Technology adoption, mortality and population dynamics

Technology Adoption, Mortality and Population Dynamics[*]

John Hejkal[†]   B. Ravikumar[‡]   G. Vandenbroucke[§]

May 2021

Abstract

We develop a quantitative theory of mortality trends and population dynamics. Our theory emphasizes individual choices on costly adoption of healthy technologies and diffusion of knowledge about infections as a key channel for reducing mortality. Our theory is consistent with three observations on mortality: (i) The cross-country correlation between levels of mortality and income is negative; (ii) mortality in poor countries has converged to that of rich countries despite no convergence in income; and (iii) economic growth is not a prerequisite for mortality to decline. We calibrate our model to the time series of crude death rates in Sweden. We then simulate the time series of crude death rates for 87 countries from 1960 to 2018. Our model accounts for the static negative correlation, 99% of the convergence of mortality; and, as in the data, countries with negative growth do experience decreasing mortality, and no country with increasing mortality experiences negative growth. The model reproduces the change in population and its distribution across countries. For instance, total population increased by 3.1 billion and the model accounts for 97% of the increase and the fact that almost one half of this increase is due to poor countries.

JEL codes: I12, I15, J11, E13
Keywords: Mortality, population dynamics, technology adoption, diffusion.

---

1 INTRODUCTION

The search for an economic theory of population is as old as economics itself: Both Smith (1776) and Malthus (1798) theorized the link between standards of living and population growth. The interest in demographic questions was later revived by Becker (1960) via an *economic* theory of fertility. Becker's approach was to introduce individual decisions into the analysis of fertility and, thereby, explain the systematic correlation between economic variables and fertility.[1] Following Becker's work a large literature investigated the economic determinants and implications of fertility decisions. Yet, population dynamics cannot be explained by changes in fertility alone: The world's population *growth* rate has increased for at least 200 years *despite* a declining birth rate (see Figure 1). It took 82 years for the world population to increase from 1 billion to 1.5 billion, an average annual growth rate of less than 0.5 percent; the crude birth rate (CBR) during this period was above 40, implying an average 5.7 births per woman. In contrast, it took only 8 years for the world population to increase from 3 billion to 3.5 billion; the population growth rate quadrupled to almost 2 percent per year, while the CBR in this period was 35 (5 births per woman). Thus, for an economic theory of population we need a theory of mortality.

Similar to Becker's economic approach, our theory of mortality is based on individual decisions. Our objective is to explain the observed patterns on mortality and economic outcomes quantitatively, both in cross-section and time series. Our theory emphasizes infections as causes of death and diffusion of knowledge about avoiding the infections as a channel for reducing mortality. An essential feature of our theory is costly adoption of "best practices," which we label technologies. Our model has two technologies—obsolete and modern—that are characterized by age-specific survival probabilities, with the modern technology offering a higher chance of survival at all ages. Individuals with the obsolete technology have to spend resources—time and/or goods—in order to increase their probability of acquiring the modern technology. We further assume that adoption confers a dynamic externality: As more individuals adopt the modern technology, it becomes less expensive to acquire. The exogenous forces that change over time in our model are total factor productivity (TFP) and crude birth rates.

Our theory accounts for a well-known fact: negative correlation between levels of real gross domestic product per capita (GDP) and mortality. The cross-country elasticity of crude death rate (CDR) with respect to GDP in 1960 is around -0.25: i.e., if a country had 1 percent higher GDP than another in 1960, the former's CDR was 0.25 percent lower than the latter's. In our model richer countries have higher TFP relative to poorer countries and they spend more resources to acquire the modern technology, which results in a lower death rate.

As Davis (1956) noted, if the only cause of death was a lack of economic means, as suggested by both

---

[1] Becker (1960, p. 209) wrote:

> "Two considerations encouraged me to analyze family size decisions within an economic framework. The first is that Malthus's famous discussion was built upon a strongly economic framework (...) Second, although no single variable (...) explained more than a small fraction of the variation in fertility, economic variables did better than others."
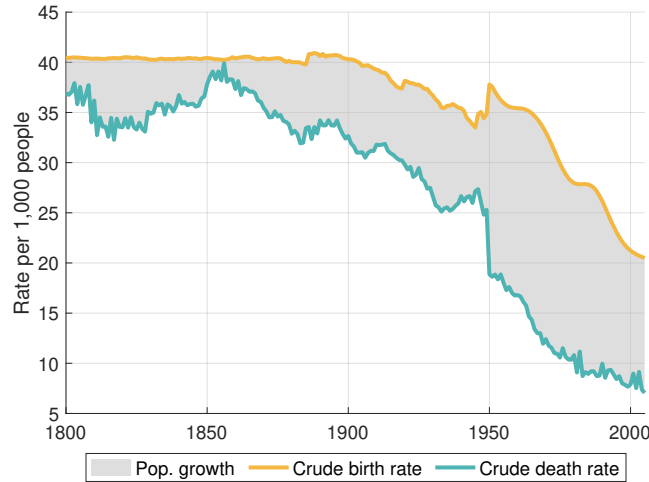
Figure 1: The growth rate of the world's population and its components

Note: The growth rate of the world's population is the difference between the crude birth rate and the crude death rate.
Source: https://gapminder.org/.

Smith (1776) and Malthus (1798), it would be trivial to explain the negative correlation between the levels of GDP and mortality. It would not be possible, however, to explain two less-known facts: the convergence of mortality rates and decline in mortality without an economic takeoff.

Mortality in poor countries has converged over time to that in rich countries despite no convergence in income. The GDP elasticity of CDR is around zero in the current decade. Benin, for example, has the same CDR as that of the U.S. in 2020 but only 1% of its GDP. Panel A of Figure 2 shows GDP in rich countries and in poor sub-Saharan African countries since 1960. A well-known pattern emerges—the absence of income convergence between the two groups. Poor sub-Saharan African countries are becoming poorer than rich countries, from 9% of the GDP of rich countries in 1960 to about 4% in 2018. Despite the divergence in income, there is convergence in CDR, as illustrated in Panel B: The mortality rate was 2.4 times higher in sub-Saharan African countries than in rich countries in 1960 but becomes essentially the same by 2018. Overall in our sample of 87 countries, there is no convergence of GDP between the top and the bottom decile, but there is convergence in CDR. GDP in the bottom decile was 1.8% of that in the top decile in 1960, and 3.5% in 2018; but CDR in the top decile was 2.4 times that of the bottom decile in 1960 and 15% lower in 2018.[2]

On economic growth and mortality, increases in GDP are not a necessary cause of mortality decline; mortality can decline without an economic takeoff. Figure 3 shows a scatter plot of the (annualized) growth rate of GDP versus growth rate of CDR between 1960 and 2018 for all countries in our sample. The message from this figure is that higher economic growth does not correlate with more rapid

---

[2]Davis (1956) noted the convergence of mortality, and Acemoglu and Johnson (2007) noted that since the 1930s life expectancy of poor countries converged to that of rich countries even though GDP did not.
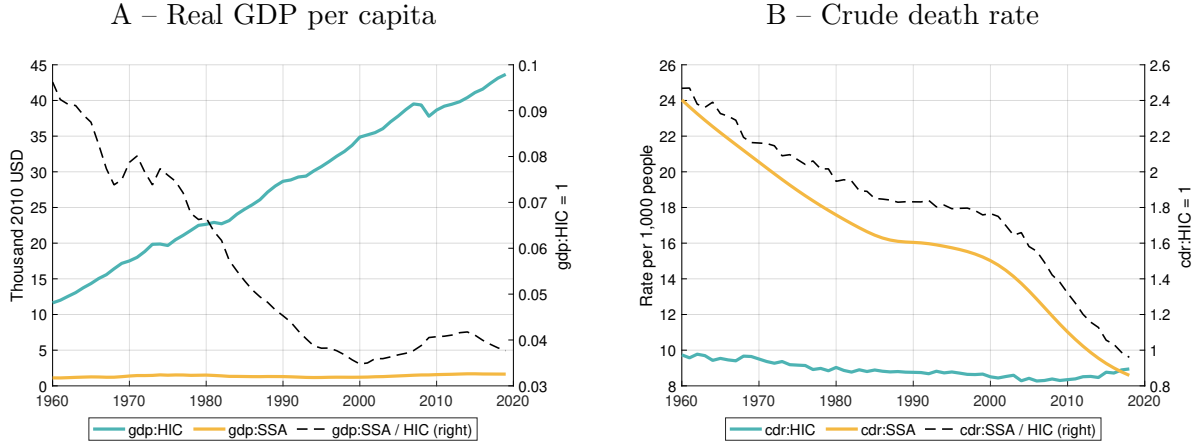
A – Real GDP per capita

B – Crude death rate

Figure 2: Real GDP per capita and mortality in high-income countries (HIC) v. sub-Saharan Africa (SSA)

Note: Panel A shows real GDP per capita in high-income countries (HIC) and sub-Saharan African (SSA) countries (excluding high income) with solid lines. The dashed line indicates the SSA-to-HIC ratio of real GDP per capita. Panel B shows the crude death rate (cdr) with solid lines in the two regions, and the SSA-to-HIC ratio of crude death rates (dashed line). sub-Saharan Africa comprises Angola, Ethiopia, Niger, Benin, Gabon, Nigeria, Botswana, Gambia, Rwanda, Burkina Faso, Ghana, Sao Tome and Principe, Burundi, Guinea, Senegal, Cabo Verde, Guinea-Bissau, the Seychelles, Cameroon, Kenya, Sierra Leone, the Central African Republic, Lesotho, Somalia, Chad, Liberia, South Africa, Comoros, Madagascar, South Sudan, the Congo (Dem. Rep.), Malawi, Sudan, the Congo, Mali, Tanzania, the Ivory Coast, Mauritania, Togo, Equatorial Guinea, Mauritius, Uganda, Eritrea, Mozambique, Zambia, Eswatini, Namibia, Zimbabwe. High-income countries are those with a gross national income per capita of $12,376 or more in 2018.
Source: World Development Indicators.

declines in mortality. There are countries with negative growth during this period but none of them experienced increasing mortality. Madagascar's GDP, for example, decreased at an annual rate of 0.87% between 1960 and 2018 while its CDR decreased at the rate of 2.39% per year. Conversely, there are countries with increasing mortality but none of them experienced negative growth.

Historians such as Fogel (2004) and Livi-Bacci (1991) noted this disconnect between economic takeoff and mortality decline, as some countries experienced a transition toward low mortality prior to their economic takeoff. As a case in point Figure 4 shows demographic and economic data for England from the 16th to the 19th centuries. Panel A indicates that the crude death rate declined throughout the 18th century. Panel B, however, indicates that economic take-off did not occur before the early part of the 19th century. Specifically, TFP in both the agriculture and manufacturing sectors are flat throughout the 18th century, the production of food per capita is not increasing, and wages increase only at the onset of the 19th century.

While countries with higher GDP tend to have lower CDR at a point in time, an increase in a country's GDP *over time* is not a prerequisite for a decrease in its CDR. The country-specific dynamics of CDR in our model are determined by the diffusion of the modern technology and the speed of diffusion depends upon TFP and the dynamic externality.
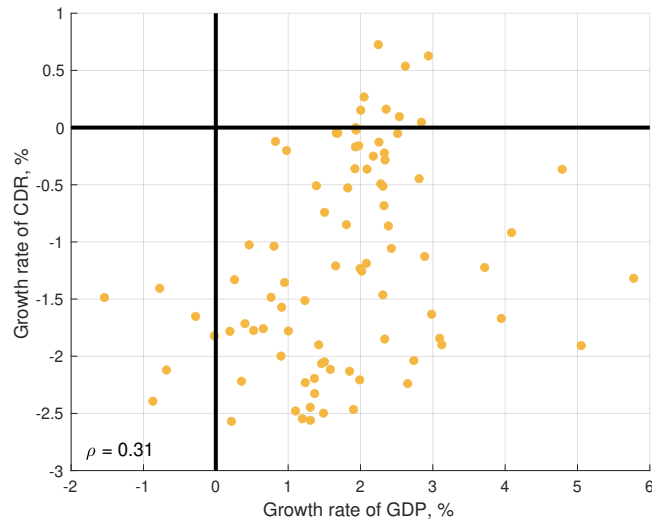
4

Figure 3: Cross-country growth rates of real GDP per capita and crude death rate, 1960-2018
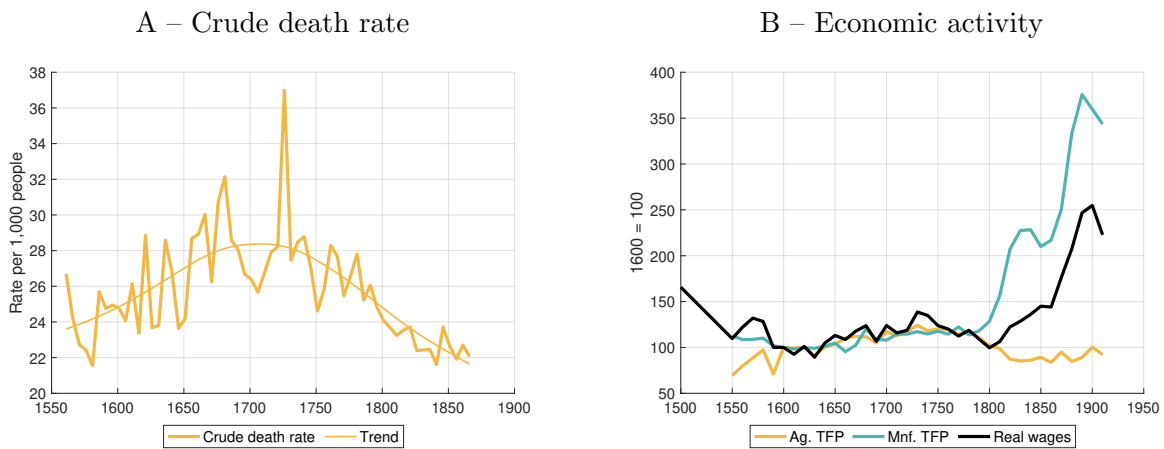
Source: World Development Indicators.

Figure 4: Mortality and the economy in England from the 16th to the 19th centuries

Source: Wrigley et al. (1997, Tables 6.3 & A9.1), Clark (2002, Tables 1, 3 & 5) and Leukhina and Turnovsky (2016).

The two-technologies feature keeps our model tractable: At any point in time, cross-sectional heterogeneity is limited to age and whether individuals have the obsolete or modern technology (obsolete and modern individuals, for short). As TFP increases, more resources—time and/or goods depending on their relative costs—are spent trying to adopt the modern technology. As more individuals adopt the modern technology, the marginal benefit of spending resources to acquire it increases as well because of the dynamic externality. Starting with an exogenous proportion of modern individuals, the evolution of this proportion is *endogenous*, and its speed dictated by the dynamic externality. The evolution follows an S-curve as in models of diffusion, and the mortality rate declines as the proportion of modern individuals increases.

The crucial parameters in our model are those of the function that maps proportion of individuals with the modern technology and expenditures on time and goods into the probability of acquiring the modern technology. We calibrate the model to Swedish time series data from 1751 to 2017. Specifically, we find parameters such that the model fits the time series of the Swedish CDR, given the time series of its crude birth rate and a time series for TFP that matches GDP growth in Sweden during this period. By construction, the calibrated model is consistent with the age distribution of the Swedish population throughout the period.

We simulate a time series of crude death rates from 1960 to 2018 for each country in our sample. We use their actual CBR and GDP as exogenous variables. The model is consistent with the well-known negative correlation between CDR and GDP across countries. It is also consistent with the fact that CDR in poor countries converges to that of rich countries. Quantitatively, the model accounts for almost all of the convergence: The model-data correlation of the poor-to-rich CDR ratio is 99%. The model also implies a positive correlation between changes in CDR and changes in GDP, as in Figure 3. In the model, countries with zero or negative GDP growth do experience declining mortality, and no country with increasing mortality experiences negative growth of GDP. As in the data, economic growth is not necessary for a decline in mortality in our model.

Finally, our model of mortality explains observed patterns on population dynamics quantitatively. Total population in our sample increases by 3.1 billion over roughly 60 years. The model accounts for 97% of the change. The increase in population in poor countries is almost one half of the increase in total population, while the rich countries' increase in population is just 5% of the total increase. The model delivers on the cross-country distribution of change in population as well.

These increases in population are delivered by the model via mortality that is consistent with the observed mortality dynamics. For instance, CDR declined by 14.5 deaths per thousand people in poor countries; in the model, the corresponding decline is 12.

As noted, almost one half of the increase in total population is due to poor countries. This increase occurred despite the CBR decline in poor countries from 43 births per thousand people to 20 over a period of almost sixty years. To account for the increase in the population of poor countries (almost 1.5 billion), the decline in their mortality is a crucial quantitative component.

2 MODEL

Time is discrete and indexed by t. The economy is populated by overlapping generations of individuals living up to a maximum age J. The preferences of an individual of generation t (i.e., an individual of age 0 at date t) and age a are represented by

$$E_{t,a} \left[\sum_{j=a}^{J} \beta^{j-a} U\left(c_{t,j}\right) \right],$$

where $\beta \in (0,1)$ is the subjective discount factor, U is a utility index, $c_{t,j}$ represents consumption at age j, and $E_{t,a}$ is the expectation operator conditional on the information available to the individual at age a. Mortality is the only source of uncertainty, so the expectation is calculated using the appropriate age-specific survival probabilities.

There are two health-relevant technologies available at any point in time: "obsolete" and "modern." These technologies are characterized by sequences of age-specific survival probabilities, $\left\{ s_j^i \right\}_{j=0}^{J}$, where s_j^i denotes the probability of survival of individual i from age j to $j+1$, conditional on being alive at age j. We will label individuals with the modern technology as "modern individuals" ($i=m$); similarly, individuals with the obsolete technology are "obsolete individuals" ($i=o$). Since J is the maximum length of life, we assume $s_J^o = s_J^m = 0$.

Individuals become "economically active" at age $k > 0$. From age k to J they are endowed with one unit of time each period. The wage rate is denoted by w_t in period t. There is no saving. Having $k > 0$ is necessary for tracking the entire age distribution, which in turn affects mortality and population dynamics.

Modern individuals allocate all their time to working, their choice problem is trivial. The value of a modern individual of generation t and age $j \geq k$ is

$$V_{t,j}^m = U(w_{t+j}) + \beta s_j^m V_{t,j+1}^m. \tag{1}$$

An obsolete individual can adopt the modern technology with some probability and become a modern individual. The probability of adoption for an age $j > k$ member of generation t is

$$Q(y_{t,j}, h_{t,j}, \pi_{t+j}), \tag{2}$$

where $y_{t,j}$ represents goods resources, $h_{t,j}$ represents time, and π_{t+j} is the proportion of economically-active modern individuals at the beginning of period $t+j$. The value of an obsolete individual is

$$V_{t,j}^o = \max_{y,h} U\left(w_{t+j}(1-h) - y\right) + \beta Q(y_{t,j}, h_{t,j}, \pi_{t+j}) s_j^m V_{t,j+1}^m$$
$$+ \beta \left(1 - Q(y_{t,j}, h_{t,j}, \pi_{t+j})\right) s_j^o V_{t,j+1}^o. \tag{3}$$

Since the probability of survival is 0 at age J, the terminal condition for modern and obsolete individuals is

$$V_{t,J}^m = V_{t,J}^o = U(w_{t+J}).$$

At an interior, the optimal choice of an obsolete individual is characterized by the first-order conditions

$$y_{t,j} \quad : \quad 0 = U_1\left(w_{t+j}\left(1-h\right)-y\right) - \beta Q_1(y_{t,j}, h_{t,j}, \pi_{t+j})\Delta_{t,j}, \tag{4}$$

$$h_{t,j} \quad : \quad 0 = U_1\left(w_{t+j}\left(1-h\right)-y\right)w_{t+j} - \beta Q_2(y_{t,j}, h_{t,j}, \pi_{t+j})\Delta_{t,j}, \tag{5}$$

where $\Delta_{t,j} \equiv s_j^m V_{t,j+1}^m - s_j^o V_{t,j+1}^o$. (We allow for corner solution in our quantitative experiments.)

We use the term "TFP" channel to refer to the effect of w_t on the solution to problem (3) and the term "dynamic externality" to refer to the effect of π_t. TFP is exogenous in our model whereas π_t evolves endogenously but is taken as given by individuals.

We assume that resources, $y_{t,j}$ and $h_{t,j}$, are expended only once for adoption. Modern individuals do not have to expend resources every period to "operate" the modern technology. None of our results depend upon this assumption.

A few remarks are in order at this stage. We interpret death in our model as the result of infections that can be avoided by adopting the modern technology. Appendix A provides examples of deadly infections and steps individuals can take to avoid contracting them: e.g. boiling water, washing hands, wearing a mask. In the same spirit, Oral Rehydration Therapy (ORT) is a cheap procedure that can be performed at home. ORT does not prevent one from becoming infected with, say cholera, but effectively reduces the death rate from such infections.

Second, the dynamic externality affects the speed of knowledge diffusion and hence the speed of mortality decline in our model. As more individuals acquire the modern technology, the easier the best practices to reduce mortality from infections are to come by.

Third, our assumption that resources affect the *probability* of acquiring the modern technology is a modeling convenience. The alternative assumption, i.e. that adoption is certain once resources are expended, would imply that all individuals in an age group adopt simultaneously. Age-specific survival probabilities implied by the model would then be step functions of time: low until the age group in question adopts, then high. This would be at odds with data. One could introduce within-age heterogeneity to avoid this issue. We decided, instead, to model adoption as random in order to keep the model simple. Our modeling can be interpreted as allowing people to choose over lotteries, à la Prescott and Townsend (1984) and Rogerson (1988).

2.1 Population dynamics

Let $p_{t,j}^i$ denote the population of obsolete ($i = o$) and modern ($i = m$) individuals of age j in generation t. The population of obsolete individuals at t is then $\sum_{j=0}^{J} p_{t-j,j}^o$ and, similarly, the population of

modern individuals at t is $\sum_{j=0}^{J} p_{t-j,j}^m$.

Let CBR_t denote the crude birth rate, number of births per person, at time t. We make the following simplifying assumptions: (i) Fertility rates are not age specific; (ii) CBR is the same for modern and obsolete individuals; and (iii) children of obsolete individuals are obsolete and children of modern individuals are modern. The last assumption represents the notion that a person cannot impart knowledge to their offspring that they do not themselves have. Thus, the age-0 populations of modern and obsolete individuals in cohort t is

$$p_{t,0}^i = \text{CBR}_{t-1} \sum_{j=0}^{J} p_{t-1-j,j}^i \qquad \text{for } i = o, m. \tag{6}$$

From age 0 to age $k-1$, the populations of modern and obsolete individuals evolve according to

$$p_{t,j+1}^i = s_j^i p_{t,j}^i \qquad \text{for } j = 0, \ldots, k-1 \ \text{ and } i = o, m. \tag{7}$$

The economically active $(j \geq k)$ populations of modern and obsolete individuals of generation t evolve according to

$$p_{t,j+1}^o = s_j^o p_{t,j}^o \left[1 - Q(y_{t,j}, h_{t,j}, \pi_{t+j}) \right] \tag{8}$$

$$p_{t,j+1}^m = s_j^m \left[p_{t,j}^m + Q(y_{t,j}, h_{t,j}, \pi_{t+j}) p_{t,j}^o \right]. \tag{9}$$

Equation (8) indicates that a fraction $1 - Q(y_{t,j}, h_{t,j}, \pi_{t+j})$ of the age-j obsolete remain obsolete and that s_j^o of them survive to the next age. Equation (9) indicates that $Q(y_{t,j}, h_{t,j}, \pi_{t+j}) p_{t,j}^o$ age-j obsolete become modern and that, together with the $p_{t,j}^m$ already-modern, they face survival probability s_j^m.

Recall that π_t refers to the proportion of economically-active modern individuals at the beginning of period t, i.e. $\pi_t = \tilde{p}_t^m / (\tilde{p}_t^m + \tilde{p}_t^o)$ where $\tilde{p}_t^i = \sum_{j=k}^{J} p_{t-j,j}^i$, $i = o, m$.

The deaths to account for in a period are: deaths of those who started the period as modern individuals; deaths of those who started as obsolete but adopted the modern technology; deaths of those who started as obsolete and did not adopt the modern technology; and, finally, deaths of all age-J individuals.

$$\text{CDR}_t = \frac{1}{p_t} \underbrace{\sum_{j=0}^{J-1} p_{t-j,j}^m \left(1 - s_j^m\right)}_{\text{modern}} + \frac{1}{p_t} \underbrace{\sum_{j=0}^{J-1} p_{t-j,j}^o Q(y_{t-j,j}, h_{t-j,j}, \pi_t) \left(1 - s_j^m\right)}_{\text{obsolete adopting}}$$

$$+ \frac{1}{p_t} \underbrace{\sum_{j=0}^{J-1} p_{t-j,j}^o \left(1 - Q(y_{t-j,j}, h_{t-j,j}, \pi_t)\right) \left(1 - s_j^o\right)}_{\text{obsolete not adopting}} + \frac{1}{p_t} \underbrace{(p_{t-J,J}^m + p_{t-J,J}^o)}_{\text{age } J}. \tag{10}$$

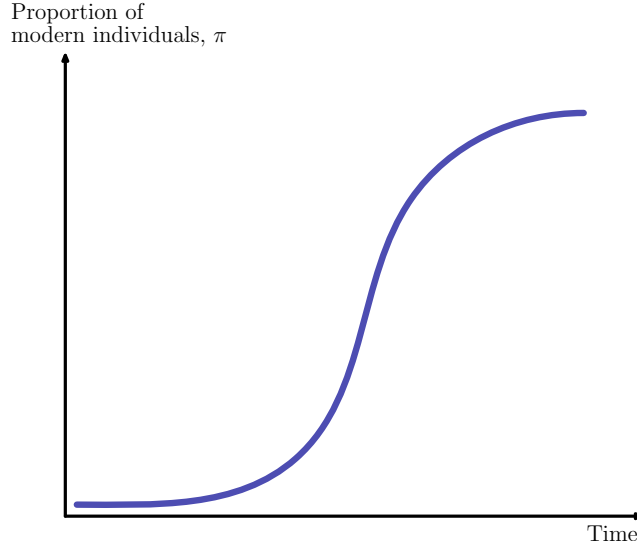where $p_t = \sum_{j=0}^{J} p_{t-j,j}^m + p_{t-j,j}^o$.

Figure 5: The S-shaped diffusion of the modern technology

2.2 Technology diffusion

The flow of individuals of age j from obsolete to modern during period t is

$$Q(y_{t-j,j}, h_{t-j,j}, \pi_t)p^o_{t-j,j}.$$

Let Q be of the form $Q(y, h, \pi) = \pi G(y, h)$—as we assume in Section 3 for our quantitative exercise—then the flow from obsolete to modern is

$$Q(y_{t-j,j}, h_{t-j,j}, \pi_t)p^o_{t-j,j} = \pi_t p^o_{t-j,j} G(y_{t-j,j}, h_{t-j,j}).$$

This equation is similar to the "flow-of-infection" equation found in SIR-type models in epidemiology. The difference is that instead of describing the spread of a disease, our model describes the spread of life-preserving practices. As in SIR-type models, our model exhibits the familiar S-shaped diffusion pattern. When the proportion of modern individuals is close to 0, the flow of adopters is "small." When the proportion of modern individuals is close to 1, there are few obsolete individuals and, therefore, the flow of adopters is "small" as well.

The S-shaped pattern of diffusion is important for the convergence of mortality between poor and rich countries. This is illustrated in Figure 5. Consider two countries, rich and poor, at a given point in time. The proportion of modern individuals is high in the rich country and, therefore, mortality is low. That is, the rich country is on the flat, upper part of the diffusion curve in Figure 5. Adoption (the slope of the S-curve) is therefore slow and mortality declines slowly. The poor country, on the other hand, has a lower proportion of modern individuals and thus a higher mortality rate. That is, it is on the lower part of the diffusion curve. Figure 5 illustrates that both adoption and the decline

in mortality are then faster than in the rich country.

3 QUANTITATIVE ANALYSIS

3.1 Functional forms

We use the following functional forms. The probability of adoption of the modern technology is

$$Q(y, h, \pi) = \pi G(y, h) \tag{11}$$

where

$$G(y, h) = \Lambda \left(1 - e^{-\lambda \left(\alpha_y y^\theta + \alpha_h h^\theta \right)^{1/\theta}} \right). \tag{12}$$

We use a utility index from the CARA family to represent preferences:

$$U(c) = \Sigma - \exp(-\sigma c).$$

The CARA form for U, together with the formulation for Q and G, yields an analytical solution to the first-order conditions of an individual's optimization problem and, therefore, reduces the computational cost of fitting the model to the data. (We have to track $J \times 2$ individuals at each point in time.) We check, in our computations, that U is always positive. This is an important restriction in models of this nature, as noted by Rosen (1988): When the utility index is negative an extra year of life reduces utility. Appendix B describes the optimal y and h.

3.2 Calibration

We calibrate the model to Swedish data. The Palgrave Macmillan Ltd (2013)'s vital statistics for Sweden are the longest of all developed countries. The Swedish data also illustrate a case where the decline in CDR precedes the economic takeoff—see Panel A of Figure 6. We use 1751 as our first date since the age distribution data from the Human Mortality Database starts in 1751.

There are two exogenous variables: CBR and TFP. We use the observed CBR of Sweden from 1751 to 2017 and Sweden's GDP for w_t. Figure 6 shows Sweden's GDP and its trend and CBR. The sequence of growth rates for w_t that best reproduces (in a least-squares sense) Sweden's GDP are 0% prior to 1800, 1.2% from 1800 to 1900, and 2.19% from 1900 to 2017 (Maddison Project Database 2020).

Age-specific survival probabilities for Sweden in 1751 and 2017, which are available from the Human Mortality Database, represent the survival probabilities of o- and m-individuals, respectively. The age-specific survival probabilities are available for all ages from 0 to 111. Thus, we set $J = 111$. Panel A of Figure 7 shows the survival probabilities.

We set the model period to be 1 year and the discount factor β to 0.95. We consider the economy from

A – Crude death rate and
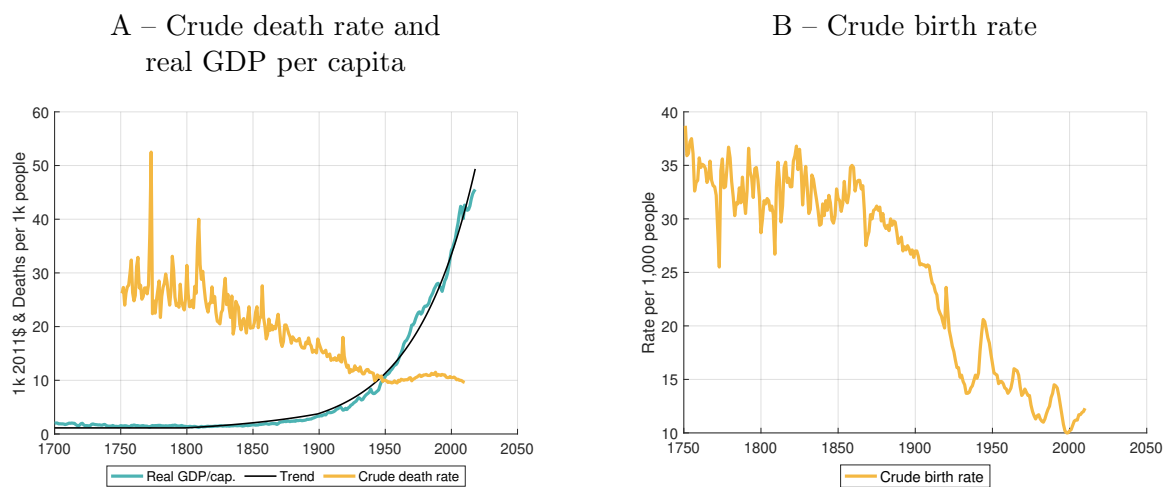real GDP per capita

B – Crude birth rate

Figure 6: Vital statistics and GDP in Sweden

A – Survival probabilities from
0 to 90 in 1751 and 2017

B – Age distribution
in 1751
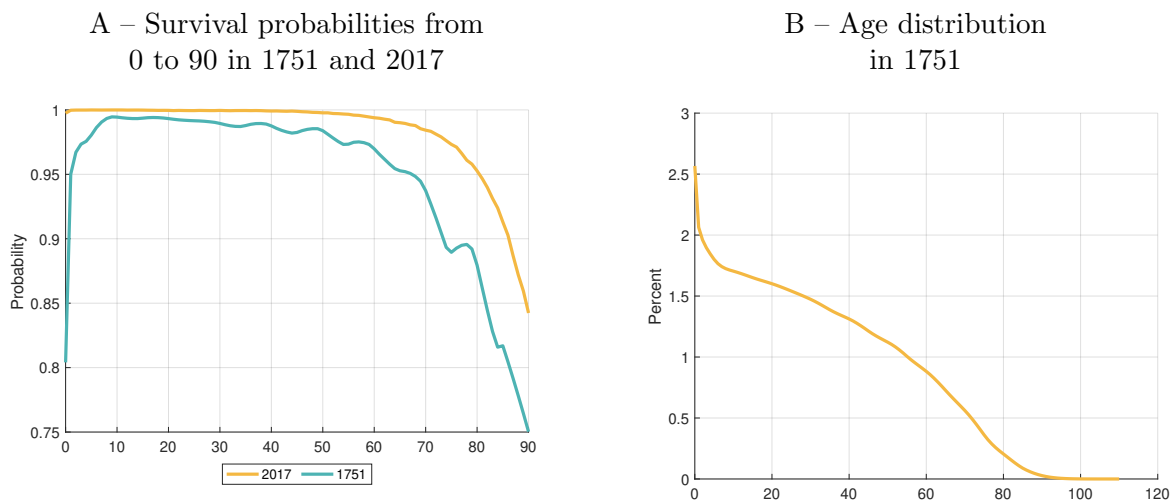
Figure 7: Survival probabilities and the initial age distribution

date $t = 1, \ldots, T$, where date 1 corresponds to 1751 and date T to 2017. The size of the population and income w at date 1 are both normalized to 1. The age distribution of population at date 1 is that of Sweden in 1751. Panel B of Figure 7 shows the 1751 age distribution. We set the economically active age, k, equal to 15.

Let $\omega = (\lambda, \Lambda, \sigma, \Sigma, \alpha_y, \alpha_h, \theta, \pi_1)$ denote the list of parameters to be determined and where π_1 represents the initial proportion of m-individuals. Let $\text{CDR}_t(\omega)$ denote the crude death rate implied by the model and defined in Equation (10). We determine ω as the solution to the distance-minimization problem:

$$\min_{\omega} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left[\text{CDR}_t(\omega) - \text{CDR}_t(\text{data})\right]^2, \tag{13}$$

where $\text{CDR}_t(\text{data})$ represents the observed crude death rate in Sweden. We use 8 parameters to match a time series of $T = 267$ observations.

Table 1 reports the calibrated parameters. Figure 8 shows the model's fit to the time series of CDR (Panel A). The correlation between model and data is 91%. There are, however, spikes in mortality that our model does not reproduce. For instance, two sharp increases in mortality occur in 1772 and 1773 and in 1834. Larsson (2020) argues that "dysentery (...), fevers, and smallpox caused the sharp rise in mortality in the severe crisis of 1772 and 1773." The high mortality in 1834 is due to a cholera epidemic. Our model does not have an exogenous shock during these specific years to account for the spikes in mortality.

The model is also consistent with the timing of mortality decline despite the *absence* of economic takeoff. The mortality rate declined from 28 in 1751 to 20 in 1870; Jörberg (1965) notes that the Industrial Revolution occurred in Sweden in the 1870s.

Non-calibrated moments The annual rate of population growth in the model is consistent with the observed population dynamics in Sweden. The correlation between model and data is 80%. Recall that the model does not have migration into and out of Sweden, and the rate of population growth in the model is the difference between CBR and CDR.

Figure 9 shows a comparison between the model's age distribution of population and the data in 1850 and 1950. Recall that the the initial period in the model is 1751. The age distributions in Figure 9 are 100 years and 200 years later. The model's age distribution is broadly consistent with the data. The evolution of the age distribution is governed by the number of births and the number of people who survive in each age group in each period. The former is exogenous and matches the data by construction. The latter is the result of forces in the model: Given the age-specific mortalities in 1751 and 2017, o-individuals choose to spend the resources to acquire the m-technology and thereby dictate the number of people who survive in each age group.

Table 1: Model parameters

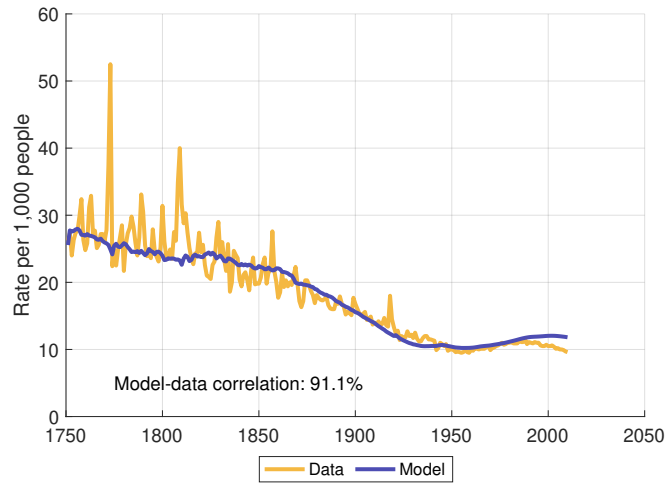|	
Exogenous	$\beta = 0.950$, $k = 15$, $J = 111$
Calibrated	
Preferences	$\sigma = 3.1468$, $\Sigma = 0.1682$,
Technology	$\alpha_y = 0.0359$, $\alpha_h = 1.0088$, $\theta = 0.2086$
	$\lambda = 0.0178$, $\Lambda = 1.3907$
Initial prop.	$\pi_1 = 0.4828\%$
of m-individuals	

Figure 8: Crude death rate in Sweden, 1751-2017

Source: Palgrave Macmillan Ltd (2013) and authors' calculations.

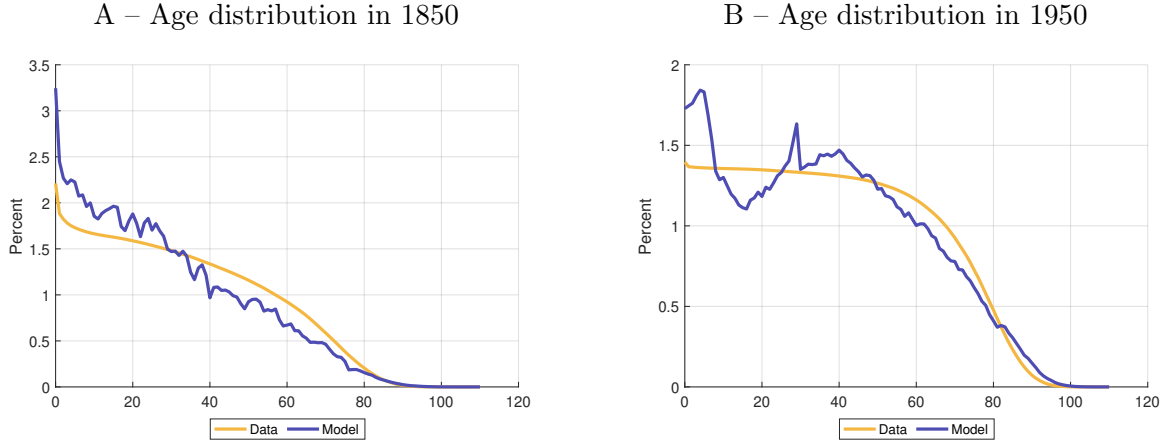A – Age distribution in 1850   B – Age distribution in 1950

Figure 9: Age distribution in Sweden, 1850 and 1950

Note: Data on population by age are interpolated between decennial dates while model-generated age distribution is not. Hence, the data line is smoother than that in the model line in Panels C and D.
Source: Panels A and B: Palgrave Macmillan Ltd (2013) and authors' calculations; Panel C and D: Human mortality database and authors' calculations.

3.3 Mortality and population dynamics across countries: 1960-2018

Using the World Bank's World Development Indicators we assemble annual data on CBR, CDR, population, and GDP from 1960 to 2018 for a group of 87 countries. We exclude China from the sample as China's mortality in the early 1960s during the Great Leap Forward (1958-62) and the ensuing famine increased by 150% in a span of a few years. As a share of the world excluding China, the 87 countries in our sample cover 74% of the population in 1960, 78% of deaths, and 79% of GDP. In 2018, these numbers are 78%, 76%, and 80%, respectively. These countries also span large differences in income: The average GDP in the poorest decile is 1.1% of that in the richest decile in 1960, and 2.1% in 2018. The annual rate of population growth for our sample is 1.77%. It is almost entirely driven by a rate of natural increase (birth rate minus death rate); the rate of net migration (immigration minus outmigration rate) is 0.18%.

We use the calibrated model to simulate time series of CDR for each country in our sample. That is, for each country i in year t we use its GDP for w_t and its CBR_t. We then find the year t_i when Sweden's CDR was the closest (in a least-squares sense) to country i's CDR in 1960, and we use the age distribution of Sweden's population in year t_i as the 1960 age distribution for country i. We adopt this procedure because we do not have the 1960 age distribution of the population for most countries in our sample. Finally, we find the initial proportion of m-individuals for country i to fit the time series of country i's CDR from 1960 to 2018 in a least-squares sense. Compared with the calibration of Section 3.2, we use only *one* parameter to match country i's CDR entire time series. We then simulate the time series of CDR and population from 1960 to 2018.
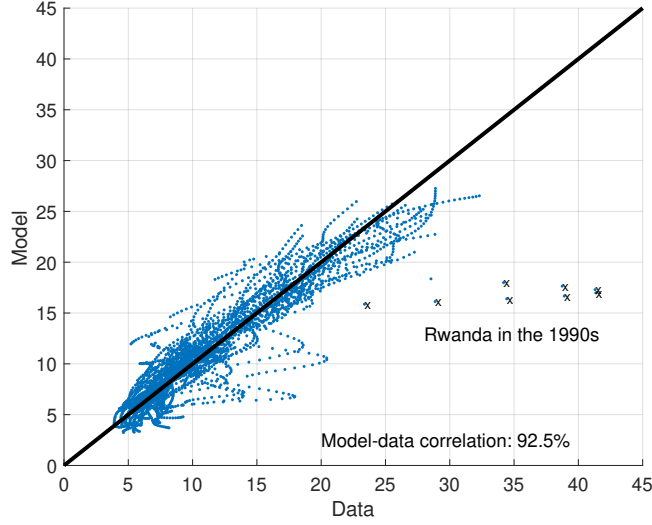
Figure 10: Model fit

Note: Each point denotes CDR for a country-year and "x" denotes data- and model-CDR for Rwanda.
Source: World Development Indicators and authors' calculations.

Figure 10 shows the model fit. Each point represents CDR for a country-year. Rwanda in the 1990s is a noticeable outlier: The civil war started in 1990 and culminated with the 1994 genocide. The model-data correlation is 92.5%.

3.3.1 *Change in population*

The model performs well on the dynamics of total population and on the distribution of population across countries. Total population in our sample increases by 3.1 billion between 1960 and 2018. The corresponding change in the model is 3.0 billion i.e., the model accounts for 97% of the data.

The *change* in population is uneven across countries and is predominantly due to poor countries. To compute the contribution of various countries to the change in total population, we group countries into quintiles of GDP in 1960, and track these countries over time, keeping the composition of the quintiles the same at all dates. Let $P_{j,t}$ denote the population of countries in quintile j and P_t denote total population at date t. Then $P_t = \sum_{i=1}^{10} P_{j,t}$ and $P_{2018} - P_{1960} = \sum_{j=1}^{5} P_{j,2018} - P_{j,1960}$. The contribution of countries in quintile j to the change in total population is then

$$\frac{P_{j,2018} - P_{j,1960}}{P_{2018} - P_{1960}}. \tag{14}$$

Table 2 reports the contributions. The bottom quintile (poor countries) accounts for almost one half of the change in total population. The top quintile (rich countries) accounts for less than 10%.

To compute the model-equivalent of these statistics we assign to each country i its observed population

16

Table 2: Contributions of poor and rich countries to the change in total population

	Data	Model
Poor	0.47	0.49
Rich	0.08	0.05

Note: Columns 2 and 3 use Equation (14) to compute observed and model-generated populations.
Source: World Development Indicators and authors' calculations.

in 1960. We then use the CDR implied by the model and the observed CBR for country i to calculate:

$$P_{i,t+1} = P_{i,t} \left(1 + \text{CBR}_{i,t} - \text{CDR}_{i,t}\right).$$

The population of quintile j is the sum of populations of the countries that belong to quintile j. The total population is the sum of the populations across all quintiles. The model-implied contributions of poor and rich countries to the change in total population is reported in Table 2. The poor countries contribute 49% of the change in total population, which overshoots the data. The rich contribute 5%, which accounts for 63% (5/8) of the data.

The model delivers the increase in total population as well as the cross-country increase in population in a manner that is consistent with the observed mortality dynamics. For instance, the mortality rate in the entire sample declined from 16.4 in 1960 to 7.8 in 2018 i.e, the mortality rate declined by 8.6 deaths per thousand people. The corresponding decline in the model is 7.5. Similar to the change in total population, the observed change in mortality in the entire sample is predominantly due to mortality dynamics in poor countries: CDR declined by 14.5 deaths per thousand people in poor countries. In the model, the corresponding decline in poor countries is 12.

As noted, almost one half of the increase in total population is due to poor countries. This increase occurred despite the CBR decline in poor countries from 43 births per thousand people in 1960 to 20 in 2018. To account for the increase in the population of poor countries (almost 1.5 billion), the decline in their mortality is a crucial quantitative component.

3.3.2   *Mortality*

The model is consistent with the well-known negative correlation between GDP and CDR. In 1970, for instance, the correlation is -65% in the data and -70% in the model. Countries with higher GDP tend to have lower CDR. It does not imply, however, that an increase in a country's GDP *over time* would result in a decrease in its CDR. The country-specific dynamics of CDR in our model are determined by the diffusion of the modern technology and the speed of diffusion depends upon TFP and the dynamic externality.
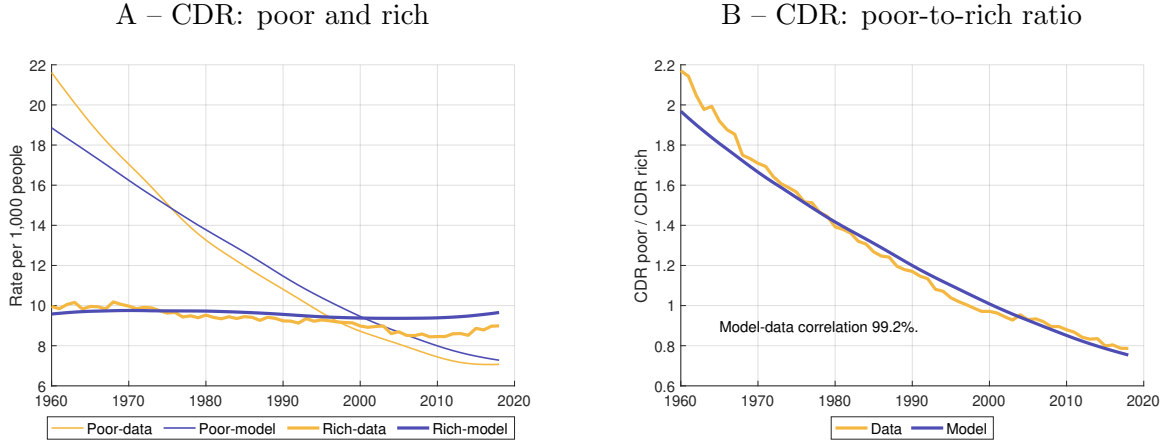
Figure 11: Crude death rates in poor and rich countries

Note: Crude death rates are population weighted in both model and data.
Source: World Development Indicators and authors' calculations.

Convergence   Figure 11, Panel A illustrates a less-known fact: convergence in the (population-weighted) CDR of rich and poor countries. CDR of poor countries remains higher than that of the rich, but the difference shrinks over time in the data and model. Panel B shows the poor-to-rich ratio of CDR in the data and model. Together, Panels A and B indicate convergence in CDR. Quantitatively, the model accounts for almost all of the convergence: The model-data correlation of the poor-to-rich CDR ratio is 99.2%.

As noted earlier, the convergence in CDR occurs despite the divergence in GDP across countries (Figure 2, Panel A). This feature is readily seen in Figure 12, where we illustrate the cross-country elasticity of CDR with respect to GDP in the model and data. The model-data correlation is 98.4%. While the elasticity is negative in the cross section every year, the elasticity has steadily increased from almost -0.25 to 0.

Convergence in mortality despite the lack of convergence in income also points to another less-known fact: *Economic growth is not a prerequisite for mortality decline.* Figure 13 is the model equivalent of Figure 3. Recall that the GDP used in the model is that of the data and, thus, the only difference between Figures 3 and 13 is the behavior of CDR implied by the model.

The correlation between the growth rate of GDP and that of CDR is positive in the data and model, i.e., there is no evidence that countries experiencing faster growth also experience larger decline in mortality.

Mortality miracles   Economic takeoff is unnecessary for mortality declines is also reinforced by *mortality miracles* in the data. In Table 3 we present all the countries in our sample that experienced negative growth of GDP. All of these countries also experienced a decline in mortality. In the model the
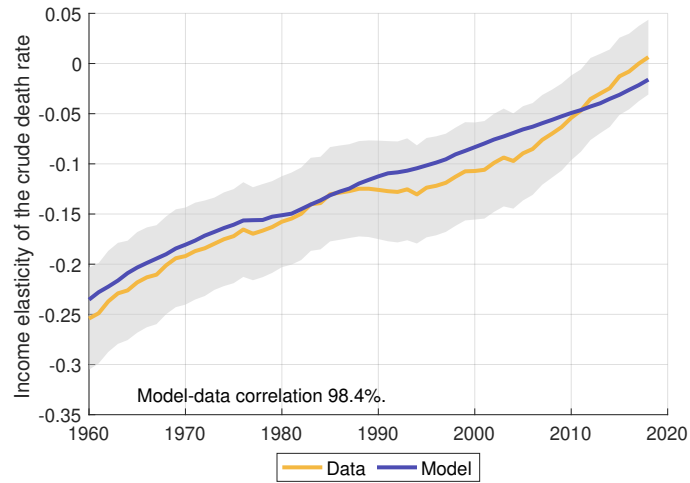
Figure 12: Cross-country elasticity of CDR with respect to GDP

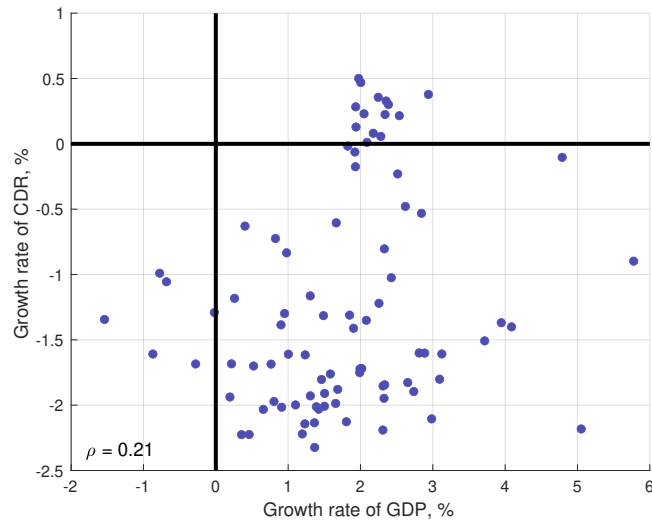*Source*: World Development Indicators and authors' calculations.

Figure 13: The growth rates of real GDP per capita and the crude death rate in the model, 1960-2018

Source: World Development Indicators.

Table 3: Mortality miracles

	Growth rate of GDP	Growth rate of CDR	
		data	model
Burundi	−0.02	−1.82	−1.29
Central African Republic	−0.78	−1.41	−0.99
Congo (Kinshasa)	−1.54	−1.49	−1.34
Haiti	−0.28	−1.65	−1.68
Madagascar	−0.87	−2.39	−1.61
Niger	−0.68	−2.12	−1.06

Note: The table presents the annual growth rate of GDP and CDR from 1960 to 2018 in both data and model.
Source: World Development Indicators and author's calculations.

list of countries that experienced negative growth is exactly the same as in the data by construction. These countries also experienced decline in mortality in the model. In Madagascar, for instance, GDP declined at an annual rate of 0.87% while its mortality declined at an annual rate of 2.39% in the data and 1.61% in the model.

3.4 World and Europe: 19th and 20th centuries

3.4.1 *World's Population Dynamics*

We use our model to generate a time series of the world's population from 1800 to 2010. (The length of the time series is constrained by the data available from gapminder.org.) We proceed as follows. The exogenous forces, and the preference and technology parameters are the same as those in Table 1. The exogenous time-varying variables we feed into the model are the world's CBR (Figure 1) and GDP from Maddison's Project Database. The initial value for w_t is scaled to match the world's GDP in 1800 relative to that of Sweden. The sequence of growth rates for w_t that best reproduces (in a least-squares sense) the world's GDP are 1% prior to 1950 and 2.4% after. Finally, we use the world's observed population in 1800 as the initial population size and Sweden's age distribution in 1751 as the initial age distribution (we do not have age-distribution data for the world's population).

We find two *time-invariant* parameters to fit the time series of the world's CDR from 1800 to 2010 in a least-squares sense. The two parameters are: the initial proportion of m-individuals, π_1, and a factor multiplying the sequence of age-specific survival probabilities of o-individuals.[3] Unlike the calibration in Section 3.2 for Sweden, we use only two parameters to match a time series of 211 observations for the world. We find $\pi_1 = 0.09\%$ and the factor multiplying the survival probabilities of o-individuals

---

[3]The world's CDR in 1800 is abnormally high at 37. Even with 100% of obsolete individuals in 1800, the model cannot generate a CDR of this magnitude. In other words, the world's obsolete technology in 1800 was worse than that of Sweden in 1751.
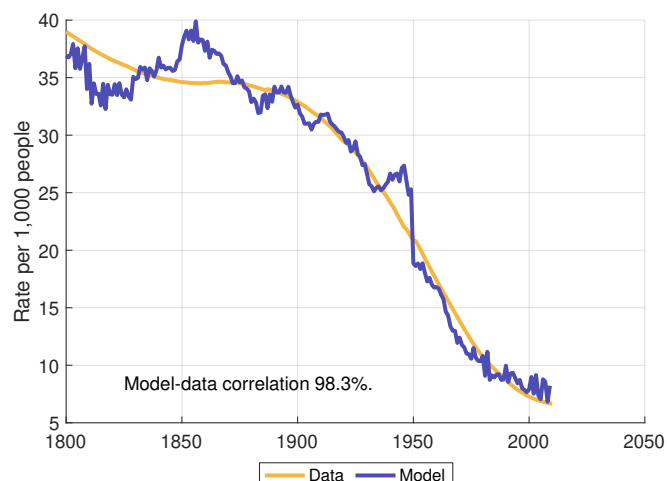
Figure 14: Model fit: Crude death rate in the world, 1800-2010

Source: https://gapminder.org/ and authors' calculations.

is 0.99. Figure 14 shows the model fit: The correlation is 98.3%.

Table 4 shows the implications of this experiment for population growth. It took 82 years for the world's population to increase from 1 billion to 1.5, and 8 years to go from 3 to 3.5. That is, it took 74 fewer years to increase population by the same quantity. The model reproduces this acceleration well: It takes 73 fewer years for the model's population to increase from 3 billion to 3.5 billion than from 1 to 1.5 billion. Thus, the model accounts for almost all (99%) of the acceleration of the world's population.

Table 4: The world's population growth

Population change	Data		Model	
	years	Annual growth	years	Annual growth
1.0 bil. to 1.5	82	0.49%	82	0.49%
1.5 bil. to 2.0	38	0.76%	38	0.76%
2.0 bil. to 2.5	24	0.93%	22	1.01%
2.5 bil. to 3.0	11	1.66%	11	1.66%
3.0 bil. to 3.5	8	1.93%	9	1.71%

Source: https://gapminder.org/ and authors' calculations.

A theory of population dynamics based solely on fertility cannot account for this acceleration. As Figure 1 shows, CBR exhibits a declining trend from 1800 to 2010. A model of fertility with exogenous
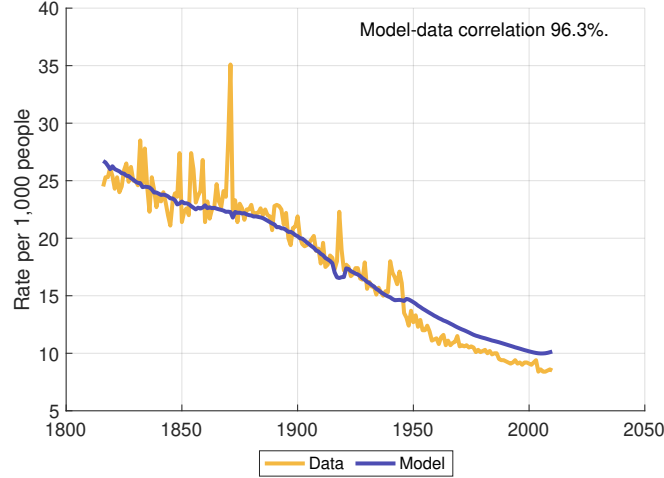
Figure 15: Model fit: Crude death rate in France, 1816-2010

Source: Palgrave Macmillan Ltd (2013) and authors' calculations.

mortality will have to rely almost exclusively on the exogenous decrease in CDR to account for the accelerating population.

3.4.2 *European countries*

We proceed in a manner similar to that of Section 3.4.1. For country i we use CBR from Palgrave Macmillan Ltd (2013) and GDP from the Maddison Project Database as exogenous variables. The initial value for w_t is scaled to match country i's GDP in year t_i (the first year for which data are available for country i) relative to that of Sweden in year t_i. We choose a sequence of growth rates for w_t to best reproduce (in a least-squares sense) country i's GDP. We use data from the Human Mortality Database for the initial age distribution of country i in year t_i.

As in Section 3.4.1, the preference and technology parameters are the same as those in Table 1. We calibrate *one time-invariant* parameter, namely the initial proportion of m-individuals, to fit the entire time series of country i's CDR. In the case of France, for instance, this means that we use one parameter to match a time series of 195 observations from 1816 to 2010. Figure 15 shows the model's fit to France's CDR. The correlation between model and data is 96.3%. Appendix C shows the model's fit to a few other European countries.

4   CONCLUSION

The World's population growth rate has increased for at least 200 years despite a declining birth rate. Thus, an economic theory of population dynamics requires an economic theory of mortality. The theory we develop emphasizes infections as causes of death and individual choices on costly adoption of a modern technology to avoid or survive the infections as a channel for longer life spans. The diffusion of modern technology is S-shaped and its speed is dictated by a dynamic externality. As the technology diffuses, individual choice to adopt the modern technology becomes more attractive and mortality decreases.

We calibrated our model to a long time series of crude death rates in Sweden, using the Swedish GDP and crude birth rate as exogenous driving variables. The model fits the Swedish data well and is consistent with the changing age distribution of the Swedish population over time. We then used the calibrated model to simulate crude death rates for a sample of 87 countries from 1960 to 2018, using their birth rates and GDP as exogenous driving variables. The model is consistent with the well-known negative cross-country correlation between crude death rate and GDP. It is also consistent with two less known facts: (i) the crude death rates in poor countries converges to those of rich countries despite a lack of convergence in GDP across countries and (ii) countries experience declining mortality even when GDP growth is zero or negative. The model accounts for 99% of the convergence. Countries with negative GDP growth do experience decreasing mortality in the model and in the data, and economic growth is not necessary for mortality declines.

We also show that the model reproduces the change in total population, as well as its distribution across poor and rich countries. Total population in our sample increases by 3.1 billion and the model accounts for 97% of the change. The increase in population is predominantly due to poor countries (almost one half) and the model is consistent with this fact. Furthermore, the model delivers the increase in population while being consistent with the observed mortality dynamics.

Finally, world population has been accelerating: It took 82 years for the world's population to increase from 1 billion to 1.5, and 8 years to go from 3 to 3.5. Our model reproduces this fact and accounts for 99% of the acceleration. A theory of population dynamics based solely on fertility will have to rely almost exclusively on the exogenous decrease in crude death rate to account for the accelerating population.

References

Acemoglu, Daron and Simon Johnson, "Disease and development: The effect of life expectancy on economic growth," *Journal of political Economy*, 2007, *115* (6), 925–985.

Becker, Gary S, "An economic analysis of fertility," in "Demographic and economic change in developed countries," Columbia University Press, 1960, pp. 209–240.

Clark, Gregory, "The agricultural revolution and the industrial revolution," *Unpublished Working Paper, UC Davis*, 2002.

Davis, Kingsley, "The amazing decline of mortality in underdeveloped areas," *The American Economic Review*, 1956, *46* (2), 305–318.

Fogel, Robert W., *The escape from hunger and premature death, 1700-2100: Europe, America, and the Third World* (Cambridge Studies in Population, Economy and Society in Past Time, number 38) 2004.

Jörberg, Lennart, "Structural change and economic growth: Sweden in the 19th century," *Economy and History*, 1965, *8* (1), 3–46.

Krantz, Olle, "Swedish GDP 1300-1560 A Tentative Estimate," *Lund Papers in Economic History: General Issues*, 2017, (152).

Larsson, Daniel, "Diseases in Early Modern Sweden: A Parish-level Study 1631-1775," *Scandinavian Journal of History*, 2020, *45* (4), 407–432.

Leukhina, Oksana M. and Stephen J. Turnovsky, "Population size effects in the structural development of England," *American Economic Journal: Macroeconomics*, 2016, *8* (3), 195–229.

Livi-Bacci, Massimo, *Population and nutrition: An essay on European demographic history* 1991.

Malthus, Thomas R., *An Essay on the Principle of Population*, J. Johnson, London, 1798.

Palgrave Macmillan Ltd, "A6 Vital Statistics: Rates per 1,000 Population," in Palgrave Macmillan Ltd, ed., *International Historical statistics*, 2013. https://doi.org/10.1057/978-1-137-30568-8_393.

Prescott, Edward C and Robert M Townsend, "Pareto optima and competitive equilibria with adverse selection and moral hazard," *Econometrica: Journal of the Econometric Society*, 1984, pp. 21–45.

Rogerson, Richard, "Indivisible labor, lotteries and equilibrium," *Journal of monetary Economics*, 1988, *21* (1), 3–16.

Rosen, Sherwin, "The value of changes in life expectancy," *Journal of Risk and Uncertainty*, 1988, *1* (3), 285–304.

Schön, Lennart and Olle Krantz, "New Swedish Historical National Accounts since the 16th Century in Constant and Current Prices," *Lund Papers in Economic History*, 2017, (140).

Smith, Adam, *An Inquiry into the Nature and Causes of the Wealth of Nations*, W. Strahan and T. Cadell, London, 1776.

Wrigley, E. A., R. S. Davies, J.E. Oeppen, and R.S. Schofield, *English population history from family reconstitution 1580-1837*, Cambridge University Press, 1997.

A   Example of life-saving practices

Below is a list of deadly diseases and the steps that individuals can take to prevent getting and spreading them. These steps require knowledge of how the diseases spread, but they do not require large-scale public interventions. Vaccination campaigns, when a vaccine exists, are not included below in the measures one can take on their own. The information below is available from the Center for Disease Control and Prevention, the North Dakota Department of Health, and the Kentucky Department of Public Health.

AIDS

AIDS is caused by the human immunodeficiency virus (HIV). The spread of the virus can be prevented by abstinence or using condoms.

Cholera

Cholera is a disease that one gets by drinking water or eating food contaminated with the bacteria Vibrio Cholerae. Casual contact between people does not spread cholera. One can reduce the risk of getting cholera by drinking only clean (bottled, boiled, or chemically treated) water, washing hands often, using clean water to wash dishes, brush teeth, wash and prepare food, and make ice and, finally, dispose of feces in a manner that prevents contamination of water and food.

Ebola

The Ebola Virus Disease (EVD) spreads to people through direct contact with the blood, body fluids, and tissues of an infected animal. The virus then spreads from person to person through direct contact with blood or body fluids, and contaminated objects.

One can reduce the risk by avoiding contact with blood and body fluids, of sick people and people who died from EVD and avoiding contact with contaminated objects. One should avoid funeral or burial practices that involve touching the body of someone who died from EVD. One should avoid contact with bats, forest antelopes, and nonhuman primates blood, fluids, or raw meat prepared from these or unknown animals (bushmeat).

The plague (Yersinia pestis)

The plague bacteria can be transmitted to humans via flea bites, contact with contaminated fluid or tissue and infectious droplets. Transmission of these droplets is the only way that plague can spread between people. One can reduce the risk of becoming infected by limiting rodent habitat around

home, work place, and recreational areas. One should wear gloves when handling infected animals, use repellent and keep fleas off of pets. Wearing a surgical mask also reduces the risk of transmitting and catching the Plague.

Polio (Poliomyelitis)

Polio is caused by the poliovirus which is found in the stool and throat of infected people. Besides the vaccine, one can prevent the spread of polio by washing hands with soap and water after using the bathroom and changing diapers, and before preparing food and eating.

B Optimization

Consider the first-order conditions (4)-(5), abstracting from time and generation subscripts for simplicity,

$$
\begin{aligned}
0 &= U_1\left(w\left(1-h\right)-y\right) - \beta Q_1\left(y,h,\pi\right)\Delta \\
0 &= U_1\left(w\left(1-h\right)-y\right)w - \beta Q_2\left(y,h,\pi\right)\Delta.
\end{aligned}
$$

Given the functional form for Q (Section 3), these conditions imply

$$
\begin{aligned}
0 &= U_1\left(w\left(1-h\right)-y\right) - \beta\pi\Lambda\lambda\exp\left(-\lambda\chi\left(y,h\right)\right)\chi_1\left(y,h\right)\Delta \\
0 &= U_1\left(w\left(1-h\right)-y\right)w - \beta\pi\Lambda\lambda\exp\left(-\lambda\chi\left(y,h\right)\right)\chi_2\left(y,h\right)\Delta
\end{aligned}
$$

where $\chi\left(y,h\right) = \left(\alpha_y y^\theta + \alpha_h h^\theta\right)^{1/\theta}$. This implies $1/w = \chi_1\left(y,h\right)/\chi_2\left(y,h\right)$, where $\chi_1\left(y,h\right) = \chi\left(y,h\right)^{1-\theta}\alpha_y y^{\theta-1}$ and $\chi_2\left(y,h\right) = \chi\left(y,h\right)^{1-\theta}\alpha_h h^{\theta-1}$. Hence, $h = yX\left(w\right)$ where,

$$
X\left(w\right) = \left(w\frac{\alpha_y}{\alpha_h}\right)^{1/(\theta-1)}.
$$

Note that, at the optimum,

$$
\chi\left(y,h\right) = y\left(\alpha_y + \alpha_h X\left(w\right)^\theta\right)^{1/\theta}
$$

and

$$
\chi_1\left(y,h\right) = \left(\alpha_y + \alpha_h X\left(w\right)^\theta\right)^{1/\theta-1}\alpha_y \equiv \chi_1(w),
$$

where the identity, abusing notations, indicates that the first-derivative of χ with respect to y is a function of w only at the optimum. Consumption is $c = w - y\left(1 + wX\left(w\right)\right)$. Given the functional form for U (Section 3), the first-order condition for y is

$$
\sigma\exp\left(-\sigma w + \sigma y\left(1 + wX\left(w\right)\right)\right) = \beta\Delta\pi\Lambda\lambda\exp\left(-\lambda y\left(\alpha_y + \alpha_h X\left(w\right)^\theta\right)^{1/\theta}\right)\chi_1\left(w\right)
$$

or

$$
y\left[\sigma\left(1 + wX\left(w\right)\right) + \lambda\left(\alpha_y + \alpha_h X\left(w\right)^\theta\right)^{1/\theta}\right] = \sigma w + \ln\left(\frac{\beta\Delta\pi\Lambda\lambda}{\sigma}\chi_1\left(w\right)\right).
$$

A – Norway, 1846-2010

B – Switzerland, 1876-2010

C – Denmark, 1835-2010
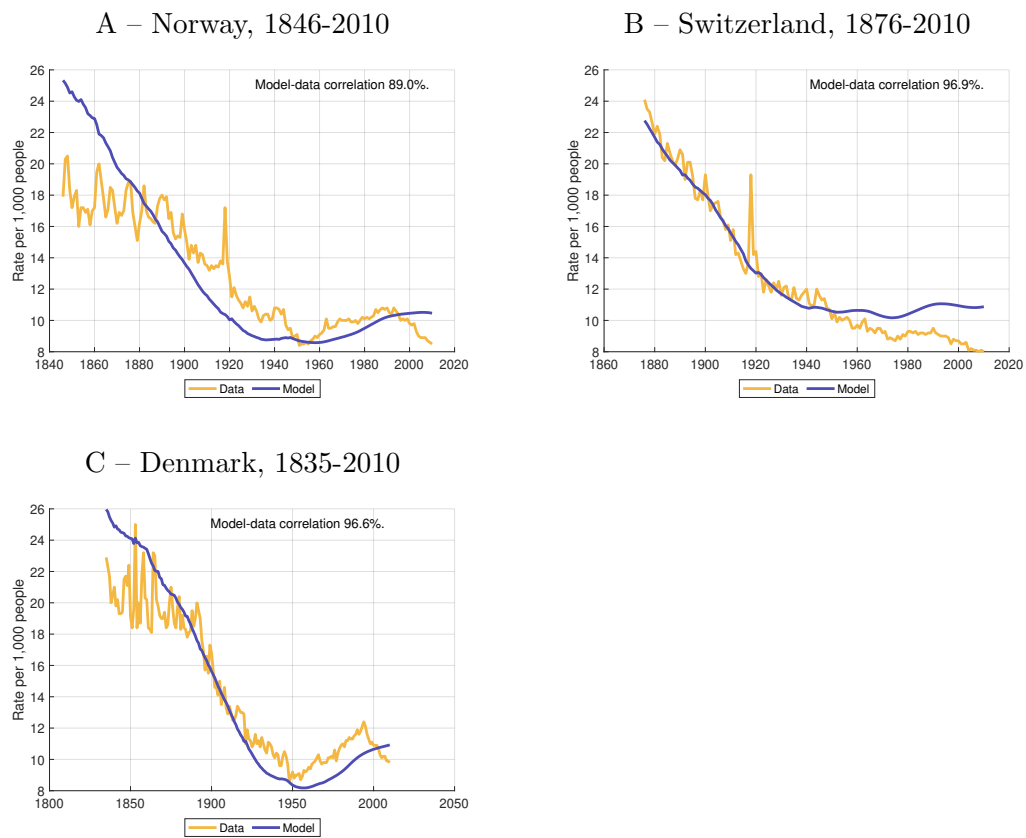
Figure C.1: Crude death rates in Europe