

# The Role of Mortality in the Transmission of Knowledge\*

Michael Bar<sup>†</sup> and Oksana Leukhina<sup>‡</sup>

October 2009

We investigate, both theoretically and quantitatively, a previously unexplored link between adult mortality and growth. Our mechanism allocates a central role to individuals as carriers of useful ideas and to personal contact as an important means of passing on the useful ideas to future generations. It thus implies that disrupting a human life will impede the process of knowledge transmission across generations. We embed this mechanism in a simple model of endogenous growth and fertility and use it to study its relevance in the application to the long-run growth experience of England.

**Keywords:** Economic growth, adult mortality, longevity, knowledge transmission, knowledge creation, ideas, human capital

**JEL Codes:** O40, O31, O33, J10, N10

\*We thank Michele Boldrin, Oded Galor, Fidel Pérez Sebastián, John Seater and David Weil for their comments, as well as the participants of the 2008 ASSA meetings and March 2008 Growth Workshop sponsored by Duke University.

<sup>†</sup>Bar: Department of Economics, San Francisco State University, San Francisco, CA 94132.

<sup>‡</sup>Leukhina: Department of Economics, University of North Carolina, Chapel Hill, NC 27599.

# 1. Introduction

Motivated by the strong empirical association between gains in longevity and growth, found both in the time series and cross-sectional studies (e.g. Kelley and Schmidt (1995), Shastry and Weil (2003), Lorentzen et al. (2005)),<sup>1</sup> we investigate a previously unexplored link between adult mortality and growth.

Our mechanism is rooted in the premise that in preindustrial and traditional societies, a large body of useful knowledge was tacit, uncodifiable and embodied collectively in individuals. An effective transfer of tacit knowledge requires extensive personal contact. In fact, we observe that most learning in traditional societies was accomplished through master-apprentice and parent-child relationships. In a society where knowledge is embodied in individuals and personal contact is essential for knowledge transfer, a premature death of the master (parent) interrupts the technique transfer to their apprentices (children), and it also implies a loss of their knowledge and experience to draw upon. These factors represent the direct impact of mortality on productivity growth. In addition, the anticipation of the low rate of knowledge transmission over time lowers returns to creation of new knowledge, representing the indirect impact of mortality. Our main hypothesis is that reduction in adult mortality facilitated the take-off in per capita output by improving knowledge transmission across generations and by encouraging more innovation.

Episodes of high mortality are historically known to disrupt the process of knowledge transmission. Recorded episodes of such losses of useful knowledge due to mortality are numerous. Archaeological findings reveal that bubonic plague and scarlet fever epidemic that largely depopulated the Senecas in the 1630s eliminated much of their knowledge of ceramic craftsmanship (Halverson 2007). Most recently, the AIDS epidemic in Sub-Saharan Africa is negatively affecting agricultural productivity through the loss of indigenous farming practices, specialized skills and experience.<sup>2</sup> According to the Food and Agricultural Organization of the United Nations, premature death blocks the transmission of knowledge and practices to children and contributes to the spread of ineffective farming techniques. Waterhouse (2005) documents a large negative impact of HIV/AIDS on the farmers' knowledge of seed by conducting farmer surveys in several provinces of Mozambique. The negative impact of mortality on local productivity is also revealed through a well-documented practice of preindustrial trade guilds to send their skilled craftsmen to places that recently experienced an epidemic. There is also a record of several preindustrial episodes of productivity growth, whose interruption is associated with a high mortality episode (Black Death for the case of 14th c. Europe, wars with the Ottoman empire for the case of the Italian Golden Age, Jurchen and Mongol invasions for the case of Sung China). Monteiro and Pereira (2007) provide a description of these.<sup>3</sup>

We first derive the impact of adult mortality on productivity, starting with reasonable assumptions about how ideas embodied in individuals are transferred to the next generations and how these ideas aggregate to deliver the aggregate productivity level. Our explicit modeling of the mapping of ideas residing in individual's minds into aggregate productivity is closely related to the approach developed in Lucas (2008). We obtain a simple form of adverse impact of adult mortality on productivity growth that captures ideas described above. The destruction rate of productivity is an increasing and convex function of adult mortality. Convexity arises due to the assumption of diminishing marginal

<sup>1</sup>Acemoglu and Johnson (2005) find no evidence that increases in output per capita levels are associated with increases in life expectancy. However, the association of gains in life expectancy and gains in output per capita growth is also positive in their dataset.

<sup>2</sup>This information is collected from [www.un.org](http://www.un.org) and [www.fao.org](http://www.fao.org).

<sup>3</sup>Note that even though we are focusing on mortality episodes, any other abrupt displacement of a group of people should deliver the same effects. For example, the fall of the Western Roman Empire caused a dramatic loss in productivity in the west. Roads deteriorated, the rule of law diminished, political organization fell apart, trade diminished greatly, and all sorts of technology and knowledge was truly forgotten.

returns to experience associated with a given idea and to the number of carriers of this idea. We then embed this adverse impact of adult mortality on productivity retention into a general equilibrium model of endogenous technological change and fertility.<sup>4</sup> The model explicitly incorporates the process of knowledge creation and its transmission within and across generations. We allow for knowledge diffusion across locations in order to moderate the adverse impact of a local mortality shock. Innovation and improvement of existing techniques is carried out by young adults, whose sole motivation for doing so is greater productivity in the future, should they remain alive.

To assess the quantitative relevance of our hypothesis, we apply our model to study the case of the English take-off. We argue that our model built on the notion of embodied human capital is suitable for studying the case of England. Indeed, Mokyr (1998) argues that "most practical knowledge in the 18th c. was informal, often uncodified and passed on vertically from master to apprentice or horizontally across agents... Printed text might have remained secondary to personal contact throughout the 19th c..." David Mitch, summarizing a number of historical sources, concludes that informal human capital accumulation was at the heart of the Industrial Revolution: "technological advance [in the 19th c.] was primarily due to the practical experience by men of little or no formal scientific training." He further documents very small gains in literacy rates in the 19th c. England (Mitch 1998). It is also true that even in preindustrial time, landlords and tenant farmers actively engaged in the time-consuming process of improving their methods of production. Macdonald (1979) using postal records uncovers a great variety of ways farmers actively used to improve their production techniques, such as traveling and corresponding with other farmers, sending their children to another farm for a year to learn a more advanced technique and attempting new techniques themselves.

We perform the following calibration exercise: We first parameterize the model using the method of moments and the assumption of balanced growth to capture the available observations for England around 1600-1700. We employ historical data on population density, growth of output per capita, age-specific mortality, birth rates, real wage response to the Black Death epidemic and records on farmers' willingness to pay for sending a worker away for a year to learn a better technique. We also estimate the evolution of the stochastic process of adult mortality shocks using the maximum likelihood approach and time series data on adult mortality rates, finding that the frequency and severity of shocks has declined beginning around the mid 18th c.<sup>5</sup> Taking the mortality shock process as exogenously given, we then quantitatively study its role in the take-off from stagnation. We find that the decline in adult mortality can generate about 70% of the rise in the log of output per capita observed in the data, with most of the effect transpiring according to our hypothesis.

An interesting feature of the calibrated model is the dramatic difference in the effect that temporary and permanent changes in mortality introduce on output per capita. A temporary unexpected high mortality shock is actually "good" for output per capita because of its effect on per capita land holdings, and despite its negative impact on productivity. This is consistent with positive Malthusian checks. However, a permanent and expected rise in mortality has a negative impact on output per capita as, in anticipation of low retention of knowledge and low probability of own survival to the next period, labor moves into production of output and away from innovation. Productivity growth suffers more relative to the case of a temporary rise in mortality, because in addition to the decline in knowledge retention, less time

<sup>4</sup>Bar and Leukhina (forthcoming) estimate productivity in England during the period 1600-1910. They show that gains in productivity were the main driving force behind the economic transformation of England. Hence, we focus on changes in productivity, abstracting from physical capital here.

<sup>5</sup>It is important that we are able to use age-specific mortality data from Wrigley et al. (1997), the data not available for other countries. This is the reason for focusing on England.

is allocated to innovating. Although the positive Malthusian check, i.e. the rise in land per capita, is nearly as large, the negative impact on TFP growth due to the permanent rise in mortality counteracts the positive Malthusian check.

The 19th c. England is often associated with a number of developments that could potentially improve the speed and efficiency of knowledge transfers such as the rise in informal communication and learning networks and scientific communities, the rise in population density which would make idea exchange easier, the decline in transportation and post costs and a movement away from small family farms towards larger scale production (See Mokyr (1998) and ). In the context of our framework, these developments (1) could further reduce the negative impact of adult mortality on productivity growth over time and hence reinforce the mechanism examined here and (2) could also increase diffusion of knowledge across locations. To explore the potential effect of the latter hypothesis, we model the rate of absorption of the gap between local and frontier technology as positively dependent on population density. We find this latter channel to be quantitatively insignificant.

Finally, let's put our work in the context of existing literature on take-offs. Previous work in the field of growth has already emphasized the intimate empirical link between gains in life expectancy and increases in per capita output growth, resulting in many papers allocating a central role to gains in life expectancy in the process of income growth. An incomplete list of such papers would include Erlich and Lui (1991), Kalemli-Ozcan et al. (2000), Boucekine et al. (2002, 2003), Lagerlöf (2003), Cervellati and Sunde (2005), Soares (2005), Tamura (2006).<sup>6</sup>

Two mechanisms linking declines in mortality to increases in growth are commonly used in the literature. First, a decline in child mortality reduces parental costs of educating each surviving child, inducing greater investments in children's human capital, and in turn, leading to a take-off in output per capita and a drop in fertility through the quantity-quality trade-off. The timing of gains in adult mortality, however, is better aligned with the period of take-off. In fact, for the case of England, output per capita accelerates around 1800, while child mortality declines only late in the 19th c. This is why we focus on the causal from *adult* mortality to growth.

The second mechanism encountered in the literature is one whose logic dates back to Ben-Porath (1967); it has become the conventional wisdom regarding the causal effect from life expectancy to growth. Declines in adult mortality lead to the rise in the lifetime labor input, thus raising returns to human capital accumulation and consequently encouraging more human capital accumulation and growth. This particular link has been recently criticised by Hazan (2009) who documents that for cohorts of American men born in 1840-1970, the labor input has actually declined despite the dramatic gains in life expectancy. He further argues that a rise in lifetime labor input is a necessary implication of the Ben-Porath mechanism and concludes that gains in longevity could not have caused human capital accumulation and hence growth via the Ben-Porath mechanism. Our mechanism reconciles Hazan's findings with the possibility of causation from gains in longevity to growth. First, the direct impact of adult mortality decline on knowledge retention over time and consequently on growth is independent of lifetime labor input. Second, the indirect impact of adult mortality decline on growth is through increasing innovation time. This effect can transpire even in the presence of a decline in lifetime labor input, and hence it is not subject to Hazan's criticism. Intuitively, this can happen because falling mortality directly raises the return to human capital accumulation by improving knowledge retention over time, i.e. by raising productivity of time spent innovating. We should point out that there exist other mechanisms that can raise returns to human capital accumulation even in the presence of the declining lifetime labor input. For example, the idea that productivity of the human capital creation increases in population density, developed in Lagerlöf (2008), could

<sup>6</sup>Servellati and Sunde (2005), Lagerlöf (2003), Tamura (2006) endogenize mortality.

be also adopted to create a plausible causal link from gains in adult longevity to growth. Nonetheless, our mechanism complements the existing literature on the causal effect from declines in mortality to growth.

Because we show that high adult mortality regime can contribute to long periods of stagnation, our mechanism is also complementary to the Malthusian mechanism, the common wisdom behind long periods of stagnation. See Galor (2005) for the description of the Malthusian epoch. Recently, the effect has been subjected to quantitative analysis by Crafts and Mills (2007) and Ashraf and Galor (2008).

Our work also complements existing work on growth theory over long time scales that emphasizes the importance of technological progress (e.g. Galor and Weil (2000), Fernandez-Villaverde (2001), Hansen and Prescott (2002), Greenwood and Seshadri (2002), Bar and Leukhina (2009)). The later paper using price data estimates large productivity changes during 1600-1910 and finds that these changes were largely responsible for the economic transformation of England. This is why in this paper we focus on understanding changes in productivity and abstract from capital accumulation. Even though fertility is endogenous in our model, we abstract from the quantity-quality trade-off. The analysis of the demographic transition is outside the scope of this paper. We find it very plausible though that there exist important interactions between fertility and technological progress, especially progress biased towards skilled labor. In addition, the decline in birth rates in the late 19th c. can be attributed to factors such as compulsory education and child labor reforms (Doepke and Zilibotti 2005, Hazan and Berdugo 2006).

The rest of the paper is organized as follows. In Section II, we present the environment, including the microfoundations of the knowledge destruction rate, and equilibrium analysis. In Section III, we present our quantitative analysis. This includes estimating the time-dependent parameters of the mortality shock distribution, calibrating the model and performing experiments within the parameterized model. We conclude in Section IV.

## 2. The Benchmark Model

Time is discrete and indexed by  $t = 0, 1, 2, \dots$ . There are  $J$  locations within the country, indexed by  $j \in \{1, 2, \dots, J\}$ , each endowed with  $\Lambda_j$  acres of land, total land thus given by  $\Lambda = \sum_j \Lambda_j$ . We denote total population at location  $j$  at time  $t$  by  $N_{t,j}$ , and the total country population by  $N_t = \sum_j N_{t,j}$ . In each location, there is a representative dynasty.<sup>7</sup> Every generation lives for at most three periods, as a child, a young adult and an old adult, referenced by superscripts  $c$ ,  $y$  and  $o$ . Populations of children, young and old adults in location  $j$  at time  $t$  are denoted by  $N_{t,j}^c$ ,  $N_{t,j}^y$  and  $N_{t,j}^o$ .

### 2.1. Mortality Shocks, Population and Land Holdings Dynamics

In each period, random fractions of children and young adults in location  $j$ , denoted by  $m_{t,j}^y$  and  $m_{t,j}^c$ , do not survive to the next period. Similar to Lagerlöf (2003), we model mortality rates as a function of a log-normally distributed random variable,  $\ln(\omega_t) \sim N(\mu_t^\omega, \sigma_t^\omega)$ , drawn at the end of each period. Precisely, we assume realized mortality rates are given by

$$m_{t,j}^y = m^y(\omega_{t,j}) = \frac{\omega_{t,j}}{1 + \omega_{t,j}} \text{ and } m_{t,j}^c = \zeta_t m^y(\omega_{t,j}), \quad (1)$$

where  $\omega_{t,j}$  denotes location- and time-specific realization of mortality shock  $\omega_t$ , and  $\zeta_t$  captures the difference in trends of child and adult mortality, present in the data. Thus, the survival rate at location  $j$  is a strictly decreasing function

<sup>7</sup>Alternatively, we could reinterpret the setup as consisting of  $J$  representative dynasties within a single location.

of  $\omega_{t,j}$ . Note that the distribution parameters for  $\omega_t$  are indexed by time, and will be estimated, along with the time series for  $\zeta_t$ , to match mortality data.<sup>8</sup> While all locations face identical  $\mu_t^\omega, \sigma_t^\omega, \zeta_t$ , they will differ in mortality shock realizations.

The following relationships are implied

$$N_{t+1,j}^y = (1 - m_{t,j}^c) N_{t,j}^c, \quad (2)$$

$$N_{t+1,j}^o = (1 - m_{t,j}^y) N_{t,j}^y, \quad (3)$$

$$N_{t,j}^c = n_{t,j} N_{t,j}^y, \quad (4)$$

where  $n_{t,j}$  denotes the time- and location-specific fertility net of infant mortality and will be determined endogenously.

The above three relations imply the evolution of old and young adult populations:

$$N_{t+1,j}^y = (1 - m_{t,j}^c) n_{t,j} N_{t,j}^y, \quad (5)$$

$$N_{t+1,j}^o = (1 - m_{t,j}^y) N_{t,j}^o (1 - m_{t-1,j}^c) n_{t-1,j} / (1 - m_{t-1,j}^y). \quad (6)$$

We assume that land is jointly and equally owned by young and old adult cohorts:

$$\lambda_{t,j}^y = \frac{0.5\Lambda_j}{N_{t,j}^y}, \quad (7)$$

$$\lambda_{t,j}^o = \frac{0.5\Lambda_j}{N_{t,j}^o}. \quad (8)$$

This assumption together with (3) implies that the land holdings of the young adults surviving to old adulthood is incremented with those of non-surviving members of their cohort, thus capturing the link from high mortality realization to an increase in resources:

$$\lambda_{t+1,j}^o(\omega_{t,j}) = \frac{\lambda_{t,j}^y}{1 - m^y(\omega_{t,j})}. \quad (9)$$

## 2.2. Intergenerational Knowledge Transmission

Productivity of individuals in location  $j$  at time  $t$  is denoted by  $A_{t,j}$ . Production technology is given by  $A_{t,j} L^\theta \Lambda^{1-\theta}$ , where  $L$  is labor input,  $\Lambda$  is land input,  $\theta \in (0, 1)$ . Future productivity is a function of current productivity  $A_{t,j}$ , average time spent on improving existing ways of production by the young adult cohort  $i_{t,j}$ , increment due to a diffusion of ideas from other locations  $d_{t,j}$ , and the realization of mortality  $m_{t,j}^y$ , that is,  $A_{t+1,j} = f(A_{t,j}, i_{t,j}, d_{t,j}, m_{t,j}^y)$ . The innovation and diffusion terms are familiar determinants of future productivity, the precise form of dependence on which can be borrowed from the growth literature. However, modeling the dependence of future productivity on adult mortality requires a more thoughtful approach. We aim to capture the idea that premature death interferes with the

<sup>8</sup>Lagerlöf (2003) instead keeps the distribution parameters for  $\omega$  fixed, but allows the constant in the denominator to decline as the amount of human capital in the economy rises. In that model, only two periods of life are assumed, hence it is the mortality in the first period of life (those who do not make decisions) that is modeled. In contrast, we model both child and young adult mortality, focusing on the causal link from young adult mortality to growth.

intergenerational knowledge transmission, and hence adversely affects future productivity.

For simplicity, we separate the timing of productivity augmentation due to innovation and diffusion from the timing of the adverse impact of mortality, by assuming the augmentation takes place during the period, while mortality impacts the augmented productivity at the end of the period, i.e. upon the realization of mortality shock  $\omega_{t,j}$ . Borrowing from the growth literature, the end of period productivity, which represents future productivity in the absence of frictions in intergenerational knowledge transmission, i.e. its upper bound, is given by  $A_{t,j}^{end} = A_{t,j} + \zeta A_{t,j} i_{t,j}^\eta + d_{t,j}$ ,  $\eta \in (0, 1)$ . Note that productivity in the innovation sector depends positively on the current level of productivity, i.e. innovation is done by building on what is already known. We will discuss the form of  $d_{t,j}$  later.

The challenge is to motivate a reasonable form of future productivity dependence on  $A_{t,j}^{end}$  and  $m_{t,j}^y$  :  $A_{t+1,j} = g(A_{t,j}^{end}, m_{t,j}^y)$ . To do so, we follow Lucas (2008) and start with the premise that all ideas and techniques reside in individual heads. These embodied ideas, through complex interactions among their carriers, determine average productivity at any point in time. In what follows, we lay out the microfoundations of idea transmission across generations and productivity dependence on the stock of ideas and their carriers. The advantage of our approach, as opposed to directly assuming a particular functional form  $g(\cdot)$ , is that it allows us to be explicit about the assumptions that yield a particular form of dependence of future productivity on adult mortality.

Suppose that individual ideas known collectively at the end of the period are indexed by  $i = 1, 2, \dots, I_{t,j}$  and vary by quality  $q_{t,j}(i)$ . A person can embody any number from 0 to  $I_{t,j}$  of useful ideas. The number of young and old adult carriers of idea  $i$  are denoted by  $N_{t,j}^y(i)$  and  $N_{t,j}^o(i)$ . We assume that children, even if involved in the learning of techniques, cannot contribute to the aggregate knowledge stock prior to engaging in production as young adults. Suppose the average productivity at any point of time is a Dixit-Stiglitz aggregate of individual idea contributions  $a_{t,j}(i)$ . Note that we assume that although ideas and knowledge are embodied in private individuals, they aggregate to yield publically available (non-rival) productivity level. Thus, the end of period productivity is given by

$$A_{t,j}^{end} = \left[ \sum_{i=1}^{I_{t,j}} a_{t,j}(i)^\rho \right]^{1/\rho}.$$

The idea-specific contribution  $a_{t,j}(i)$  depends positively on its quality  $q(i)$ , the fraction of productive population that embodies this idea,  $\frac{N_{t,j}^y(i) + N_{t,j}^o(i)}{N_{t,j}^y + N_{t,j}^o}$ , and a factor reflecting the available experience associated with using this idea in production, which, for simplicity, we assume to be the fraction of the old adult carriers of idea  $i$  in total productive population. The later dependence aims to capture the idea that the older generation's expertise facilitates the application of a particular technique by its current carriers. Having the experience of the elder to draw upon is particularly important in farming economies (Rozenzweig and Wolpin, 1996). Precisely,

$$a_{t,j}(i) = q_t(i)^{1-\theta_m} \left[ \left( \frac{N_{t,j}^o(i)}{N_{t,j}^y + N_{t,j}^o} \right)^\mu \left( \frac{N_{t,j}^y(i) + N_{t,j}^o(i)}{N_{t,j}^y + N_{t,j}^o} \right)^{1-\mu} \right]^{\theta_m},$$

with  $\mu, \theta_m \in (0, 1)$ . Note the diminishing returns to experience and the number of idea carriers are the key assumptions.

By the law of large numbers, fraction  $1 - m_{t,j}^y$  of young adult carriers of idea  $i$  survive to the next period and become old adult carriers of idea  $i$  :

$$N_{t+1,j}^o(i) = \left( 1 - m_{t,j}^y \right) N_{t,j}^y(i).$$

Further suppose that the young adults that live through the entire period and survive to old adulthood (i.e. fraction  $1 - m_{t,j}^y$  of their cohort) pass all of their techniques on to their children, while those who do not survive do not pass on their knowledge. Hence, the law of motion of the young adult carriers of idea  $i$  is given by

$$N_{t+1,j}^y(i) = (1 - m_{t,j}^y) N_{t,j}^y(i) (1 - m_{t,j}^c) n_{t,j}. \quad (10)$$

To see this, note that the number of surviving young adult carriers at time  $t$  is  $(1 - m_{t,j}^y) N_{t,j}^y(i)$ , and each of those carriers can transmit this idea to  $n_{t,j}$  children,  $1 - m_{t,j}^c$  of which survive to  $t + 1$ .

The above two relations imply  $N_{t+1,j}^o(i) (1 - m_{t-1}^c) n_{t-1} = N_{t,j}^y(i)$ , and hence, the number of old adult carriers of idea  $i$  evolves according to

$$N_{t+1,j}^o(i) = (1 - m_{t,j}^y) N_{t,j}^o(i) (1 - m_{t-1}^c) n_{t-1}. \quad (11)$$

Productivity in the beginning of period  $t + 1$  is then

$$A_{t+1,j} = \left[ \sum_{i=1}^{I_{t,j}} \left[ q_t(i)^{1-\theta_m} \left( \frac{[N_{t+1,j}^o(i)]^\mu [N_{t+1,j}^y(i) + N_{t+1,j}^o(i)]^{1-\mu}}{N_{t+1,j}^y + N_{t+1,j}^o} \right)^{\theta_m} \right]^\rho \right]^{1/\rho} = \left( (1 - m_{t,j}^y)^\mu (1 - m_{t,j}^c)^{1-\mu} \right)^{\theta_m} \cdot \left[ \sum_{i=1}^{I_{t,j}} \left[ q_t(i)^{1-\theta_m} \left( \frac{[(1 - m_{t-1,j}^c) n_{t-1,j} N_{t,j}^o(i)]^\mu [(1 - m_{t,j}^c) n_t N_{t,j}^y(i) + (1 - m_{t-1,j}^c) n_{t-1,j} N_{t,j}^o(i)]^{1-\mu}}{(1 - m_{t,j}^c) n_{t,j} N_{t,j}^y + \frac{1 - m_{t,j}^y}{1 - m_{t-1,j}^y} (1 - m_{t-1,j}^c) n_{t-1,j} N_{t,j}^o} \right)^{\theta_m} \right]^\rho \right]^{\frac{1}{\rho}},$$

where we used equations (5), (6), (10) and (11).

If the mortality is the same in periods  $t - 1$  and  $t$ , and the economy is on a balanced growth path, to be discussed in Section 2.5, so that  $n_{t,j} = n_{t-1,j}$ , then the above simplifies exactly to  $A_{t+1,j} = \left( (1 - m_{t,j}^y)^\mu (1 - m_{t,j}^c)^{1-\mu} \right)^{\theta_m} A_{t,j}^{end}$ , i.e.<sup>9</sup>

$$A_{t+1,j} = (1 - m_{t,j}^y)^{\theta_m} [A_{t,j} + \zeta A_{t,j} i_{t,j}^\eta + d_{t,j}], \quad (12)$$

which implies that the fraction of potential productivity lost due to premature death interfering with intergenerational knowledge transmission is an increasing and convex function of the adult mortality rate:

$$\frac{A_t^{end} - A_{t+1}}{A_t^{end}} = 1 - (1 - m^y(\omega_{t,j}))^{\theta_m}. \quad (13)$$

There are several important messages to take. First, for reasonable micro assumptions on idea aggregation and their intergenerational transfer, we derived an adverse effect of adult mortality on the evolution of average productivity, i.e.  $(1 - m_{t,j}^y)^{\theta_m}$  is less than one and decreases in  $m_{t,j}^y$ . Second, the adverse impact of adult mortality on knowledge transmission reflects two main forces at play: (1) A premature death of the master (parent) eliminates his own contribution to future productivity and interrupts the technique transfer to their apprentices (children), eliminating their contribution as well; (2) A premature death of the parent also implies a loss of their experience to draw upon in the next period. We

<sup>9</sup>If the economy is not on a balanced growth path, but the mortality distribution parameters change slowly, the knowledge destruction function (12) serves as an approximation. Using this approximation simplifies our analysis tremendously and eliminates the need to identify  $\rho$  and  $\mu$ .

see the adverse impact decomposed into these two effects,  $(1 - m_t^y)^{(1-\mu)\theta_m}$  and  $(1 - m_t^y)^{\mu\theta_m}$ . Third, the destruction rate  $\frac{A_t^{end} - A_{t+1}}{A_t^{end}}$  is an increasing and convex function of mortality.

The above derivations explicitly reveal the assumptions that give rise to convexity of productivity destruction due to premature death. It arises due to the assumption of diminishing marginal returns to the level of experience associated with idea  $i$  and to the number of carriers of idea  $i$ . Intuitively, if the adult mortality rate is low, a young adult survivor has a very low marginal contribution to productivity, as ideas embodied by this person are embodied by many other survivors and their children, and as many other survivor carriers are available to provide their expertise. Thus, the negative impact of one death on productivity is low. However, if mortality is high, the marginal contribution of an adult survivor is large, as both the experience and the number of idea carriers are scarce in the next period. The negative impact of one death on productivity thus rises with mortality.

Note that although we model the increments to productivity due to innovation and knowledge diffusion directly by borrowing from the growth literature, our assumptions regarding these terms could be made consistent with the mathematical link from embodied knowledge to productivity that we used to derive the dependence of future productivity on mortality. For example, let's take the innovation increment. We assumed that productivity  $A$  increases to  $A + \xi A i^\eta$  provided that young adults each spend  $i$  units of time thinking how to improve the existing production process. This result could be derived through the Dixit-Stiglitz aggregator used above if the time spent innovating resulted in raising the quality of each existing idea by a factor of  $\left(1 + \xi i^\eta\right)^{\frac{1}{1-\theta_m}}$  without changing the number of ideas or the relevant fractions of idea carriers.

### 2.3. Young Adults' Problem

We assume that decisions are made prior to realizations of  $\omega_{t,j}$ . For a given mortality shock realization, a young adult may or may not survive to old adulthood. We assume that the utility of a young adult surviving to the next period is given by

$$U_{t,j}^S(\omega_{t,j}) = \frac{\left(\left(c_{t,j}^y\right)^\alpha \left(n_{t,j}\right)^{1-\alpha}\right)^{1-\sigma}}{1-\sigma} + \beta \frac{\left(c_{t+1,j}^o(\omega_{t,j})\right)^\alpha}{1-\sigma},$$

where  $c_{t,j}^y$  and  $c_{t+1,j}^o(\omega_{t,j})$  denote consumption when young and old, respectively,  $n_{t,j}$  denotes fertility net of infant mortality. Hence, the composite good consumed when young consists of the consumption good and the quantity of children, while the composite good consumed when old consists of consumption good only. Parameter  $\sigma > 0$  governs substitutability of the two composite goods. Note the explicit dependence of survivor's utility on the realization of a mortality shock: although it does not influence the decision-making at time  $t$ , it will influence future consumption through resource and productivity evolution constraints. We further assume that the utility of the young adult not surviving to old adulthood is given by  $U_{t,j}^{NS} = \frac{\left(\left(c_{t,j}^y\right)^\alpha \left(n_{t,j}\right)^{1-\alpha}\right)^{1-\sigma}}{1-\sigma}$ . The young adult's expected utility associated with a particular realization  $\omega_{t,j}$  is therefore

$$\begin{aligned} E(U_{t,j}|\omega_{t,j}) &= (1 - m^y(\omega_{t,j})) U_{t,j}^S(\omega_{t,j}) + m^y(\omega_{t,j}) U_{t,j}^{NS} = \\ &= \frac{\left(\left(c_{t,j}^y\right)^\alpha \left(n_{t,j}\right)^{1-\alpha}\right)^{1-\sigma}}{1-\sigma} + (1 - m^y(\omega_{t,j})) \beta \frac{\left(c_{t+1,j}^o\right)^{1-\sigma}}{1-\sigma}. \end{aligned}$$

Children make no decisions. When young, adults produce and consume, improve existing techniques or innovate

new ones, rear their children, and transmit their knowledge to children. When old, adults produce and consume.<sup>10</sup> We abstract from capital accumulation in order to maintain tractability and focus on the role of knowledge accumulation and its interaction with adult mortality.

The young adult cohort, taking  $A_{t,j}$ ,  $\lambda_{t,j}^y$ ,  $\bar{A}_t$  as given, chooses consumption  $c_{t,j}^y$ , time allocated to production  $l_{t,j}$ , time allocated to innovation  $i_{t,j}$ , and fertility net of infant mortality<sup>11</sup>  $n_{t,j}$  to solve

$$\max_{c_{t,j}^y, l_{t,j}, n_{t,j}} E_{\omega_{t,j}} E(U_{t,j} | \omega_{t,j}) = \frac{\left( (c_{t,j}^y)^\alpha (n_{t,j})^{1-\alpha} \right)^{1-\sigma}}{1-\sigma} + \frac{\beta}{1-\sigma} E_{\omega_{t,j}} \left[ (1 - m^y(\omega_{t,j})) (c_{t+1,j}^o(\omega_{t,j}))^\alpha \right]^{1-\sigma} \text{ s.t. (OP)}$$

$$c_{t,j}^y = A_{t,j} l_{t,j}^\theta \left( \lambda_{t,j}^y \right)^{1-\theta}, \quad (14)$$

$$A_{t+1,j}(\omega_{t,j}) = (1 - m^y(\omega_{t,j}))^{\theta_m} \left[ A_{t,j} + A_{t,j} \zeta i_{t,j}^\eta + \sigma_{A,t} (\bar{A}_t - A_{t,j}) \right], \quad (15)$$

$$l_{t,j} + q n_{t,j} + i_{t,j} = 1, \quad (16)$$

$$c_{t+1,j}^o(\omega_{t,j}) = A_{t+1,j}(\omega_{t,j}) \left( \frac{\lambda_{t,j}^y}{1 - m^y(\omega_{t,j})} \right)^{1-\theta}, \quad (17)$$

where in the last constraint we substituted for  $\lambda_{t+1,j}^o(\omega_{t,j})$  from (9) and used the optimal choice to allocate the entire unit of time to production when old. Constraint (14) states that all output produced when young is used as consumption. Constraint (15) is the law of motion for the knowledge stock (12) derived in Section 2.2, in which we further assumed, following Acemoglu (2007), that  $d_{t,j} = \sigma_{A,t} (\bar{A}_t - A_{t,j})$ , i.e. the increment to productivity due to diffusion of knowledge from other locations increases in the gap between the frontier technology available  $\bar{A}_t := \max \{A_{t,j}\}_{j \in \{1, \dots, J\}}$  and own productivity. The rate of absorption of this gap is denoted by  $\sigma_{A,t} \in (0, 1)$ ; it dictates the ease of knowledge flow across locations. It is possible that this parameter positively depends on population density and factors emphasized in Mokyr (2002), such as accessibility to communication devices (railroads, post) and the extent of availability of scientific communities. Because these factors experienced dramatic changes over the last two hundred years, we index the absorption rate by  $t$  for now and will explore its role in the development process in Section 3.3.

We model the process of transmitting the current stock of knowledge to children as the time cost  $q$ . Although it represents a fixed time amount per child, it implies that the (opportunity) cost of teaching a child/apprentice increases in the complexity of the current technology. Thus,  $q n_{t,j}$  is the time spent transmitting knowledge to the next generation. Amounts of time spent producing, teaching children and improving existing techniques must add up to the time endowment according to the time constraint (16).

### *Optimal Time Allocation*

To ease notation, we drop the dependence on  $j$  as we discuss the solution to young adults' problem. After substituting

<sup>10</sup>Note that we need at least two adult periods in order to explicitly model the interaction of adult mortality with the costly process of technology improvement.

<sup>11</sup>This is essentially an assumption that children not surviving to their first birthday are costlessly replaced with newborns.

from the constraints into the objective function, the time  $t$  problem simplifies to

$$\begin{aligned} \max_{l_t, n_t} \frac{1}{1-\sigma} & \left[ \left( A_t l_t^\theta (\lambda_t^y)^{1-\theta} \right)^\alpha n_t^{1-\alpha} \right]^{1-\sigma} + \\ \frac{\beta}{1-\sigma} E_{\omega_{t,j}} & \left[ (1 - m^y(\omega_{t,j})) \left( (1 - m^y(\omega_{t,j}))^{\theta_m} \left[ \tilde{A}_{t,j} + A_{t,j} \zeta (1 - l_{t,j} - q n_{t,j})^\eta \right] \left( \frac{\lambda_{t,j}^y}{1 - m^y(\omega_{t,j})} \right)^{1-\theta} \right)^{\alpha(1-\sigma)} \right]. \end{aligned} \quad (18)$$

where  $\tilde{A}_t = A_t + \sigma_{A,t} (\bar{A}_t - A_t)$ .

At the interior optimum, the ratio of  $n_t$  and  $l_t$  is fixed:

$$n_t = \frac{1-\alpha}{q\theta\alpha} l_t. \quad (19)$$

Using this relationship, the optimization problem of the young adult cohort can be restated as:

$$\max_{l_t} \frac{\left[ l_t^{(\alpha\theta+1-\alpha)} \right]^{1-\sigma}}{1-\sigma} + B_t \frac{\left[ 1 + \sigma_{A,t} (\bar{A}_t - A_t) / A_t + \zeta \left( 1 - \left( 1 + \frac{1-\alpha}{\theta\alpha} \right) l_t \right)^\eta \right]^{\alpha(1-\sigma)}}{1-\sigma}, \quad (20)$$

where  $B_t = \beta E_{\omega_t} \left[ (1 - m^y(\omega_t)) (1 - m^y(\omega_t))^{\theta_m \alpha(1-\sigma)} (1 - m^y(\omega_t))^{\alpha(\theta-1)(1-\sigma)} \right] / \left( \frac{1-\alpha}{q\theta\alpha} \right)^{(1-\alpha)(1-\sigma)}$  is the relative weight on future utility. This formulation is convenient for discussing the impact of changes in mortality on the intratemporal time allocation. Choosing a higher  $l_t$  (and hence spending less time on innovation) raises the amount of the composite good today and hence raises utility when young. Conversely, choosing a lower  $l_t$ , i.e. spending more time on innovation, raises utility when old.

Because labor is supplied inelastically in the last period of life, all decisions of the time  $t$  cohort are independent of the time  $t$  realization of the mortality shock. So the only way the time  $t$  parameters of the mortality shock distribution affect time  $t$  decision making is through the expectation term  $E_{\omega_t} [\cdot]$  that enters  $B_t$ . The time  $t-1$  realization of mortality, however, does affect the state variables  $A_t$  and  $\bar{A}_t$  and hence the time  $t$  decision making. Note that the state variables  $\lambda_t^y$  and  $N_t^y$  have no bearing on decision making.

Note that the expectation appearing inside  $B_t$  is taken over the product of three factors. Formulation (18) should remind of their origin. The first,  $D(\omega_t) := (1 - m^y(\omega_t))$ , represents the discount factor on future consumption due to uncertainty of lifespan. As mortality declines and people adjust their expectations accordingly, the weight on utility when old increases, which encourages more innovation. This factor can be also interpreted as the standard Ben-Porath effect of life extension on human capital accumulation, in our formulation appearing independent of the elasticity of substitution between current and future consumption. The second factor is the knowledge transmission factor  $K(\omega_t) := (1 - m^y(\omega_t))^{\theta_m \alpha(1-\sigma)}$ , it is a function of the rate of productivity retention across time. Declining mortality increases the term  $(1 - m^y(\omega_t))^{\theta_m}$ , which works to increase returns to innovation, as the amount of total knowledge transmitted to the next period increases. The impact of declining mortality on time allocation through the knowledge channel depends on the substitutability of the two composite goods. If the goods are substitutes, an increase in returns to innovation leads to labor reallocation towards innovation, and substitution of future for present consumption. However, if the goods are complements, an increase in returns to innovation raises the demand for both present and future consumption, and consequently induces labor reallocation away from future consumption, the sector experiencing productivity gains.

The last factor is the resource dilution channel  $M(\omega_t) := (1 - m^y(\omega_t))^{\alpha(\theta-1)(1-\sigma)}$ , appearing due to the assumed land redistribution among the surviving cohort members, and branded Malthusian because of the active role that resources per capita played in driving the Malthusian population dynamics (see Malthus, 1798). As mortality declines, people adjust to expect less augmentation of their land holding over the lifecycle in case of survival, which lower returns to innovation and, if the composite goods are substitutes, encourages labor reallocation towards having more children and producing consumption good when young.

The first order condition to the optimization problem in (20) is given by

$$\theta l_t^{((1-\alpha)+\alpha\theta)(1-\sigma)-1} = B_t \frac{\xi \eta \left(1 - \left(1 + \frac{(1-\alpha)}{\theta\alpha}\right) l_t\right)^{\eta-1}}{\left[\tilde{A}_t/A_t + \xi \left(1 - \left(1 + \frac{(1-\alpha)}{\theta\alpha}\right) l_t\right)^\eta\right]^{1-\alpha(1-\sigma)}}. \quad (21)$$

The first term gives the marginal benefit from increasing  $l_t$ , i.e. marginal utility from the resulting increase in consumption when young and fertility. The second term is the marginal cost of doing so, i.e. the loss, in terms of utility, of knowledge tomorrow, i.e. the loss in own productivity and productivity of children, associated with the decline in the time spent innovating.

## 2.4. Equilibrium Dynamics

**Definition 1** For given parameter values, sequence  $\{\zeta_t\}$  and time-dependent parameters of the mortality shock distribution  $\{\mu_t^\omega, \sigma_t^\omega\}$ , initial conditions  $\{N_{0,j}^y, N_{0,j}^o, A_{0,j}\}_{j=1,\dots,J}$  and a history of mortality shock realizations  $\{\omega_{t,j}\}_{t=0, j \in \{1,\dots,J\}}^\infty$  that imply a history of mortality rates according to (1), an equilibrium consists of sequences  $\{N_{t+1,j}^y, N_{t+1,j}^o, A_{t+1,j}, \lambda_{t+1,j}^y, \lambda_{t+1,j}^o, c_{t,j}^y, c_{t,j}^o, l_{t,j}, n_{t,j}, i_{t,j}\}_{t=0, j \in \{1,\dots,J\}}^\infty$  such that  $\forall t, j$ , the young adult's problem (OP) is solved and the laws of motion (2), (3), (7) and (15) hold.

**Proposition 1** If  $\sigma > 0$ ,  $\alpha, \eta \in (0, 1)$ , then there exists a unique equilibrium. Moreover, the equilibrium time paths must satisfy conditions (19) and (21), i.e. the unique equilibrium is interior.

**Proof.** See the appendix. ■

The economy starts off with the initial conditions  $\{N_{0,j}^y, N_{0,j}^o, A_{0,j}\}_{j=1,\dots,J}$ . Next, we will show how given the state variables in period  $t$ , the remaining variables in period  $t$  and the state variables in period  $t + 1$  are determined. Given  $\{N_{t,j}^y, N_{t,j}^o, A_{t,j}\}_{j=1,\dots,J}$ , equations (7) and (8) give  $\lambda_{0,j}^y, \lambda_{0,j}^o$ . Equation (17) gives consumption of the current old adults  $\{c_{t,j}^o\}_{j=1,\dots,J}$ . The optimization problem (OP) is solved for  $c_{t,j}^y, l_{t,j}, n_{t,j}$  in each location. Then (4) pins down  $N_{t,j}^c$ . The mortality shocks  $\{\omega_{t,j}\}_{j=1,\dots,J}$  are realized, so location-specific mortality rates  $m^c(\omega_{t,j})$  and  $m^y(\omega_{t,j})$  are determined for each  $j$  according to (1). Given the realizations  $\{\omega_{t,j}\}_{j=1,\dots,J}$ , equations (2), (3) and (15) determine the state variables  $\{N_{t+1,j}^y, N_{t+1,j}^o$  and  $A_{t+1,j}\}_{j=1,\dots,J}$ .

## 2.5. Balanced growth

Consider a version of our model economy in which  $\zeta_t, \mu_t^\omega, \sigma_t^\omega$  are fixed, and mortality shocks are identical across time and locations and equal their expectation value:  $m^y = E_\omega m^y(\omega)$ . To ease notation, we drop the superscript, so that  $m^y = m$  and  $m^c = \zeta m$ . Assuming further that locations are identical in their initial conditions, there will be no diffusion of knowledge across locations, and locations will remain identical. It suffices to drop the subscript  $j$ . It is

useful to discuss a balanced growth equilibrium of this deterministic economy, as it represents the trend for the behavior of the model with uncertainty, in which parameters of the mortality shock distribution are held fixed.<sup>12</sup>

In what follows, we will also assume that our model under the assumption of fixed  $\zeta_t$  and fixed distribution of mortality shock parameters, i.e. our model exhibiting fluctuations around a balanced growth path equilibrium, can generate behavior consistent with the behavior of the English economy in the beginning of the 17th century. This assumption will be used to identify the model parameter values that can capture the trends in the 17th century England.

**Proposition 2** Denote the solution to equations (22) – (24) by  $n^{BG}, l^{BG}, \gamma_A^{BG}$  :

$$\gamma_A = (1 - m)^{\theta m} [1 + \zeta (1 - l - qn)^\eta], \quad (22)$$

$$\left(\frac{1 - \alpha}{q\theta\alpha}\right)^{(1-\alpha)(1-\sigma)} \theta l^{((1-\alpha)+\alpha\theta)(1-\sigma)-1} = \frac{\beta E \eta \zeta (1 - l - qn)^{\eta-1}}{[1 + \zeta (1 - l - qn)^\eta]^{1-\alpha(1-\sigma)}}, \quad (23)$$

$$n = \frac{(1 - \alpha)l}{q\theta\alpha}, \quad (24)$$

where  $E = E_{\omega_t} [(1 - m^y(\omega_t)) (1 - m^y(\omega_t))^{\theta m \alpha (1-\sigma)} (1 - m^y(\omega_t))^{\alpha(\theta-1)(1-\sigma)}] = E_{\omega_t} [(1 - m^y(\omega_t))^{1+\alpha(\theta m + \theta - 1)(1-\sigma)}]$ .

Assume an arbitrary number ( $J$ ) of locations. Suppose the initial conditions in each location are identical and satisfy  $\frac{N_0^c}{N_0^y} = n^{BG}$ ,  $\frac{N_0^o}{N_0^y} = \frac{(1-m)}{(1-\zeta m)n^{BG}}$ ,  $\lambda_0^y = \frac{0.5\Lambda}{N_0^y}$ ,  $\lambda_0^o = \frac{0.5\Lambda}{N_0^o}$ . Then the equilibrium paths exhibit balanced growth behavior from period 0 and onward. In particular,  $n_t = n^{BG}$  and  $l_t = l^{BG} \forall t$ , and  $A_t$  grows at the rate  $\gamma_A^{BG}$ . Moreover, the growth rates of all population subgroups, output per capita, and land holdings are then given by

$$\gamma_{pop}^{BG} = \frac{N_{t+1}^y}{N_t^y} = \frac{N_{t+1}^o}{N_t^o} = \frac{N_{t+1}^c}{N_t^c} = (1 - \zeta m) n^{BG}, \quad (25)$$

$$\gamma_y^{BG} = \gamma_A^{BG} \left(\frac{1}{\gamma_{pop}^{BG}}\right)^{1-\theta}, \quad (26)$$

$$\gamma_{\lambda^y}^{BG} = \gamma_{\lambda^o}^{BG} = \frac{1}{\gamma_{pop}^{BG}}. \quad (27)$$

**Proof.** We want to show that the balanced growth equilibrium described in Proposition 2, i.e.  $\{A_t = (\gamma_A^{BG})^t A_0, N_t^y = (n^{BG} (1 - \zeta m))^t N_0^y, N_t^c = (n^{BG} (1 - \zeta m))^t N_0^c, N_t^o = (n^{BG} (1 - \zeta m))^t N_0^o, l_t = l^{BG}, n_t = n^{BG}, \lambda_t^y = \lambda_0^y / (n^{BG} (1 - \zeta m))^t, \lambda_t^o = \lambda_0^o / (n^{BG} (1 - \zeta m))^t\}$ , indeed comprises the equilibrium of the model economy with the parametric restrictions and initial conditions specified in the supposition. We proceed by showing that the candidate solution satisfies all conditions sufficient for the equilibrium. First, the candidate solution satisfies (24), hence it satisfies the equilibrium relationship (19). The candidate solution also satisfies  $\frac{N_0^c}{N_0^y} = n^{BG}$ ,  $\frac{N_0^o}{N_0^y} = \frac{(1-m)}{(1-\zeta m)n^{BG}}$  and (25) and hence it satisfies the equilibrium relationships (4) – (6). The candidate solution satisfies  $\lambda_0^y = \frac{0.5\Lambda}{N_0^y}$ ,  $\lambda_0^o = \frac{0.5\Lambda}{N_0^o}$  and (27), and hence it satisfies the equilibrium law of motion for land holdings (7) and (8). Because the candidate solution satisfies (22) and because all locations are ex-ante and ex-post identical along the candidate path, i.e.  $\bar{A}_t = A_{t,j}$  for all  $t$  and  $j$ , the candidate solution also satisfies the equilibrium relationship (15). Finally, since the candidate solution satisfies (23) and (24), it also satisfies the first order condition (21). Hence, we showed that the candidate balanced

<sup>12</sup>Adding mortality shocks to this deterministic model would amount to adding shocks to the stable dynamical system characterizing it.

growth solution satisfies all of the sufficient conditions for the equilibrium. Since, by Proposition 1, there is a unique equilibrium solution, the candidate balanced growth solution is the equilibrium solution. ■

As seen from (22), we obtain that the rate of productivity growth positively depends on the rate of knowledge transmission across generations,  $(1 - m^y)^{\theta_m}$ , and time spent on innovation. Relation (26) elucidates that productivity growth in turn contributes directly to the growth of output per capita, while population growth slows it down, which simply captures the negative impact of dilution of the non-reproducible factor. Population growth is given by the child survival rate multiplied by fertility net of infant mortality.

### 3. Quantitative Analysis

#### 3.1. Estimating evolution of the parameters of the mortality shock distribution

Because one of the questions we ask is how changes in frequency and severity of epidemics affected the process of development over time, we need to estimate changes in the parameters of the distribution of  $\omega_{j,t}$ . For the case of England, we gather the data on age-specific mortality for 10 year intervals: 1640-49,...1990-2000. The data up to 1800 is taken from Wrigley et al., while the data for later years is available through the human mortality database. We take a stand that each period in the model lasts 25 years. What we need is mortality rate for the age groups 1-25 and 25-50. Recall that the mortality at the end of the last period (50-75) is 100% by the structure of the model. Also note that even though Crude Death Rate data is available with higher frequency (5 years), it is not informative for the purpose at hand as we lack population structure data for the earlier years.

From (1), we solve for  $\omega_{t,j}$  obtaining  $\omega_{t,j} = \frac{m_{t,j}^y}{(1-m_{t,j})}$ , i.e.,  $\ln \omega_{t,j} - \ln 1 = \ln \left( \frac{m_{t,j}^y}{1-m_{t,j}} \right)$ . We denote observations of mortality 25-50 by  $m_t$ . The data reveals a declining time series of  $\ln \left( \frac{m_t}{1-m_t} \right)$ , which also becomes smoother over time. To introduce smooth behavior of parameters over time, we model  $\mu_t$  and  $\sigma_t$  as logistic functions of time and estimate parameters of these functions to maximize the log-likelihood of the observed sequence of adult mortality rates. We discuss the details of the estimation in the appendix. Figure 1 plots the resulting time series of  $\{\mu_t, \sigma_t\}$ .

Our estimate of  $\{\zeta_t\}_{t=1600}^{2000}$  captures the difference in trends of mortality of the age groups 1-25 and 25-50. We estimate  $\zeta_t$  to start at 0.65 in 1600, then during 1640-1865 increase to .9 and decrease linearly to 0.25 in 1990. This fraction grows initially because the mortality rate of the age group 1-25 is roughly fixed in the data, while adult mortality declines. It eventually decreases because in this period child mortality falls more than adult mortality.

#### 3.2. Calibration of the Benchmark Model

One goal of this study is to quantitatively assess the importance of the causal link proposed between mortality and growth from the historical perspective. In order to accomplish this goal, we parameterize the model to match several key features of preindustrial England assuming that the behavior of the 17th c. England can be accurately captured by balanced growth behavior of our model economy. We thus employ equations from Proposition 2, where recall  $m^y = E_{\omega_t} m^y(\omega_t)$  is the mortality shock identical across time and locations. Since locations are identical ex-ante and because they are hit with the identical shock, there will be no knowledge flows across locations, and locations will remain identical. It suffices to ignore the subscript  $j$  in the discussion of calibration. Also note that as will be discussed in Section 3.2.1 below, as long as all the locations face an identical distribution of mortality shocks, even if they are hit

with different mortality shocks, convergence tends to be towards balanced growth with no knowledge diffusion. Hence, we choose to assume balanced growth with no diffusion of knowledge for the purpose of calibration.

From the estimation of the mortality process we have  $m^y = 0.3817$  and  $\zeta = 0.65$ . Clark (2001b) reports that the land share in total income during 1600-1700 is around 0.3, so we set  $\theta = 0.7$ .

We set the length of each period to 25 years. From Clark (2001b), output per capita in the 17th century rose at about 0.2% a year, which implies the growth rate of 1.053 per a 25-year period. From Wriggley et al. (1997), population growth rate in the 17th c. England is roughly 0.25% per year, or 6.5% per 25 years.

Then (25) and (26) imply the model must predict fertility and productivity growth along the calibrated balanced growth path to be  $n^{BG} = 1.453$  and  $\gamma_A^{BG} = 1.073$ .

Wriggley et al. (1997) documents the total population size around 1600 to be 4, 109, 981. Hence,  $N_0 = 4, 109, 981/J$  per location. Balanced growth implies a constant age structure, i.e.  $N_0^c/N_0^y = n^{BG}$  and  $N_0^o/N_0^y = \frac{1-m^y}{(1-\zeta m^y)n^{BG}}$ . Since we already determined the mortality and fertility rates, we can determine population structure along the calibrated BGP. We must set  $N_0^y = 4, 109, 981/J / (1 + n^{BG} + \frac{1-m^y}{(1-\zeta m^y)n^{BG}}) = 4, 109, 981/J / (1 + 1.453 + \frac{1-0.3817}{(1-0.7(0.3817))1.453}) = 1354690/J$ , which implies 33% of the total population are of ages 25-50. We also have  $N_0^o = 786, 438/J$  per location.

Next we calibrate  $\theta_m$ , an important parameter that governs the rate of knowledge destruction due to mortality. To identify this parameter, we use the documented response of real agricultural wages to the Black Death epidemic in England. Along the calibrated balanced growth, realizations of mortality shocks equal their expected value every period, and both average and marginal products of labor of the young and the old grow at the rate of  $(A_{t+1}\lambda_{t+1}^{1-\theta}) / (A_t\lambda_t^{1-\theta}) = \gamma_A^{BG} ((1 - \zeta m^y) n^{BG})^{\theta-1} = 1.053$ . However, if there is a surprise mortality shock larger than its expected value, then the average and marginal products of labor will rise due to the growth in per capita land holdings for the survivors. Our knowledge destruction hypothesis implies that wages should rise less than what is implied by the simple arithmetic of recalculating the marginal product in case of complete retention of knowledge. How much less depends on  $\theta_m$ . So suppose a period  $t^*$  mortality rate is larger than its expected value and it affects equally children and young adults. Because the shock is unexpected and occurs at the end of the period, the choice variables in period  $t^*$  are still BGP choices. Clark (2001a) estimates that over the 25 year period following the Black Death, population of England shrunk by approximately 45%. We need to find the (unexpected) mortality rate that would imply a comparable destruction of population for our model economy. Denoting by  $t^*$  the time of the mortality shock, we find it by imposing that

$$\frac{N_{t^*+1}}{N_{t^*}} = \frac{(1 - m_{t^*}^y) n_{t^*+1} n^{BG} N_{t^*}^y + (1 - m_{t^*}^y) n^{BG} N_{t^*}^y + (1 - m_{t^*}^y) N_{t^*}^y}{n^{BG} N_{t^*}^y + N_{t^*}^y + (1 - m^y) N_{t^*}^y / (1 - \zeta m^y) n^{BG}} = 0.55,$$

where we used the population age structure implied by balanced growth. Since we assume the shock does not affect expectations and consider no diffusion, we have  $n_{t^*+1} = n^{BG}$  and hence the above equality then implies  $m_{t^*} = 0.63$ . In response to this shock, our model implies the following rise in the marginal (or average) product of labor of young adults:

$$\begin{aligned} \frac{A_{t+1}\lambda_{t+1}^{y1-\theta}}{A_t\lambda_t^{y1-\theta}} &= \left( \frac{1 - m_{t^*}^y}{1 - m^y} \right)^{\theta_m} \gamma_A \left( \frac{N_{t+1}^y}{N_t^y} \right)^{\theta-1} = \\ &= \left( \frac{1 - m_{t^*}^y}{1 - m^y} \right)^{\theta_m} \gamma_A ((1 - m_{t^*}^y) n^{BG})^{\theta-1}, \end{aligned} \quad (28)$$

where we used (22) and the fact that the optimal decisions at  $t^*$  are the same as on the original BGP. Clark (2001a)

presents data on real agricultural wages during the 14th c. revealing that 25 years following the Black Death, real wages were 20% higher. Setting (28) to 1.2 gives  $\theta_m = 0.15$ .

Note that this procedure of identifying  $\theta_m$  implies a very low rate of knowledge destruction of 6.88% over the period of 25 years. The smaller this value is, the smaller will be the effect of gains in life expectancy.<sup>13</sup> What we aim to identify is the influence (if any) of these quantitatively small gains in knowledge retention on the course of development.

Some of the moments predicted by our model of human capital accumulation, in particular, time spent raising children  $qn$  and time spent innovating  $i$  are difficult to map to and measure in the data. What can be shown numerically, however, is that choosing lower values for these two moments tends to reduce human capital accumulation in response to declining mortality. So, we choose to go with the low values  $i = 0.05$  and  $qn = 0.05$ . Clearly, this means the model must imply  $l^{BG} = 1 - qn - i = 0.9$ . Since we already determined  $n^{BG}$ , we obtain  $q = qn/n^{BG} = 0.035$ .

We set  $\eta = 0.7$ , which seems to be the middle of the range of estimates of labor share in human capital accumulation production functions. Then equation (22) implies

$$\xi = \left( \frac{\gamma_A}{(1 - m^y)^{\theta_m}} - 1 \right) / i^\eta,$$

i.e.  $\xi = 1.245$ .

Since equation (24) gives  $l = \frac{q\theta\alpha}{(1-\alpha)}n$ , we obtain the relationship between  $\alpha$  and  $(1 - \alpha)$

$$\begin{aligned} (1 - \alpha) &= \frac{qn\theta}{l}\alpha, \\ \alpha &= 1 / \left( 1 + \frac{qn\theta}{l} \right) = 1 / \left( 1 + \frac{.05(.7)}{.9} \right) = .96. \end{aligned}$$

We fix  $\beta = \left( \frac{1}{1.02} \right)^{25} = 0.61$ .

We have one equation left

$$\left( \frac{1 - \alpha}{q\theta\alpha} \right)^{(1-\alpha)(1-\sigma)} \theta l^{(1+\alpha(\theta-1))(1-\sigma)-1} = \frac{\beta E \eta \xi i^{\eta-1}}{[1 + \xi i^\eta]^{1-\alpha(1-\sigma)}}, \quad (29)$$

where  $E = E_{\omega_t} \left[ (1 - m^y (\omega_t))^{1+\alpha(\theta_m+\theta-1)(1-\sigma)} \right]$ . Note that  $E$  depends on  $\sigma$  and it must be computed numerically. The solution to this remaining BGP equation (23), which has to be found numerically, gives  $\sigma$ . We obtain the unique solution to be  $\sigma = 0.58$ .

Macdonald (1979) reports that people were willing to pay as much as 60 pounds to spend a season at a more progressive farm learning a better technique of production. Depending on the estimate of money wages around 1800 (Clark 2001 or Mitchell 1989), this would translate into 1.5-2 annual farm wages. Our calibration implies that relative to the corner solution with no innovation, to move one year of labor ( $1/25=0.04$ ) into the innovation sector, a young adult would be willing to give up around 1.7 times the marginal product of labor at the optimum. This instills further

<sup>13</sup>Our theory of knowledge aggregation implies that  $\delta(m, \theta_m) = 1 - (1 - m)^{\theta_m}$ . Then  $\delta_{\theta_m}(m, \theta_m) = -(1 - m)^{\theta_m} \ln(1 - m) > 0$ . So we have to minimize  $\theta_m$  in order to minimize the destruction rate. What about minimizing the contribution of the fall in mortality to the rise in knowledge retention? First, note that indeed the fall in mortality contributes to the rise in the knowledge retention because  $\delta_m(m, \theta_m) = \theta_m (1 - m)^{\theta_m-1} > 0$ . Let's consider  $\delta_{m\theta_m}(m, \theta_m) = (1 - m)^{\theta_m-1} + \theta_m (1 - m)^{\theta_m-1} \ln(1 - m) = (1 + \theta_m \ln(1 - m)) (1 - m)^{\theta_m-1}$ . It is negative for a large enough  $m$ . For  $\theta_m = .5$ , for example,  $\delta_{m\theta_m}(m, \theta_m) < 0$  for all  $m > .86466$ . So, for all  $m$  that we work with, we have  $\delta_{m\theta_m}(m, \theta_m) > 0$ , meaning that minimizing  $\theta_m$  implies minimizing the contribution of changes in  $m$  to increasing knowledge retention.

confidence in our choices of  $i$  and  $\eta$ .

Note that  $1/\sigma$  is the elasticity of substitution between the composite good when young and the composite good when old. The implied elasticity of intertemporal substitution between consumption when young and old is actually  $\frac{1}{1-\alpha(1-\sigma)} = 1.67$ , i.e. it is smaller than  $\frac{1}{\sigma}$ . In any case, 1.67 certainly falls in the range of micro estimates of intertemporal elasticity of substitution (IES). In particular, Gruber (2008) gives a strong case for the IES to be greater than 1. He claims that the large literature that estimates IES has produced very mixed results, and most estimates suffer from the endogeneity problem. He uses an exogenous variation in tax rates across individuals to identify  $IES = 2$ , which is even higher than our 1.67. Finally, existing estimates are based on an Euler equation from the neoclassical growth model, which is different from the one in our model, which would also include fertility.

Also note that it is standard in Ben-Porath type models to maximize income over lifetime, i.e. assume infinity  $IES = \infty$ , which is a far more extreme assumption than 1.67.

Finally, we normalize  $\Lambda = 10000$  and  $A_{0,j} = 1$ .

Table 1: Calibrated Parameter Values

Consumption Technology:	$A_0 = 1, \theta = 0.70$
Idea Production Technology:	$\Lambda = 1, \eta = 0.7, \theta_m = 0.15$
Preferences:	$\alpha = .96, \sigma = 0.58$
Time Cost of Children:	$q = 0.035$
Mortality	$m^y = 0.3817, \zeta = 0.65$

### 3.2.1. Simulating Preindustrial Economy with Knowledge Diffusion

Before testing the quantitative importance of our hypothesis, we simulate the economy calibrated for the case of no diffusion. With simulated mortality shocks, locations are no longer identical and hence knowledge diffusion will take place. The question is whether or not the 17th c. trends can still be captured by this simulation.

First, we need to understand the role of knowledge diffusion. To understand the effect of introducing diffusion of knowledge, consider 2 identical locations that are on a BGP described above, except that one of them is hit with a higher than expected mortality rate in period  $t^*$ . Formally, the two locations face identical  $t^*$  initial conditions from Proposition 2 and identical stationary distribution of mortality  $\{\mu^\omega, \sigma^\omega\}$  and  $\zeta$  in each period, but sequences of mortality rate realizations are given by  $\{m_{t,1}^y\} = \{m^y, m^y, \dots, m^y\}$  for the first location, where  $m^y = E_\omega m^y(\omega)$ , and by  $\{m_{t,2}^y\} = \{m^y, m^y, \dots, (1 + \varepsilon)m^y, \dots, m^y\}$  for the second location, where  $\varepsilon > 0$  and  $(1 + \varepsilon)m^y$  is the mortality rate realized at time  $t^*$ .

For  $t = 0, 1, \dots, t^*$ , the two locations experience balanced growth characterized in Proposition 2 and innovation time given by  $i^{BG} = 1 - l^{BG} - qn^{BG}$ . Productivity at time  $t^*$  is also identical at the level of  $\gamma_A^{t^*} A_0$ . At time  $t^*$ , however, location 2 passes less knowledge to the next period and consequently becomes less productive:  $A_{t^*+1,2} = (1 - (1 + \varepsilon)m^y)^{\theta_m} [\gamma_A^{t^*} A_0 + \gamma_A^{t^*} A_0 \zeta (i^{BG})^\eta + \sigma_A(0)]$  is less than  $A_{t^*+1,1} = (1 - m^y)^{\theta_m} [\gamma_A^{t^*} A_0 + \gamma_A^{t^*} A_0 \zeta (i^{BG})^\eta + \sigma_A(0)] = \gamma_A^{t^*+1} A_0$ . The time  $t^* + 1$  gap between the frontier productivity (of location 1) and productivity of location 2 is  $A_{t^*+1,1} - A_{t^*+1,2} = \gamma_A^{t^*+1} A_0 \left(1 - \left(\frac{1 - (1 + \varepsilon)m^y}{1 - m^y}\right)^{\theta_m}\right) > 0$ . Fraction  $\sigma_A$  of this gap will be absorbed by location 2 in period  $t^* + 1$ . The significance of this absorption is that it mitigates the negative impact of the high

mortality episode. At the extreme end, assuming  $\sigma_A = 1$ , there is complete absorption of time  $t^* + 1$  frontier technology by location 2. However, even in the case of  $\sigma_A = 1$ , productivity in the second location will be lower in period  $t^* + 2$  due to the discouraging impact of knowledge diffusion on innovation. The intuition can be obtained from formulation (20) of the young adult cohort time allocation problem. The presence of diffusion increases productivity of labor when old and hence lowers the marginal utility of future consumption consequently causing time to reallocate towards current production and fertility. Note that the effect of diffusion on innovation is always negative, irrespective of substitutability of present and future composite goods. Hence,  $i_{t^*+1,2} < i^{BG}$ , while  $i_{t^*+1,1} = i^{BG}$ , and  $A_{t^*+2,1} = (1 - m^y)^{\theta_m} \left[ \gamma_A^{t^*+1} A_0 + \gamma_A^{t^*+1} A_0 \zeta (i^{BG})^\eta \right] = \gamma_A^{t^*+2} A_0$  and  $A_{t^*+2,2} = (1 - m^y)^{\theta_m} [A_{t^*+1,2} + A_{t^*+1,2} \zeta (i_{t^*+1,2})^\eta + 1 (A_{t^*+1,1} - A_{t^*+1,2})] = (1 - m^y)^{\theta_m} \left[ \gamma_A^{t^*+1} A_0 + A_{t^*+1,2} A_0 \zeta (i_{t^*+1,2})^\eta \right]$ . While the first term in the brackets is the same, i.e. the presence of diffusion gives a boost to productivity of the follower location, the second term is lower for the follower location and prevents the gap from closing immediately. As the gap shrinks, however, time spent innovating converges back to its balanced growth level, productivity equalizes, and both locations exhibit balanced growth behavior of an economy with  $\sigma_A = 0$ . Figure 2 reports the results of this experiment for  $\sigma_A \in \{0.3, 1\}$  and the calibration described above. As intuition suggests, convergence time declines in the level of  $\sigma_A$ .

The significance of introducing diffusion is twofold. First of all, it moderates the negative impact of a high mortality episode on knowledge transmission across generations, i.e. working against the hypothesis we aim to test here. In particular, it makes the negative impact of high mortality episodes only short-lived. Second, recall that fertility increases in the level of knowledge inflow from other locations. Because the level of knowledge inflow is positively associated with increases in resources per capita in our model (both increase following a high mortality shock), the presence of diffusion also creates a positive correlation of per capita resources and fertility in the model, one link that played an important role in population dynamics of preindustrial societies.

To simulate the economy, we need to choose  $J$  and  $\sigma_A$ . We find that the volatility of the growth rate of total population in the country is sensitive to the number of locations assumed, decreasing with the number of locations. To match the coefficient of variation of population growth of 0.048 around 1600, we choose  $J = 5$ . We set the absorption rate to a very low level of 0.3, which allows to have a low level of fertility variation across locations, documented for the preindustrial period.

Figure 3 reports the simulated paths for the 5 locations and their empirical counterparts. For data sources, refer to the appendix. Panel *c* reports the total country population, rather than location-specific population, in order to facilitate comparison to the population time series in the data. In panel *d*, we report the change in  $B_t$ , the relative weight on the second term in the objective function of the simplified maximization problem (formulation (20)), which changes only due to changes in the expectation term  $E_{\omega_t} \left[ (1 - m^y(\omega_t)) (1 - m^y(\omega_t))^{\theta_m \alpha (1 - \sigma)} (1 - m^y(\omega_t))^{\alpha (\theta - 1) (1 - \sigma)} \right]$ . Because in this simulation we keep the parameters of the mortality distribution fixed at their 1600 values,  $B_t$  is constant.

It is clear that the model with diffusion and differential mortality shocks across locations closely captures the empirical trends in 17th c. England, aimed to be captured by our calibration procedure. Panels *a* and *c* reveal that output per capita and population size closely follow their empirical counterparts for the 17th c. From panel *c*, we see that CBR net of infant mortality in the model fluctuates around a constant value.<sup>14</sup>

What we see is that on average, time spent on innovation is slightly below the value we used in the calibration. At a

<sup>14</sup>To facilitate comparison to the model, in which we assume  $n$  stands for net fertility, the empirical series for CBR is adjusted by infant mortality according to  $(1 - IMR/1000) CBR$ . CBR net of infant mortality in the model is computed as  $\frac{1000}{25} n N^y / N$ .

given point in time, only the frontier productivity location will innovate at that level. The rest of the locations experience a positive inflow of knowledge from the frontier locations, which prompts shifting of resources into production of output and children and away from innovating. Nonetheless, the 17th c. trends of output per capita and population growth are captured well by the calibrated model with diffusion.

### 3.3. Results

To understand our main results, it is useful to point out that there is a stark difference between the effect of a temporary shock and a permanent mortality shock in our model. For clarity of this discussion, assume there is no diffusion. Assume three locations. Location 1 is hit with expected mortality,  $m^y = E_\omega m^y(\omega)$ , i.e. it exhibits trends we aimed at by the calibration, it is reported just to have a benchmark. Location 2 differs from Location 1 in that in period 5 it is hit with a larger by the factor 1.1 than expected mortality shock. The shock is temporary and unexpected. Location 3 is hit with the same shock, but it is permanent and unexpected. With this experiment, we attempt to understand the difference between shocks along the simulated balanced growth path and permanent changes in mortality that we introduce in the main experiment. The results are reported in Figure 4. Panel *b* reports the growth rate of  $y_t^y$  and the remaining panels report the components of this growth rate, i.e. the growth rates of  $A_t$ ,  $l_t^\theta$  and  $(\lambda_t^y)^{1-\theta}$ . With a temporary and unexpected rise in mortality, time allocation is not affected, but the growth rate of per capita land holdings temporarily increases because the fixed resource is now shared among fewer survivors, the so called positive Malthusian check. Even though the growth rate of productivity temporarily declines, it is not enough to reverse the positive Malthusian check, consequently, the output per capita growth temporarily rises. So, a temporary high mortality shocks that often hit the simulated economy from Figure 3 were good for the survivors.

In light of this result, it seems surprising that permanent rises in mortality would introduce a negative effect on the economy (and permanent declines would result in improvement of living standards). The key difference in the result is driven by the difference in the effect that temporary and permanent changes in mortality introduce on time allocation. A permanent and expected rise in mortality initially (before the shock is realized) also has a positive effect on output per capita as, in anticipation of low retention of knowledge and low probability of own survival to the next period, labor moves into production of output and away from innovation. As the first high mortality shock hits, productivity growth suffers more relative to the case of the temporary rise in mortality, because in addition to the decline in knowledge retention  $(1 - m_5^y)^{\theta m}$ , less time is allocated to innovation, thus the decline in  $(1 + \xi i_5^\eta)$  also contributing to the drop in TFP growth (See equation (12)). Although the positive Malthusian check is nearly as large (the rise in  $(\lambda^y)^{1-\theta}$  is slightly smaller because of higher fertility in period 5), the negative impact on TFP growth due to the permanent rise in mortality is now large enough to reverse the positive Malthusian check.

This discussion should help understand why permanent declines in mortality will contribute to permanent increases in output per capita growth, as we will show next.

#### 3.3.1. Main Experiment: Overall Effect of Mortality Decline on Takeoff

In the main experiment, we simulate the model subject to exogenously changing parameters of the mortality process according to our estimation from Section 3.1, that is, we allow  $\{\mu_t^\omega, \sigma_t^\omega, \zeta_t\}$  change according to our estimation and track the magnitude of the resulting change in per capita output as well as other variables of interest.

Figure 5 reports the results from one such simulation, while column 3 of Table 2 summarizes the results from

averaging over 1000 simulations. The main experiment generates about 70% (i.e. 2.129/3) of the increase in the log of output per capita observed in the data. Before discussing in detail the forces behind the increased output per capita growth, we discuss population dynamics. Panel *c* reveals that the fall in birth rates that the model generates is very small compared to the one that comprises the demographic transition beginning in 1880, about 24% of the empirical fall. This result is further reassurance of the validity of our hypothesis. In fact, if the decline were of the magnitude of the demographic transition, the model would predict the transition too early, starting around 1800. In this model, we only aim to examine the role of declining adult mortality in the process of productivity evolution, and consequently output dynamics. We decided to have endogenous fertility to obtain approximately correct population dynamics, which is also important for output dynamics, but the analysis of the demographic transition is outside the scope of this paper. There is no quantity-quality tradeoff in our model, and the demographic transition had to occur because of changing relative value of quantity and quality of children, due to possibly the nature of technological progress (e.g. Galor and Weil (2005)) or introduction of compulsory education and child labor reforms (Doepke and Zilibotti 2005, Hazan and Berdugo 2006). If people can respond to children's death by having more children and aim for a target number of surviving children, then the drop in child mortality (0-10), which began around the end of the 19th c., would also contribute to falling fertility. Again, we do not model this here.

We also see that the model fails to produce the rise in fertility (and consequently increase in population growth rate) of the second half of the 18th c. This rise is almost entirely attributed to changes in the timing and prevalence of marriage, with marital fertility remaining unchanged until the demographic transition (Floud and Johnson (2004) and Wilson and Woods (1991)). Whatever factors allowed earlier and more prevalent marriages are not captured by the mechanism we examine here.

This experiment generates a substantial rise in the time spent on innovation (from 5% to 14%) due to the rise in the expectation term  $E_{\omega_t} [(1 - m^y(\omega_t)) (1 - m^y(\omega_t))^{\theta_m \alpha(1-\sigma)} (1 - m^y(\omega_t))^{\alpha(\theta-1)(1-\sigma)}]$ , and consequently, the rise in the relative weight on future utility, shown in Panel *d*. Recall this term  $E()$  consists of three factors that we referred to as  $D()$ ,  $K()$  and  $M()$ . The declining mortality trend raises the discount and knowledge factors but lowers the Malthusian channel. The first two effects are stronger than the last one, which leads to the rise in  $E$ .

Panel *e* elucidates the sources of output growth: while resources per capita were diluted over time implying that output should shrink by 47% by 2000 (From Table 2,  $\left(\frac{\lambda_{2000}}{\lambda_{1600}}\right)^{1-\theta} = 0.526$ ), and even the declining labor input implied a further 7% drop in output per capita, the growth rate of productivity was large enough to generate the increase in output per capita comparable to 70% of its increase in the data (in log terms).

So, the growth rate of  $A_t$  was the driving force behind the growth rate of output per capita. In turn, the growth rate of  $A_t$  increased due to the direct effect coming from the rise in the rate of knowledge retention and due to the indirect effect coming from increased innovation activity. Both were important channel. The rate of knowledge retention,  $(1 - m^y)^{\theta_m}$ , increased from 0.929 to 0.999 (7.5% increase), while the growth of  $A_t$  over the period,  $1 + \xi i^n$ , increased from 1.15 to 1.3 (13.7%). Together, these changes implied 22.5% rise in the growth rate of  $A_t$  (from 6.9% in 1600 to 31% in 2000, over 25 year periods).

### 3.3.2. The Contribution to the Overall Effect of the Intergenerational Knowledge Transmission Channel

It is convenient to think of different channels through which the decline in mortality influences the dynamics of the main experiment. As already discussed, it acts to influence time allocation through its influence on the expectation term  $E$ , which itself is composed of three factors. In addition, mortality influences the evolution of TFP and population

Table 2: Results

	Data	Main Exp.	Main with Destruction of A of 1600	Main with Increasing Diffusion
Change in CBR	-11	-2.659	-2.494	-2.661
Change in ln y	3	2.129	1.422	2.138
$\frac{y_{2000}^y}{y_{1600}^y}$		7.596	3.773	7.659
$\frac{A_{2000}}{A_{1600}}$		15.519	7.786	15.653
$\left(\frac{l_{2000}}{l_{1600}}\right)^\theta$		0.934	0.94	0.934
$\left(\frac{\lambda_{2000}}{\lambda_{1600}}\right)^{1-\theta}$		0.5259	0.5172	0.5257
$\frac{A_{1625}}{A_{1600}}$		1.069	1.07	1.0694
$(1 - m_{1600}^y)^{\theta_m}$		0.929	0.9295	0.9332
$1 + \xi_{1600}^n$		1.1512	1.1512	1.1512
$\frac{A_{2000}}{A_{1975}}$		1.3097	1.2055	1.309
$(1 - m_{1975}^y)^{\theta_m}$		0.9989	0.9295	0.9991
$1 + \xi_{1975}^n$		1.3091	1.297	1.3093

dynamics directly.

In order to understand the contribution of mortality decline through its effect on improved knowledge retention alone, i.e. through its influence on factor  $K() = (1 - m^y(\omega_t))^{\theta_m \alpha(1-\sigma)}$  in the expectation term and factor  $(1 - m^y)^{\theta_m}$  in the law of motion of productivity alone, we simulate the economy again, except this time we keep these factors at their 1600 level. In other words, we eliminate their influence from the main experiment and examine how much of the overall success of the main experiment is lost. The results are reported in Figure 6 and Table 2.

Panel *d* reveals that the rise in the relative weight on future utility is only slightly smaller in this case, this is because  $\theta_m$  is very small in our calibration. Consequently, the rise in time spent innovating is only slightly smaller (5% to 13%). Nonetheless, this simulation accounts for a much lower part of the empirical rise in output per capita (only 47%), indicating that the effect of mortality decline through its influence on the intergenerational knowledge transmission comprised a major part of its the overall effect. Again, it is  $A$  that drives the increase in output per capita (panel *e*). In turn, the increase in the growth rate of  $A$  is driven by increased innovation time alone, as by construction, we shut down the knowledge retention channel (panel *f*).

### 3.3.3. Population Density and Knowledge Diffusion

It is reasonable that in our context adult mortality may have yet another channel through which it can influence evolution of productivity. This potentially important influence of mortality decline is through its influence on population density and consequently the rate of knowledge absorption across locations. To capture the idea that higher population density facilitates idea exchange across locations, we set

$$\sigma_{A,t} := \frac{(N_t^y + N_t^o) / \Lambda}{\psi + (N_t^y + N_t^o) / \Lambda}, \quad (30)$$

i.e., an increase in population density would imply an increase in the speed and efficiency of knowledge transmission across locations. Parameter  $\psi$  captures other relevant factors for idea flows across locations, such as the extent of availability of scientific communities, accessibility to communication devices, railroads, post, and others emphasized in Mokyr (2002). A decline in  $\psi$  would also increase the rate of knowledge absorption.

To maintain  $\sigma_{A,1600} = 0.3$  given the population structure in 1600 (the initial period), we need to set  $\psi = 500$ . Next, we perform the main experiment again, except that we allow for  $\sigma_{A,t}$  to evolve according to (30). Over the time period examined, population density rises and  $\sigma_A$  increases from 0.3 to 0.82. This change, however, has such a negligible effect on the dynamics generated by the model, that it is nearly identical to the main experiment dynamics reported in Figure 5. The results reported in the last column of Table 2 further confirm the lack of added influence from the increase in  $\sigma_{A,t}$ .

The main reason is that there is already little difference in TFP across locations under the low absorption rate of 0.3. A location with the worst mortality rate benefits more from diffusion in the next period and partially catches up. An increase in the rate of absorption would strengthen the catching up. However, as  $\sigma_A$  rises, mortality shocks also become less severe and differ less across locations.

### 3.4. Sensitivity Analysis

There are several moments we used in the calibration that were difficult to determine. For example, the fraction of time spent raising children,  $qn$ , the only cost of raising children in our model, is difficult to conceptualize and measure in the data. We used a low value of 0.05, because in the 17th c., children contributed substantially to family production thus greatly offsetting their costs to the parents. Here we ask how the results would change if we instead chose  $qn = 0.1$ . The calibrated values of  $\alpha$ ,  $\sigma$  and  $q$  then change to 0.92, 0.38 and 0.07, respectively. Recall that an increase in the expectation term  $E = E_{\omega_t} [(1 - m^y(\omega_t))^{1+\alpha(\theta_m+\theta-1)(1-\sigma)}]$  encourages time reallocation towards innovation. Because sigma is lower in this calibration,  $E$  increases more in response to declining mortality, strengthening the incentive to innovate. In the long run, the growth of  $A$  due to innovation is larger:  $1 + \xi i_{1975}^\eta = 1.32$  (compared to 1.309 in the benchmark model). So,  $A$  rises by a factor of 16.33 (larger than 15.5, its counterpart in the benchmark model), and it allows a larger rise in output per capita relative to that in the benchmark model. The main experiment accounts for 73% of the rise in the log of output per capita.

Next, we study how the results are affected by adjusting the calibration procedure to deliver  $\sigma > 1$ . Increasing our choice of time elasticity in the innovation sector to  $\eta = 0.78$  and further lowering the time spent innovating to 0.045 ensures that the calibration implies  $\sigma = 1.4$ . This change also implies that  $\xi = 1.7$ . Increasing  $\sigma$  to 1.4 actually increases the rise in  $E$  as a result of declining mortality. In the benchmark model, the influence through the knowledge transmission channel  $K(\omega_t) := (1 - m^y(\omega_t))^{\theta_m \alpha (1-\sigma)}$  on  $E$  (and hence on time spent innovating) was positive and the influence through the resource dilution channel  $M(\omega_t) := (1 - m^y(\omega_t))^{\alpha(\theta-1)(1-\sigma)}$  was negative. With the present and future composite goods being complements, both are reversed, but the overall rise in  $E$  is greater, because the influence of mortality is larger in magnitude through the resource dilution channel. An increase (a decrease) in returns to innovation due to the effect of mortality through the knowledge transmission (resource dilution) channel raises (lowers) the demand for both present and future consumption, and consequently induces labor reallocation away from (towards) innovating. So, the decline in mortality generates a larger increase in innovation, which results in a larger rise in productivity and per capita output:  $\frac{A_{2000}}{A_{1600}} = 16.47$  and  $\frac{y_{2000}^y}{y_{1600}^y} = 7.84$  (compared to 15.5 and 7.6 in the benchmark model). Here, 72% of the rise in the natural log of output per capita is explained.

Finally, we also examine the changes in the results to the calibration of  $\eta$ . Because we do not have other inputs in the production of human capital, the elasticity with respect to time should perhaps be set to a larger value. Setting  $\eta = 0.75$  changes the calibrated values for  $\zeta$  and  $\sigma$  to 1.43 and 0.82. The higher value of  $\eta$  increases returns to innovation and makes time allocation more responsive to changes in mortality. In addition, an increase in  $\sigma$  actually increases the rise in  $E$  due to declining mortality: although it decreases the positive influence through the knowledge channel, it also decreases the negative influence through the resource dilution channel. Thus, the main experiment yields a larger rise in innovation and a larger increase in productivity:  $1 + \zeta i_{1975}^{\eta} = 1.326$  and  $\frac{A_{2000}}{A_{1975}} = 1.328$  (their benchmark counterparts were 1.309 and 1.32). Nearly 75% of the rise in the natural log of output per capita is explained. We see that increasing  $\eta$  increases the importance of changes in mortality in accounting for the rise in output per capita.

## 4. Conclusions

We proposed a new causal link from gains in longevity to growth, based on the notion of embodied human capital and personal contact being important in the transfer of knowledge. According this link, a decline in mortality can contribute to increased growth by (1) directly improving knowledge transmission over time (2) indirectly encouraging h.c. accumulation, which can transpire even in the face of declining labor supply.

For reasonable micro assumptions on idea aggregation and their intergenerational transfer, we derived a simple form of adverse impact of adult mortality on knowledge transmission across generations. The destruction rate of productivity we obtained is an increasing and convex function of adult mortality.

In order to assess the quantitative relevance of our proposed link between mortality and growth, we performed a calibration exercise employing the available historical data on England. Our quantitative analysis suggests that declining adult mortality could be an important force behind the escape from stagnation. In fact, when we vary the time-dependent parameters of the mortality shock distribution according to our estimates, the parameterized model generates both early stagnation and the later rise in output per capita, which is approximately 70% of its empirical counterpart. We find that a large part of the overall effect is due to the direct effect of the knowledge retention channel.

Finally, we also explored the role of increasing knowledge diffusion across locations, which could be another outcome of declining mortality. We found its quantitative effect insignificant.

## 5. Appendix

### *Data Sources*<sup>15</sup>

*Index of Real GDP per capita* ( $y$ ): [1565-1865] - Clark (2001a), Table 7, p. 30, rescaled to equal 100 in 1565 (England and Wales); [1820-1990] - Maddison (1995), p. 194, rescaled to match Clark's index in 1850 (UK).

*Crude Birth and Crude Death Rates*: [1541 - 1871] - Wrigley et al. (1997) (England); [1871 - 1986] - Mitchell, 1978 (England and Wales).

*General Fertility Rate*: Computed using CBR and the fraction of females in the total population, taken from Wrigley et al. (1997) for [1541 - 1841] (England) and Human Mortality Database for [1841 - 1999] (England and Wales).

*Population Growth Rate*: [1541 - 1836] - Wrigley et al. (1997) (England); [1841 - 1999] - Human Mortality Database (England and Wales).

*Age-specific survival probabilities*: [1580-1837] - Wrigley et al. (1997) (England); [1841 - 1999] - Human Mortality Database (England and Wales).

### *Proof of Proposition 1*

It suffices to prove that the solution to

$$\max_{l_t, n_t} \frac{l_t^{\alpha\theta(1-\sigma)} n_t^{(1-\alpha)(1-\sigma)}}{1-\sigma} + \beta \frac{\left[ \tilde{A}_t/A_t + \zeta (1-l_t - qn_t)^\eta \right]^{\alpha(1-\sigma)}}{1-\sigma} E_t$$

$$l_t, n_t \geq 0, l_t + qn_t \leq 1$$

exists and is interior.

The objective function is continuous and the constraint set, which can be written as  $\{(l_t, n_t) \in \mathbb{R}^2 | l_t, n_t \geq 0, l_t + qn_t \leq 1\}$ , is compact. This guarantees existence.

Uniqueness is guaranteed by strict concavity of the objective and convexity of the constraint set. The constraint set is indeed convex. Next we derive sufficient conditions for strict concavity of the objective. It is enough that each summand in the utility is strictly concave. The first summand is

$$F(l_t, n_t) = \frac{\left[ l_t^{\alpha\theta} n_t^{(1-\alpha)} \right]^{1-\sigma}}{1-\sigma} = \frac{[f(l_t, n_t)]^{1-\sigma}}{1-\sigma}.$$

The function  $f$  is strictly concave if  $\alpha\theta < 0$  and  $\alpha\theta + 1 - \alpha < 1$ . The function  $g(x) = \frac{x^{1-\sigma}}{1-\sigma}$  is increasing and concave for  $\sigma > 0$ , so the composite  $F$  is strictly concave.

<sup>15</sup>Due to data limitations for England, we were forced to draw on the data sources available for England and Wales and UK. Although this inconsistency introduces some degree of error, we believe that it is small for the following reasons. (1) We do not consider level variables, such as GDP or population size, but instead growth rates, indices, and fractions of level variables. (2) For the period under consideration, the population of Wales is less than 6% of that of England. (3) Scotland's population size relative to that of England and Wales falls from 17% in 1820 (the earliest date for which we are forced to use UK data sources) to less than 10% today. (4) Appropriate rescaling was made in all cases.

Next we need to find restrictions on parameters that would make the second summand also strictly concave. Let

$$F(l_t, n_t) = \frac{\left[ \tilde{A}_t/A_t + \zeta (1 - l_t - qn_t)^\eta \right]^{\alpha(1-\sigma)}}{1 - \sigma} = \frac{[f(l_t, n_t)]^{\alpha(1-\sigma)}}{1 - \sigma}$$

The function  $f$  is concave if  $\eta < 1$ :

$$\begin{aligned} f(l_t, n_t) &= \tilde{A}_t/A_t + \zeta (1 - l_t - qn_t)^\eta \\ f_1 &= -\eta\zeta (1 - l_t - qn_t)^{\eta-1}, \quad f_2 = -q\eta\zeta (1 - l_t - qn_t)^{\eta-1} \\ f_{11} &= \eta(\eta - 1)\zeta (1 - l_t - qn_t)^{\eta-2}, \quad f_{22} = q^2\eta(\eta - 1)\zeta (1 - l_t - qn_t)^{\eta-2} \\ f_{12} &= q\eta(\eta - 1)\zeta (1 - l_t - qn_t)^{\eta-2} \\ f_{11} &< 0 \\ f_{11}f_{12} - f_{12}^2 &= 0 \end{aligned}$$

The function  $g(x) = \frac{x^{\alpha(1-\sigma)}}{1-\sigma}$  is increasing and strictly concave if

$$\begin{aligned} g'(x) &= \alpha x^{\alpha(1-\sigma)-1} > 0 \\ g''(x) &= \alpha[\alpha(1-\sigma) - 1]x^{\alpha(1-\sigma)-2} < 0 \end{aligned}$$

The combined sufficient conditions for concavity of the entire utility function are

$$\begin{aligned} \alpha\theta + 1 - \alpha &< 1, \\ \alpha(1 - \sigma) - 1 &< 0. \end{aligned}$$

These are obviously satisfied if  $\alpha, \eta \in (0, 1)$  and  $\sigma > 0$ .

Now consider the Kuhn-Tucker conditions for  $l_t$  and  $n_t$ :

$$\begin{aligned} \alpha\theta l_t^{\alpha\theta(1-\sigma)-1} n_t^{(1-\alpha)(1-\sigma)} - \beta\alpha [\cdot]^{\alpha(1-\sigma)-1} \eta\zeta (1 - l_t - qn_t)^{\eta-1} E_t &\leq 0, \\ (1 - \alpha) l_t^{\alpha\theta(1-\sigma)} n_t^{(1-\alpha)(1-\sigma)-1} - \beta\alpha [\cdot]^{\alpha(1-\sigma)-1} \eta\zeta (1 - l_t - qn_t)^{\eta-1} qE_t &\leq 0, \end{aligned}$$

with strict equalities only if  $l_t, n_t \geq 0$ . The strict inequalities hold when  $l_t, n_t = 0$ .

We see that as long as

$$\begin{aligned} \alpha\theta - 1 &< 0 \\ (1 - \alpha)(1 - \sigma) - 1 &< 0 \\ \eta &< 1 \end{aligned}$$

it is optimal to choose  $l_t, n_t, 1 - l_t - qn_t \geq 0$ . Again, these are satisfied with the assumptions on  $\alpha, \eta, \theta, \sigma$ .

*Estimation of parameters of the mortality shock distribution*

We assume that

$$\ln \omega_t = \frac{a_1}{1 + b_1^{t-c_1}} + d_1 + \varepsilon_t, \text{ with } \varepsilon_t \sim N(0, \sigma_t) \text{ and } \sigma_t = \frac{a_2}{1 + b_2^{t-c_2}} + d_2.$$

Note that this implies that  $\ln \omega_t \sim N(\mu_t, \sigma_t)$ , with  $\mu_t = \frac{a_1}{1 + b_1^{t-c_1}} + d_1$  and  $\sigma_t = \frac{a_2}{1 + b_2^{t-c_2}} + d_2$ . We want to estimate  $\{\mu_t, \sigma_t\}_{t=1640, 1665, \dots}$  to maximize the log-likelihood of the observed sequence of adult mortality rates. Denoting the vector of parameters by  $\Omega = \{a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2\}$  and the sequence of observations  $\left\{ \ln \left( \frac{m_t}{1 - m_t} \right) \right\}$  by  $\{x_t\}$ , the likelihood function is

$$L(\Omega | \{x_t\}) = \prod_{t=1}^T f(x_t | \Omega), \text{ where } f(x_t | \Omega) = \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{1}{2} \left(\frac{x_t - \mu_t}{\sigma_t}\right)^2\right).$$

The log-likelihood function is then

$$l(\Omega | \{x_t\}) = -\sum_{t=1}^T \ln \sigma_t - \frac{1}{2} \sum_{t=1}^T \left(\frac{x_t - \mu_t}{\sigma_t}\right)^2.$$

## References

- Acemoglu, Daron and Simon Johnson (2005), "Disease and Development: The Effect of Life Expectancy on Economic Growth," working manuscript, MIT.
- Acemoglu, Daron (2008), *Introduction to Modern Economic Growth*, unpublished book, MIT.
- Ashraf, Quamrul and Oded Galor (2008), "Malthusian Population Dynamics: Theory and Evidence," working manuscript, Brown University.
- Ben-Porath, Yoram (1967), "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, August, 75, pp. 352–365.
- Bleakley, Hoyt (2006). "Comments on Acemoglu and Johnson *Disease and Development*."
- Boucekkine, Raouf, David de la Croix, and Omar Licandro (2002), "Vintage Human Capital, Demographic Trends, and Endogenous Growth," *Journal of Economic Theory*, 104, pp. 340–375.
- (2003), "Early Mortality Decline at the Dawn of Modern Growth," *Scandinavian Journal of Economics*, 105, pp. 401–418.
- Cervellati, Matteo and Uwe Sunde (2005), "Human Capital Formation, Life Expectancy and Process of Economic Development," *American Economic Review*, 95 (5), pp. 1653–1672.
- Clark, Gregory (2001a), "Microbes and Markets: Was the Black Death an Economic Revolution?" Working Manuscript, UC Davis.
- Clark, Gregory (2001b), "The Secret History of the Industrial Revolution." Working Manuscript, UC Davis.
- Crafts, Nicholas and Terence C. Mills (2007), "From Malthus to Solow: How Did the Malthusian Economy Really Evolve?," *Journal of Macroeconomics*.

- Doepke, Matthias and Fabrizio Zilibotti (2005), "The Macroeconomics of Child Labor Regulation," *American Economic Review*, December, 95 (5), pp. 1492–1524.
- Ehrlich, Isaac and Francis T Lui (1991), "Intergenerational Trade, Longevity and Economic Growth," *Journal of Political Economy*, October, 99 (5), 1029–1059.
- Fernandez-Villaverde, Jesus (2001), "Was Malthus Right? Economic Growth and Population Dynamics." working manuscript.
- Jones, Charles (2001), "Was an Industrial Revolution Inevitable? Economic Growth over the Very Long Run." *Advances in Macroeconomics*, 1(2), pp. 1028-1028.
- (1995), "R&D-Based Models of Economic Growth," *Journal of Political Economy*, University of Chicago Press, vol. 103(4), pp. 759-84.
- Galor, Oded (2005), "From Stagnation to Growth: Unified Growth Theory." *Handbook of Economic Growth* (Aghion, Philippe and Durlauf, Steven N., eds.). Amsterdam: Elsevier, volume 1B, pp. 1113-1180.
- Galor, Oded and Weil, David N. (2000), "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review*, 90(4), pp. 806-828.
- Greenwood, Jeremy and Ananth Seshadri (2002), "The U.S. Demographic Transition," *American Economic Review*, May, 92 (2), pp. 153–159.
- Halverson, Melissa Sue (2007), "Native American Beliefs and Medical Treatments During the Smallpox Epidemics: an Evolution," *The Early America Review*, Vol. VII, No. 4.
- Hazan, Moshe (2006), "Longevity and Lifetime Labor Input: Data and Implications," working manuscript, Hebrew University.
- and Binyamin Berdugo (2002), "Child Labor, Fertility, and Economic Growth," *Economic Journal*, 112 (482), pp. 810–828.
- and Hosny Zoaby (2006), "Does Longevity Cause Growth? A Theoretical Critique," *Journal of Economic Growth*, Vol. 11, pp. 363–376, December.
- Human Mortality Database, <http://www.mortality.org>.
- Kalemli-Ozcan, Sebnem; Ryder, Harl E. and Weil, David N. (2000), "Mortality Decline, Human Capital Investment and Economic Growth." *Journal of Development Economics*, 62(1), pp. 1-23.
- Lagerlöf, Nils-Petter (2003), "From Malthus to Modern Growth: Can Epidemics Explain the Three Regimes?" *International Economic Review*, 44(2), pp. 755-777.
- Lorentzen, Peter, John McMillan, and Romain Wacziarg (2005), "Death and Development," forthcoming in the *Journal of Economic Growth*.
- Lucas, Robert (2008) "Ideas and Growth," NBER working paper No. 14133.
- Macdonald, Stuart (1979), "The Diffusion of Knowledge among Northumberland Farmers, 1780-1815," *The Agricultural History Review*, 27, 1, pp.30-39.
- Malthus, Thomas Robert (1798) *An Essay on the Principle of Population, as it affects the future improvement of society with remarks on the speculations of Mr. Godwin, M. Condorcet, and other writers*. Anonymously published.

- Maddison, Angus (2003), *The World Economy: Historical Statistics*, OECD, Paris.
- Monteiro and Pereira (2007), "From Growth Spurts to Sustained Growth: The Nature of Growth and Unified Growth Theory," working manuscript, University of York.
- Mokyr, Joel. (2002), *The Gifts of Athena: Historical Origins of the Knowledge Economy*, Princeton University Press, Princeton.
- Murphy, M.J. and Dyson, T.P.G. (1986), *Vital Statistics Time Series, 1749-1982* [computer file]. Colchester, Essex: UK Data Archive [distributor], March 1986. SN: 2093.
- Soares, Rodrigo R (2005), "Mortality Reductions, Educational Attainment, and Fertility Choice." *American Economic Review*, 95(3), pp. 580-601.
- Shastri, Gauri Kartini and David N. Weil (2003), "How much of cross-country income variation is explained by health?" *Journal of the European Economic Association*, 1, pp. 387-396.
- Tamura, Robert F (2006), "Human Capital and Economic Development." *Journal of Development Economics*, 79(1), pp. 26-72.
- Wrigley, E.A.; Davies, R.S.; Oeppen, J.E. and Schofield, R.S (1997), *English Population History from Family Reconstitution 1580-1837*. Cambridge: Cambridge University Press.

FIGURE 1: ESTIMATES OF THE MORTALITY SHOCK PROCESS











