our base sample we need data for 1940; to be in the 1930-1940 falsification tests we also need data for 1930; this sample is consequently smaller.

The following 12 countries are in our panel but not in our long differences, i.e., we have data from 1950 but not before: *Algeria, Bolivia, Egypt, Iran, Lebanon, Morocco, Tunisia, Turkey, Singapore, South Africa, Turkey, and Vietnam*. Life expectancy data for these countries are from the UN on-line database only.

The detail by country is as follows

1. *Argentina (in base sample for panel and long differences)*

UN (1967, p. 716 and 1961, p. 626) has life expectancy in 1914 of 45.2 for males and 47.5 for females, for an average of 46.35.⁷ For 1947, life expectancy for males is 56.9 and for females is 61.4, for an average of 59.15. Using our linear interpolation equation (C1), the improvement

⁷Argentina is not in UN (1949) and Preston (1975) does not have data. The IVS (p.220) reports life expectancy only for 1914; at birth this was 51.7 for the native population and 46.44 in the city of Buenos Aires (both estimates are for “both sexes”).

of 12.8 years over 33 years translates to 0.39 years of extra life for each year. This yields a life expectancy of 52.6 in 1930 and of 56.5 in 1940.

We use 52.6 for 1930 and 56.5 for 1940.

2. *Australia (in base sample for panel and long differences)*

United Nations (1949, p. 520 and 1951, p. 538) for 1932-34 has 63.48 (male) and 67.14 (female), average of 65.31; for 1946-48, it has 66.07 (male) and 70.63 (female), average of 68.35.⁸ United Nations (1968, p. 738) also reports life expectancy for 1920-22 as 59.15 (male) and 63.31 (female), for an average of 61.23.

Taking the mid-year in each of these ranges and assuming an even pace of life expectancy increase, from 1921 to 1933 (12 years) of 0.34 years per elapsed year implies life expectancy in 1930 (after 9 years) of 64.29. A similar calculation from 1933 to 1947 (14 years) gives 0.22 years more life expectancy per elapsed year; this implies life expectancy in 1940 (after 7 years) of 66.83.

We use 64.3 for 1930 and 66.8 for 1940.

3. *Austria (in base sample for panel and long differences)*

United Nations (1949, p. 516, and 1951, p. 532) for 1930-33 has 54.5 (male) and 58.5 (female), for an average of 56.50. IVS (p. 220) and Preston use the same numbers. For 1901-1905, UN 1949 has male life expectancy of 39.14 and female life expectancy of 41.06 (also in UN 1968, p. 722), average of 40.10; in the 28.5 year period from 1903 to half way through 1931 there was an average improvement of 0.58 years of life expectancy per elapsed year. This implies life expectancy in 1930 was 55.63 (calculating back from the midpoint of 1930-33).

There are no data for the late 1930s or 1940 in any of our sources. The next available number from the UN is for 1949-51, for which life expectancy is 61.91 (male) and 66.97 (female), for an average of 64.44. In the 18.5 years from half way through 1931 to 1950, there was a life expectancy improvement of 0.43 on average per elapsed year. This implies life expectancy in 1940 (after 8.5 years) was 60.16.

We use 55.6 for 1930 and 60.2 for 1940.

4. *Bangladesh (in base sample for panel and long differences)*

⁸IVS (p. 220) and Preston (1975) use the same estimates (for 1932-34), but IVS notes that this is “except aboriginals”.

We use the same estimates as for India. Standard errors in all regressions are clustered by unit of observation for original life expectancy data (i.e., pre-independence India) to account for this.

We use 26.8 for 1930 and 29.9 for 1940.

5. *Belgium (in base sample for panel and long differences)*

United Nations (1949, p. 516, and 1951, p. 532; 1968, p. 722) for 1928-32 has 56.02 (male) and 59.79 (female), for an average of 57.91. Preston and IVS (p. 220) use the same numbers. The next available estimates are for 1946-49, with life expectancy from UN 1968 (p. 722) given as 62.04 (male) and 67.26 (female), average of 64.65.

Over the 17.5 years from 1930 (midpoint of 1928-32) to half way through 1947 (midpoint of 1946-49), the average improvement in life expectancy was 0.39 years per elapsed year. This implies life expectancy in 1940 (after 10 years) was 61.76.

We use 57.9 for 1930 and 61.8 for 1940.

6. *Brazil (in base sample for panel and long differences)*

For this country, there is some discrepancy in the various UN sources. UN (1948 and 1951, p. 528) gives 37.43 as life expectancy for “both sexes” in 1920, while UN (1968, p. 716) has life expectancy for “both sexes” as 39.25 on average over the 1890-1920 period. UN (1961, p. 628 and 1968, p. 716) gives life expectancy for 1940-50, as 39.3 (male) and 45.5 (female), for an average of 42.40.⁹ Taking these latter data as representing 1945 (the midpoint) and taking the UN’s 1890-1920 as representing 1905 (preferring the later, presumably revised estimate from UN 1968), over the 40 year period there was an average improvement in life expectancy of $(42.4-39.25)/40$ or 0.08 years per elapsed year. This implies life expectancy in 1930 (after 25 years) of 41.22 and in 1940 (after 35 years) of 42.05.

Instead of the UN data for 1940, Preston (1975) prefers life expectancy estimates of 36.06 (male) and 37.25 (female), the average is 36.66; these data are from Arriaga (1968). Comparing with the earlier UN estimates suggests a worsening of life expectancy from 1905 to 1940 of 0.07 years per elapsed year. This would put life expectancy in 1930 at 37.40.¹⁰ We prefer the

⁹The table note says “for population born in Brazil” and “based on mortality rates implied from apparent survivorship rates between censuses.”

¹⁰IVS (p. 220) gives higher estimates, but this is presumably because it only reports data for two major cities. Male life expectancy in Rio de Janeiro 1939-41 is given as 40.77 and female life expectancy as 46.27. In Sao

Preston estimate for 1940 as he has a country specific source that he regards as comparable with but better than the available UN estimate. We also prefer the calculation for 1930 that uses Preston's 1940 estimate.

We use 37.4 for 1930 and 36.7 for 1940.

7. *Canada (in base sample for panel and long differences)*

United Nations (1949, p. 514) gives, for 1930-32, male life expectancy of 60.0 and female life expectancy of 62.1, with an average of 61.05; for 1940-42, it gives male life expectancy of 62.96 and female life expectancy of 66.30, for an average of 64.63. UN 1968, p. 710, has tiny adjustments to just the 1940-42 numbers: 62.95 male and 66.29 female, for an average of 64.62.¹¹

Preston (1975) uses 59.09 (male) and 61.58 (female) for 1931, for an average of 60.34; the source is Preston, Keyfitz, and Schoen 1972. We prefer the Preston estimate as he chooses a country specific study over the UN estimate.

The average annual change in life expectancy from 1931 to 1941 (midpoint of 1940-42) is 0.43 years. This implies that life expectancy in 1940 was 64.19, while in 1930 it was 59.91.

We use 59.9 for 1930 and 64.2 for 1940.

8. *Chile (in base sample for panel and long differences)*

UN (1949 and 1951, p. 528) reports life expectancy as 35.4 male and 37.7 female in 1930 (average 36.55), and as 37.9 male and 39.8 female in 1940 (average 38.85); these data were also used by IVS.¹² But there was a major revision of these data. UN (1961, p. 628, and 1968, p. 716) has life expectancy of 40.40 male and 41.03 female in 1930, average of 40.72, and 40.91 male and 43.16 female in 1940, average of 42.04.¹³ Preston uses these revised data for 1940.

We use 40.7 for 1930 and 42.0 for 1940.

Paolo for the same period male life expectancy is 46.71 and female is 51.77. IVS also gives alternative estimates of 43.04 for both sexes (not broken out separately) for Rio in 1939-40 and 50.13 for both sexes in Sao Paolo in 1939-40. For 1920-21, the "both sexes" life expectancy in Rio is given as 41; for Sao Paolo the equivalent estimate is 42.67. IVS also reports life expectation for "native population" in 1890-1920 as 39.25 (both sexes). Natives here likely means non-immigrants.

¹¹IVS (p. 220) has very similar estimates. Male life expectancy in 1941 ("except Yukon and Northwest Territories") was 62.95, while female life expectancy was 66.29. The same source gives male life expectancy as 58.46 in 1931 and female life expectancy as 60.23, for an average of 59.35. For males, life expectancy was 59.32 in 1929-31 and for females it was 61.59, for an average of 60.46.

¹²IVS also reports "both sexes" life expectancy as 41.2 in 1939 and 38.8 in 1929-32, suggesting that mortality fluctuated considerably from year-to-year at that time.

¹³The relevant footnote says "revised to take account of estimated underregistration of births."

9. *China (in base sample for panel only)*¹⁴

No life expectancy data for mainland China are available from the UN or our other sources. We therefore use estimates from Taiwan (Formosa) which was a province of China before 1950. Crude death rates indicate that this is a reasonable proxy.¹⁵

UN (1968, p. 718) has life expectancy for 1926-30 as 38.76 (male) and 43.13 (female), average of 40.95. For 1936-40, the same source gives 41.08 (male) and 45.73 (female), average of 43.41. See also United Nations (1949, p. 514, 1951, p. 530, and 1961, p. 628).

From 1928 (midpoint of 1926-30) to 1938 (midpoint 1936-40), the average annual increase in life expectancy was 0.25. This implies life expectancy in 1930 was 41.45, while in 1940 it was 43.91.

We use 41.5 for 1930 and 43.9 for 1940.

10. *Colombia (in base sample for panel and long differences)*

The earliest data in our UN sources (e.g., 1968, p. 716) are for 1950-52, with life expectancy of 44.18 (male) and 45.95 (female), average of 45.07.

Preston (1975) has estimates of life expectancy of 36.04 (male) and 37.19 (female) for 1938; average of 36.62 (his source is Arriaga).¹⁶ In the 13 years from 1938 to 1951 (midpoint of 1950-52) the average annual increase in life expectancy was 0.65. Interpolating implies life expectancy in 1940 was 37.92. We have not found any reasonable basis for calculating a 1930 estimate.

We use 37.9 for 1940 and have no data for 1930.

11. *Costa Rica (in base sample for panel and long differences)*

UN (1968, p. 710) reports life expectancy in 1949-51 of 54.65 for males and 57.05 for females, average of 55.85, but there is no earlier UN data and Preston does not have an estimate. However, IVS, p. 220, reports life expectancy for both sexes (not broken out separately) of 40.69 for 1927.

Using our interpolation formula, (C1), the improvement between 1927 and 1950 (midpoint

¹⁴The on-line UN database has no value for life expectancy in China in 1980. We interpolate between 1970 and 1990, to get a life expectancy of 65.31.

¹⁵The crude death rate on the mainland was 16 in 1950-54 and in Taiwan it was 10.

¹⁶There is a big difference with IVS, p. 220, which for 1939-41, reports male life expectancy at 46.3 and has no estimate for females. We prefer Preston because he uses a country-specific study and has data on females.

of 1949 and 1951) translates into an annual increase in life expectancy of 0.66 years, and implies that life expectancy in 1940 was 49.26, while in 1930 it was 42.67.

We use 42.7 for 1930 and 49.3 for 1940

12. *Denmark (in base sample for panel and long differences)*

UN (1951, p. 532 and 1968, p. 724) gives life expectancy for 1926-30 as 60.9 (male) and 62.6 (female), average of 61.75.¹⁷ This source also reports life expectancy for 1931-35 as 62.0 (male) and 63.8 (female), average of 62.9 (also in UN 1949, p. 516); Preston uses the 1931-35 data, but that is probably just because he was seeking data “in the 1930s”. We prefer the 1926-30 estimates as this allows us to interpolate (rather than extrapolate backwards).

UN (1968, p. 724) gives life expectancy for 1936-40 as 63.5 (male) and 65.8 (female), average of 64.65. The same source gives life expectancy for 1941-45 as 65.62 (male) and 67.7 (female), average of 66.66; also in UN 1949, p. 516.¹⁸ Ulanis indicates only a small demographic effect of the war in Denmark, but we prefer the 1936-40 data in any case as they include the year 1940.

Between 1928 (midpoint of 1926-30) and 1938 (midpoint of 1936-40), the average annual increase in life expectancy was 0.29 years. This implies life expectancy in 1930 of 62.33.

Between 1938 (midpoint of 1936-40) and 1943 (midpoint of 1941-45), the average annual increase in life expectancy was 0.4 years. This implies life expectancy in 1940 of 65.45.

We use 62.3 for 1930 and 65.5 for 1940

13. *Ecuador (in base sample for panel and long differences)*

We average of the estimates for Colombia (37.9 in 1940 and no data for 1930) and Peru (39.0 in 1930 and 40.6 in 1940).¹⁹

We use 39.3 for 1940 and have no estimate for 1930.

14. *El Salvador (in base sample for panel and long differences)*

We use an average of the estimates for Guatemala (30.4 for 1940 and no data for 1930) and

¹⁷UN (1949, p. 516) gives life expectancy for 1901-05 as 52.9 (male) and 56.2 (female); for 1911-15 as 56.2 (male) and 59.2 (female); and for 1921-25 as 60.3 (male) and 61.9 (female).

¹⁸There is a slight discrepancy with IVS (p. 220), which for Denmark “except Faroe Islands” for 1936-40, gives male life expectancy as 63.5 and female life expectancy as 65.8. The IVS estimates for 1931-35 match those of the UN.

¹⁹In 1935-39, the crude death rate in Ecuador was 24.6, while in Colombia it was 16.2 and in Peru it was 15.5. However, by 1950-54, Ecuador’s crude death rate was much closer to that of Colombia and Peru (within 3 per 1,000).

Nicaragua (34.5 in 1940 and no data for 1930).²⁰

We use 32.5 for 1940 and have no estimate for 1930.²¹

15. *Finland (in base sample for panel and long differences)*

United Nations (1949, p. 516, 1951, p. 532, and 1968, p. 726) for 1921-30 has 50.68 (male) and 55.14 (female), average of 52.91; for 1931-40 it has 54.45 (male) and 59.55 (female), average of 57.0; UN (1968, p. 728) for 1941-45 has 54.62 (male) and 61.14 (female), average of 57.88 (“excluding war losses”).²² The 1931-40 and 1941-45 data are also in UN (1961, p. 632); the note indicates these estimates exclude deaths of civilians and members of the armed forces due to military operations.²³ Preston (1975) uses 54.32 (male) and 59.48 (female) from 1936-40, which is from the UN (1968, p. 726); average of 56.9.

Between half way through 1925 (midpoint of 1921-30) and half way through 1935 (midpoint of 1931-40) there was an average annual increase in life expectancy of $(57-52.91)/10$, i.e., 0.41 years. This implies life expectancy in 1930 was $1.84+52.91=54.75$.

Between 1938 (midpoint of 1936-40) and 1943 (midpoint of 1941-45) there was an average annual increase in life expectancy of $(57.88-56.9)/5$, i.e., 0.20. This implies life expectancy in 1940 was 57.3. (Alternatively, if we extrapolate using the earlier rate of change in life expectancy, this would imply life expectancy in 1940 of 57.72).

We use 54.8 for 1930 and 57.3 for 1940.

16. *France (in base sample for panel and long differences)*

United Nations (1951, p. 534, 1968, p. 726) for 1928-33 has 54.30 (male) and 59.02 (female), average of 56.66. For 1933-38 the same source has 55.94 (male) and 61.64 (female), average of 58.79.²⁴ There are no data for exactly 1940 but for 1946-49 the UN (1968, p. 726) gives life expectancy as 61.86 (males) and 67.43 (females), average of 64.65.

The midpoint of 1928-33 is halfway through 1930 and so we use the 1928-33 data for

²⁰In 1935-39, the crude death rate in El Salvador was 21.1, while in Guatemala it was 26.8 and in Nicaragua it was 15.0.

²¹The earliest estimate in the UN sources is UN (1968, p. 710), which has life expectancy estimates for El Salvador in 1949-51 of 49.94 (male) and 52.40 (female). This implies a large, but plausible jump in life expectancy during the 1940s.

²²These data are also in IVS (p. 222).

²³UN (1949, p. 516) also has life expectancy for 1901-10 of 45.33 (male) and 48.1 (female), and for 1911-20 it has 43.41 (male) and 49.12 (female).

²⁴The IVS (p. 220) has similar data, with slight discrepancies at the second decimal place.

our 1930 estimate. Over the 12 years from halfway through 1935 (mid-point of 1933-38) to halfway through 1947 (midpoint of 1946-49) there was an average increase of 0.49 years of life expectancy per elapsed year. This implies life expectancy in 1940, assuming a linear trend, was 60.02.

Preston (1975) uses UN data of 55.12 (male) and 60.33 (female) from 1928-38; average of 57.73. We prefer the 1928-33, 1933-38 and 1946-49 UN data as they cover narrower windows that are closer to 1930 (and 1940).

We use 56.7 for 1930 and 60.0 for 1940.

17. *Germany (in base sample for panel and long differences)*

United Nations (1951, p. 534, and 1968, p. 726) for 1924-26 of 55.97 (male) and 58.82 (female), average of 57.40. The same source for 1932-34 has life expectancy of 59.86 (male) and 62.81 (female), average of 61.34; Preston and IVS use the same data. Interpolating between these two periods, over 8 years (1925 to 1933) there was an average annual increase of 0.49 years. This implies life expectancy in 1930 was 59.85.

The next available German data are for 1946-47, with male life expectancy at 57.72 and female life expectancy at 63.44 (UN, 1968, p. 728), but these data must large large effects of the recently ended war (Urlanis). There are also data for 1949-51, with life expectancy at 64.56 (male) and 68.48 (female), average of 66.52.²⁵ We therefore prefer to interpolate between 1932-34 and 1949-51 to obtain an estimate for 1940. Over the 17 year period from 1933 (midpoint of 1932-34) to 1950 (midpoint of 1949-51) there was an average annual life expectancy improvemen of 0.30.²⁶ This implies, with a linear trend, that life expectancy in 1940 was 63.47.

We use 59.9 for 1930 and 63.5 for 1940.

18. *Greece (in base sample for panel and long differences)*

UN (1968, p. 728) has life expectancy for 1926-30 of 44.95 (male) and 47.46 (female), average of 46.21.²⁷ For 1940 the same source puts life expectancy at 52.94 (male) and 55.80

²⁵The immediate post-war data are divided by the UN into Eastern Germany, Federal Republic of Germany, Berlin and West Berlin. Here we are using the numbers for the Federal Republic.

²⁶This is a lower annual increase in life expectancy than in France, for example. But given the differential effects of the war (see Urlanis), this does not seem unreasonable.

²⁷From IVS (p. 220) male life expectancy in 1928 was 49.09 and female life expectancy was 50.89.

(female), average of 54.37; Preston also uses these data.

In the 12 years between 1928 (midpoint of 1926-30) and 1940, there was an annual average increase in life expectancy of 0.68 years. This implies life expectancy in 1930 was 47.57.

We use 47.6 for 1930 and 54.4 for 1940.

19. *Guatemala (in base sample for panel and long differences)*

UN (1957, section following p. 564) reports life expectancy for Guatemala in 1921 as 28.22 years (male and female).

From United Nations (1949, p. 514, and 1951), covering 1939-41, we have life expectancy estimates of 35.97 for males and 37.09 for females; average of 36.53, but these are not for the whole country.²⁸ The UN (1968, p. 710) reports data for the whole of Guatemala only from 1949-51.

Preston (1975) uses life expectancy of 30.25 (male) and 30.46 (female) for 1940; average of 30.36 (from Arriaga). We prefer his estimates as they appear to cover the whole country, and also because Preston choose Arriaga's estimates over the UN numbers (which were available to him).

We use 30.4 for 1940 and do not have an estimate for 1930.

20. *Honduras (in base sample for panel and long differences)*

Based on crude death rates, we use the average of Guatemala (30.4 for 1940 and no data for 1930) and Nicaragua (34.5 in 1940 and no data for 1930).²⁹

We use 32.5 for 1940 and have no estimate for 1930.

21. *India (in base sample for panel and long differences)*

The United Nations (1949, p. 516, and 1951, p. 530) for 1921-31 gives male life expectancy of 26.91 and female life expectancy of 26.56, average of 26.74 (this is for "pre-partition India, including Burma"); these numbers are also used in IVS (p. 220), which says they are for "British India".³⁰

Preston (1975) uses 26.9 (male) and 26.6 (female) from 1931 data; average of 26.75. His

²⁸The note says "Department of Guatemala only".

²⁹In 1935-39, the crude death rate in Honduras was 16.2, while in Guatemala it was 26.8 and in Nicaragua it was 15.0.

³⁰The same 1949 source gives the following data for 1891-1901, 23.63 (male) and 23.96 (female); and for 1901-1911, 22.59 and 23.31.

source is Dandekar (1972) a country-specific source (although the data are obviously almost identical to those from the UN for 1921-31.) Between the 1921-31 estimate (midpoint 1926) and the 1931 estimate, there was an average annual increase in life expectancy of 0.002 years. Interpolating, this implies life expectancy in 1930 was 26.75 (to 2 significant figures).

UN (1961, p. 630 and 1968, pp. 718-720) has life expectancy for 1941-50 of 32.45 (male) and 31.66 (female), average of 32.06. Using this for the midpoint, halfway through 1945, the average annual improvement in life expectancy was 0.37 years over the 14.5 years after 1931. This implies life expectancy in 1940 was 30.05.

We use 26.8 for 1930 and 30.0 for 1940.

22. Indonesia (in base sample for panel and long differences)

There are no data in the UN Demographic Yearbooks or the IVS for Indonesia. Preston uses 32.5 (male) and 32.5 (female) for 1930-35; average of 32.5 (from a country-specific source, Nitisastro, 1970).

The UN on-line database gives 37.5 for average life expectancy in 1950-55. Using this for halfway through 1952 (the midpoint) implies that the average annual life expectancy increase in the 19.5 years after 1933 was 0.26. This implies life expectancy in 1940 was 34.29.

We use 34.3 for 1940 and have no estimate for 1930.

23. Ireland (in base sample for panel and long differences)

UN (1968, p. 730) has life expectancy in 1925-27 as 57.37 (male) and 57.93 (female), average of 57.65. The same source has life expectancy in 1935-37 of 58.20 (male) and 59.62 (female), average of 58.91; Preston uses these data.³¹ Taking the midpoints of these windows, between 1926 and 1936 there was an average annual increase in life expectancy of 0.13. This implies life expectancy in 1930 was 58.15.³²

UN (1961, p. 634, and 1968, p. 730) for 1940-42 has 59.01 (male) and 61.02 (female), for an average of 60.02. Between 1941 (midpoint of 1940-42) and 1936, there was an annual average increase in life expectancy of 0.22 years, implying that life expectancy in 1940 was

³¹From IVS (p. 222), life expectancy was 59.01 (male) and 61.02 (female) in 1940-42, 58.20 (male) and 59.62 (female) in 1935-37, and 57.37 (male) and 57.93 (female) in 1925-27.

³²United Nations (1951, p. 534) has somewhat different estimates. For 1921-30 this source has 56.2 (male) and 61 (female), average of 58.6; and for 1931-40 it has 60.9 (male) and 65.6 (female), average of 63.25. However, these estimates are not repeated in UN (1968), so we presume they were revised and we prefer the numbers in UN (1968).

59.80. There were no large demographic effects of the war on Ireland.

We use 58.2 for 1930 and 59.8 for 1940.

24. *Italy (in base sample for panel and long differences)*

United Nations (1951, p. 534 and 1968, p. 730) for 1921-22 has life expectancy of 49.27 (male) and 50.75 (female), average of 50.01. The same source for 1930-32 has 53.76 (male) and 56.0 (female), average of 54.88; and for 1935-37 it has female (only) life expectancy of 57.49.³³ After that, the next available data are for 1950-53, when life expectancy is given as 63.75 (male) and 67.23 (female), average of 65.49.

Preston (1975) uses the UN's data, but has an estimate of male life expectancy which is not UN (1968).³⁴ His estimates are 55.25 (male) and 57.49 (female) for 1935-37; average of 56.37. Using this for the midpoint of 1936 and assuming a linear trend to halfway through 1951 (midpoint of 1950-53) gives an annual average improvement in life expectancy over 15.5 years of 0.59. This implies life expectancy in 1940 was 58.72.

Between halfway through 1921 (midpoint of 1921-22) and 1931, there was an annual average increase in life expectancy of 0.51 years. This implies life expectancy in 1930 was 54.37.

We use 54.4 for 1930 and 58.7 for 1940.

25. *Korea (in base sample for panel and long differences)*

United Nations (1951, p. 530 and 1968, p. 720) for 1938 gives male life expectancy of 47.20 and female life expectancy of 50.59, for an average of 48.9. Korea is not in Preston or IVS.

The UN's on-line database has an estimate of life expectancy (average for male and female combined) of 47.5 for 1950-55. This estimate is presumably affected by the Korean war.³⁵

In the 14.5 years between 1938 and halfway through 1952, the average annual decrease in life expectancy was 0.1 years. This implies life expectancy in 1940 was 48.7.

We use 48.7 for 1940 and do not have an estimate for 1930.

26. *Malaysia (in base sample for panel and long differences)*

Based on crude death rates, we use the estimate of Thailand and adjust our standard errors

³³IVS (p. 222) also reports the 1930-32 data.

³⁴IVS, p. 222, also only has female life expectancy.

³⁵The next estimate, from UN (1968, p. 720) is for 1955-60 – 51.12 (male) and 53.73 (female), average of 52.43. But this is quite far into the treatment period, according to our reading of the history.

for clustering accordingly.³⁶

We use 42.6 for 1940 and do not have an estimate for 1930.

27. *Mexico (in base sample for panel and long differences)*

UN (1968, p. 712) has life expectancy in 1930 as 35.45 (male) and 37.08 (female), average of 36.27. For 1940, the same source reports life expectancy as 57.96 (male) and 40.42 (female). But the 1940 male datapoint must be a typo—for all other ages, the estimates are essentially unchanged from 1935 (male life expectancy of 38.94); probably the estimate for males should be 37.96, giving an average of 43.58. From IVS (p. 222) male life expectancy at birth was 37.19 in 1929-1933 and 32.44 in 1930.

Preston (1975) uses 38.94 (male) and 41.89 (female) for 1935; average of 40.42 (from the UN). We prefer the UN estimates for the exact dates of 1930 and 1940.

We use 36.3 for 1930 and 43.6 for 1940.

28. *Myanmar, previously known as Burma (in base sample for panel and long differences)*

The UN (1968, p. 718) reports life expectancy in Burma for 1921-31 as 30.61 (male) and 31.0 (female), average of 30.81. The same source gives life expectancy in 1954 as 40.8 (male) and 43.8 (female), average of 42.30.³⁷ Interpolating between 1926 (midpoint of 1921-31) to 1954, this yields an average annual increase of 0.41 additional years, and implies that life expectancy in 1930 was 32.45 and in 1940 was 36.56.

We use 32.5 for 1930 and 36.6 for 1940.

29. *Netherlands (in base sample for panel and long differences)*

United Nations (1951, p. 536, and 1968, p. 732) for 1921-30 has 61.9 (male) and 63.5 (female), average of 62.7; for 1931-40 it has 65.5 (male) and 67.2 (female), average of 65.6 (“excluding war losses”); IVS (p. 222) has very similar estimates and Preston uses exactly these numbers.³⁸ For 1947-49, UN (1968, p. 732) gives life expectancy as 69.40 (male) and 71.50 (female), average of 70.45.

From halfway through 1925 (midpoint of 1921-30) to halfway through 1935 (midpoint of

³⁶In West Malaysia (the Malay Peninsula), crude death rates were 20.8 in 1935-39; in Thailand during the same time period they were 16.4.

³⁷The UN’s on-line database gives life expectancy for 1950-55 as 36.9 (both sexes).

³⁸The same source gives male life expectancy (only) “excluding war losses” of 65.7 for 1931-40, indicating that the war losses adjustment is very small.

1931-40), there was an annual average increase in life expectancy of 0.29 years. This implies life expectancy in 1930 was 63.9.

From halfway through 1935 (midpoint of 1931-40) to 1948 (midpoint of 1947-49), there was an annual average increase in life expectancy of 0.39 years. This implies life expectancy in 1940 was 67.35.

We use 63.9 in 1930 and 67.4 in 1940.

30. *New Zealand (in base sample for panel and long differences)*

United Nations (1951, p. 538 and 1968, p. 740) has life expectancy for 1925-27 of 63.99 (male) and 66.57 (female), average of 65.28; and for 1931 of 65.04 (male) and 67.88 (female), average of 66.46. For 1934-38, the same source gives life expectancy as 65.46 (male) and 68.45 (female), average of 66.96 (IVS uses these numbers), while for 1950-52 life expectancy was 67.19 (male) and 71.29 (female), average of 69.24.³⁹

Between 1926 (midpoint 1925-27) and 1931, the average annual increase in life expectancy was 0.25 years. This implies life expectancy in 1930 was 66.21.

Between 1935 (midpoint of 1934-38) and 1951 (midpoint of 1950-52), the average annual increase in life expectancy was 0.14. This implies life expectancy in 1940 was 67.67.

We use 66.2 in 1930 and 67.7 in 1940.

31. *Nicaragua (in base sample for panel and long differences)*

Early data for this country are not in the UN Yearbooks or IVS. However, Preston (1975) has life expectancy of 33.88 (male) and 35.09 (female) for 1940 (from Arriaga), for an average of 34.49.

We use 34.5 in 1940 and have no estimate for 1930.

32. *Norway (in base sample for panel and long differences)*

United Nations (1951, p. 536 and 1968, p. 732) for 1921/22-1930/31 has 60.98 (male) and 63.84 (female), for an average of 62.41; IVS also uses these data. The same source gives life expectancy in 1931/32-1940/41 as 64.08 (male) and 67.55 (female), average of 65.82; Preston uses these data. In UN (1968, p. 732), life expectancy in 1945-48 was 67.76 (male) and 71.68 (female), average of 69.72.

³⁹Preston has for 1934-38, 65.04 (male) and 68.45 (female); but the male number seems to be an error (compared with UN 1951).

From 1926 (midpoint of 1921/22-1930/31) to 1936 (midpoint of 1931/32-1940/41), the average annual increase in life expectancy was 0.34. This implies life expectancy in 1930 was 63.77.

From 1936 (midpoint of 1931/32-1940/41) to halfway through 1946 (midpoint of 1945-48), the average annual increase in life expectancy was 0.37. This implies life expectancy in 1940 was 67.31.

We use 63.8 in 1930 and 67.3 in 1940.

33. *Pakistan (in base sample for panel and long differences)*

We use the same estimates as for India. Standard errors in all regressions are clustered by unit of observation for original life expectancy data (i.e., pre-independence India) to account for this.

We use 26.8 for 1930 and 30.0 for 1940.

34. *Panama (in base sample for panel and long differences)*

From United Nations (1949, p. 514, and 1968, p. 712), covering 1941-43, we have life expectancy of 50.54 for males and 53.46 for females, average of 52.0. Preston (1975), in contrast, has life expectancy of 41.50 (male) and 43.26 (female) for 1940, average of 42.38, from Arriaga. We prefer Preston's number as it is for exactly 1940 and from a country specific source that he preferred to the UN data (which was available to him).

For 1940 we use 42.4 and for 1930 we have no estimate.

35. *Paraguay (in base sample for panel and long differences)*

Based on crude death rates, we average the estimates of neighboring Argentina (52.6 for 1930 and 56.5 for 1940).and Brazil (37.4 for 1930 and 36.7 for 1940).⁴⁰

We use 45.0 for 1930 and 46.6 for 1940.

36. *Peru (in base sample for panel and long differences)*

There are no data in the UN Yearbooks or in Preston. However, IVS (p. 222), which gives "both sexes" life expectancy (for the city of Lima) as 38.97 in 1933-35. From the on-line UN data, life expectancy in 1950-55 was 43.9. Interpolating between 1934 (midpoint of 1933-35) to half way through 1952 (midpoint of 1950-55) implies an average annual increase of 0.27 years.

⁴⁰In 1950-54, the first years for which data are available, crude death rates in Paraguay were 11.2, while in Argentina they were 8.8 and in Brazil they were 20.6 (the latter is for 1940-54).

This implies life expectancy in 1940 was 40.61.

We use 40.6 in 1940 and do not have an estimate for 1930.

37. *Philippines (in base sample for panel and long differences)*

UN (1968, p722) has life expectancy of 25.17 (male) and 26.07 (female) in 1918, average of 25.62; it also has life expectancy of 44.80 (male) and 47.72 (female) for 1938; average of 46.26 (these data are also used by Preston). UN (1961, p. 630 and 1968, p. 722) has male life expectancy of 48.81 and female life expectancy of 53.36 in 1946-49, for an average of 51.09.

Interpolating from 1918 to 1938, there was an annual average increase of 1.02 years. This implies life expectancy in 1930 was 38.0.

From 1938 to halfway through 1947 (midpoint of 1946-49), there was an annual average increase in life expectancy of 0.51 years. This implies life expectancy in 1940 was 47.28.

We use 38.0 for 1930 and 47.3 for 1940.

38. *Portugal (in base sample for panel and long differences)*

United Nations (1951, p. 536, and 1968, p. 734) for 1939-42 has 48.58 (male) and 52.82 (female), average of 50.7; IVS uses these data ("including the Azores and Madeira islands"). For 1949-52, UN (1968, p. 734) has life expectancy of 55.52 (male) and 60.50 (female), average of 58.01.

Between halfway through 1940 (midpoint 1939-42) and halfway through 1950 (midpoint 1949-52), there was an average annual increase in life expectancy of 0.73 years. This implies life expectancy in 1940, extrapolating backwards from 1939-42 to 1940, was 50.33.

We use 50.3 for 1940 and do not have an estimate for 1930.

39. *Spain (in base sample for panel and long differences)*

United Nations (1968, p. 734) for 1930 has life expectancy of 48.38 (male) and 51.60 (female), average of 49.99; Preston uses very similar data from the UN for 1930-31. The same source gives life expectancy for 1940 as 47.12 (male) and 53.24 (female), average of 50.18 (presumably reflecting the effects of the civil war).

We use 50.0 for 1930 and 50.2 for 1940.

40. *Sri Lanka (in base sample for panel and long differences)*

UN (1968, p. 718) has life expectancy for 1920-22 of 32.72 (male) and 30.67 (female),

average of 31.70. The same source has life expectancy of 46.79 (male) and 44.72 (female) for 1945-47, average of 45.76.

Interpolating from 1921 to 1946, there was an average annual increase of 0.56 years of life expectancy. This implies life expectancy in 1930 was 36.76, while in 1940 it was 42.34.

We use 36.8 in 1930 and 42.3 in 1940.

41. *Sweden*

United Nations (1951, p. 536) has data for: 1921-30, 60.97 (male) and 63.16 (female), average of 62.07; for 1931-40, 63.76 (male) and 66.13 (female), average of 64.95; for 1941-45, 67.06 (male) and 69.71 (female), average of 68.39. The UN (1968, p. 734) gives life expectancy as 64.30 (male) and 66.92 (female) in 1936-40, average of 65.61; these data are also used by the IVS. Sweden did not have significant war losses (Uralanis).

Preston (1975) uses 62.02 (male) and 64.11 (female) for the narrower window of 1928-32; average of 63.07 (from Keyfitz and Flieger). We use these data for 1930 as that is the exact midpoint of his window and because Preston preferred these numbers to estimates from the UN.

Between 1938 (midpoint of 1936-40) and 1943 (midpoint of 1941-45), there was an average annual increase in life expectancy of 0.56 years. This implies life expectancy in 1940 was 66.72.

We use 63.0 in 1930 and 66.7 in 1940.

42. *Switzerland*

UN (1968, p. 736) has life expectancy for 1920-21 as 54.48 (male) and 57.70 (female), average of 56.09. United Nations (1951, p. 536) has life expectancy for 1929-32 as 59.17 (male) and 63.05 (female), average of 61.11 (Preston uses this estimate); for 1933-37 as 60.7 (male) and 64.6 (female), average of 62.65; for 1939-44 as 62.68 (male) and 66.96 (female), average of 64.82. Switzerland had no significant war losses (Uralanis). IVS also uses these data.

Between halfway through 1920 (midpoint of 1920-21) and halfway through 1930 (midpoint 1929-32), the average annual increase in life expectancy was 0.50 years. This implies life expectancy in 1930 was 60.86.

Between 1935 (midpoint 1933-37) and half way through 1942 (midpoint 1939-44), the average annual increase in life expectancy was 0.29. This implies life expectancy in 1940 was

64.10.

We use 60.9 for 1930 and 64.1 for 1940.

43. *Thailand*

Preston and the IVS have no data. United Nations (1949, p. 516 and 1951, p. 530) gives life expectancy at birth for Siam, 1937-38, as 36.73 for male and 43.3 for female, average of 40.02. These data are only "For Bangkok municipal area."⁴¹

The next available data are for 1947-48—UN (1968, p. 722) gives life expectancy for that period as 48.69 (male) and 51.90 (female), average of 50.30. In the 10 years between halfway through 1937 (midpoint of 1937-38) and halfway through 1947 (midpoint of 1947-48), the average annual increase in life expectancy was 1.03 years. This implies life expectancy in 1940 was 42.59.

We use 42.6 for 1940 and do not have an estimate for 1930.

44. *United Kingdom (England and Wales)*⁴²

UN (1968, p. 736) has life expectancy for 1920-22 of 55.62 (male) and 59.58 (female), average of 57.60. United Nations (1951, p. 536, and 1968, p. 736) has life expectancy for 1930-32 as 58.74 (male) and 62.88 (female), average of 60.81 (Preston uses these data); and for 1948, 66.39 (male) and 71.15 (female), average of 68.77.

Between 1921 (midpoint of 1920-22) and 1931 (midpoint of 1930-32), the average annual increase in life expectancy was 0.32 years. This implies life expectancy in 1930 was 60.49.

Interpolating between 1931 (midpoint of 1930-32) and 1948, there was an average annual increase in life expectancy of 0.47 years. This implies life expectancy in 1940 was 65.02.

We use 60.5 for 1930 and 65.0 for 1940.

45. *USA*

United Nations (1951, p. 528, and 1968, p. 714) gives 57.71 (male) and 60.99 (female), average 59.35, for 1929-31 and 61.60 (male) and 65.89 (female) for 1939-41, average 63.75, (the 1939-41 data are also in UN 1949, p. 514).⁴³

⁴¹UN (1961, p. 632) has life expectancy for 1947-48 of 48.69 (male) and 51.9 (female), for an average of 50.3.

⁴²We do not use the data for Scotland or Northern Ireland.

⁴³UN (1949, p. 514) has male life expectancy in 1900-1902 as 47.88 and female life expectancy as 50.70. For 1909-1911, it has 49.86 and 53.24. But these data are only "for the ten death-registration States of 1900." IVS reports life expectancy broken down by "white," "negro," and "nonwhite."

Preston uses 59.2 (male) and 62.8 (female), average of 61.00, for 1929-38 (from US official statistics). However, his window is so broad that we prefer the UN data, for which the midpoints are the exact dates of interest.

We use 59.4 for 1930 and 63.8 for 1940.

46. *Uruguay*

There are no life expectancy estimates in our sources. However, from the UN (1967) we know that crude death rates were quite similar to those in a neighbor, Argentina.⁴⁴ We therefore use the life expectancy estimates of Argentina (52.6 for 1930 and 56.5 for 1940).

We use 52.6 for 1930 and 56.5 for 1940.

47. *Venezuela*

Preston reports life expectancy of 33.29 (male) and 34.47 (female) for 1936; average of 33.88. In the UN on-line database, for 1950-55, life expectancy is 55.1 (average of both sexes).

Between 1936 and halfway through 1952, the average annual increase in life expectancy was 1.29 years. This implies life expectancy in 1940 was 39.02.

We use 33.9 for 1940 and have no estimate for 1930.

Additional Life Expectancy Data

Our sources provide life expectancy data for some additional countries, but we are not able to include them in our sample because we are currently missing other data (typically GDP, but sometimes population). For example, Preston (1975) has data for the Dominican Republic, Iceland and Japan.⁴⁵ IVS has data on Estontia and Latvia. The UN Demographic Yearbook has data on British Guiana, Cyprus, Jamaica, Iceland and Puerto Rico.

African Data

There are only limited data on life expectancy in Africa south of the Sahara before 1950. For Angola, UN (1961, p. 622), reports “both sexes” life expectancy of 35 in 1940. For Mozambique in 1940, the same source gives “both sexes” life expectancy of 45 in 1940 (specifically, this is

⁴⁴In 1935-39, crude death rates were 11.1 in Uruguay and 11.5 in Argentina.

⁴⁵Japan’s data are problematic because the country was at war for most of the 1930s, and this had a potentially significant effect on life expectancy.

“African population ... living according to tribal customs.”) The UN (1961, p. 622) reports both sexes life expectancy as 38 in 1948 for Ghana. For Mauritius, 1942-46, this source gives male life expectancy as 32.25 and female life expectancy as 33.83.

Kuczynski (1948) contains a great deal of qualitative description of the disease burden, particularly for West Africa. His more limited information for East Africa suggests the disease burden may have been less, but his evidence is not conclusive. Overall, in terms of the mix of diseases, his evidence suggests that Africa is very similar to India around 1940, and we assume the same distribution of diseases in Africa as in India to calculate predicted mortality.⁴⁶

Using our global mortality instrument does not require any such assumption, as in this case we take predicted mortality from a global average (however, we cannot add Africa to the calculation of that average due to lack of data.) For the sample including Africa, this instrument should therefore be preferred.

Causes of Death

Detailed breakdowns of causes of death, by country, are used as follows for our base sample. IVS here refer to Table 20 of International Vital Statistics (beginning on p.174). In some cases this source indicates when the sample is not national and we repeat that indication here. League of Nations refers to the data republished by World Health Organization (1951); this reports data for 1939-46 and we use the information for 1940 or nearest available year.⁴⁷

1. Argentina, IVS for 1936
2. Australia, IVS for 1940 (“excluding aboriginals”)
3. Austria, IVS for 1938
4. Bangladesh, same as for India
5. Belgium, IVS for 1940

⁴⁶Controversially, this implies that the burden of malaria was the same as in India. It also implies that African malaria was not of a qualitatively different nature (e.g., there are some who argue it was impossible to eradicate African malaria using the technologies of the 1940s/1950s). However, as we are missing population and GDP data for Africa before 1950, this assumption does not much affect our estimates.

⁴⁷The League of Nations reports death by disease in cities (with the number of cities varying by country); we construct an unweighted average of death rates by disease across all available cities (for some countries the city coverage varies by disease). City level data are also available for places where we have country estimates of death by disease, and we have checked that the two sets of estimates are similar. The main text reports robustness results if we drop all data from this source).

6. Brazil, IVS for 1940 (“21 cities”)
7. Canada, IVS for 1940 (“excluding Yukon and N.W.T. [North West Territories]”)
8. Chile, IVS for 1940
9. China, League of Nations
10. Colombia, IVS for 1940
11. Costa Rica, IVS for 1940
12. Denmark, IVS for 1940
13. Ecuador, League of Nations
14. El Salvador, IVS for 1940
15. Finland, IVS for 1940
16. France, IVS for 1940
17. Germany, IVS for 1938
18. Greece, IVS for 1938
19. Guatemala, IVS for 1943
20. Honduras, League of Nations
21. India, League of Nations
22. Indonesia, League of Nations
23. Ireland, IVS for 1940
24. Italy, IVS for 1940
25. Korea, League of Nations
26. Malaysia, League of Nations
27. Mexico, IVS for 1940
28. Myanmar (Burma),
29. Netherlands, IVS for 1940
30. New Zealand, IVS for 1940 (“excluding Maoris”)
31. Nicaragua, League of Nations
32. Norway, IVS for 1940
33. Pakistan, same as India
34. Panama, League of Nations

35. Paraguay, League of Nations
36. Peru, IVS for 1943 (“excluding jungle population”)
37. Philippines, League of Nations
38. Portugal, League of Nations
39. Spain, IVS for 1940
40. Sri Lanka, League of Nations
41. Sweden, IVS for 1940
42. Switerland, IVS for 1940
43. Thailand, League of Nations
44. United Kingdom (England and Wales), IVS for 1940
45. USA, IVS for 1940, League of Nations
46. Uruguay, IVS for 1940, League of Nations
47. Venezuela, IVS for 1940 (“excluding tribal Indians”)

For other death by disease data (in our extended sample), we use IVS for Egypt in 1940 (“Health Bureau Areas”) and, where relevant, for South Africa, IVS for 1939 (“Europeans”). For all other countries, we use the League of Nations.⁴⁸

⁴⁸We do not use data from IVS for the following countries (because other data issues preclude us from including these countries in our samples): Iceland, IVS for 1940; Japan, IVS for 1940; Lithuania, IVS for 1939; Northern Ireland, IVS for 1940; and Scotland, IVS for 1940.

Additional References Used in Appendix C (for other references see main text)

Arriaga, Eduardo E. (1968) *New Life Tables for Latin American Populations in the Nineteenth and Twentieth Centuries*, Population Monograph #3, Institute of International Studies, University of California, Berkeley.

Dandekar, K. (1972) “Mortality and Longevity in India, 1901-61,” in K.E.Vaidyanathan (ed.) *Studies on Mortality in India*, Grandhirgram Institute of Rural Health.

Nitisastro, Widjojo (1970) *Population Trends in Indonesia*, Cornell University Press.

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

Preston, Samuel H., Nathan Keyfitz, and Robert Schoen (1972) *Causes of Death: Life Tables for National Population*, Seminar Press, New York.