J Econ Growth (2010) 15:291–321
DOI 10.1007/s10887-010-9059-5

The role of mortality in the transmission of knowledge

Michael Bar · Oksana Leukhina

Published online: 2 December 2010

Abstract We investigate, both theoretically and quantitatively, a previously unexplored link between gains in adult mortality and productivity growth. Our mechanism allocates a central role to individuals as carriers of useful ideas and to personal contact as an important means of transferring these ideas. It thus implies that disrupting a human life impedes the process of knowledge transmission across time. We derive a simple and intuitive form of the dependence of aggregate knowledge transfer on adult mortality and incorporate it into a model of endogenous growth. We then quantitatively examine the relevance of the proposed link in application to the long-run growth experience of England. Our calibration exercise suggests that the reduction in adult mortality, by improving knowledge transmission across time and encouraging more innovation, was a quantitatively important force behind the takeoff in output per capita.

Keywords Economic growth · Total factor productivity · Adult mortality · Longevity · Knowledge transmission · Ideas · Human capital

JEL Classification O40 · O31 · O33 · J10 · N10

1 Introduction

Motivated by the strong positive association between gains in longevity and growth, found both in time series and cross-sectional studies (e.g. Kelley and Schmidt 1995; Shastry and

M. Bar
Department of Economics, San Francisco State University, San Francisco, CA 94132, USA

O. Leukhina (✉)
Department of Economics, University of Washington, Seattle, WA 98195, USA

Weil 2003; Lorentzen et al. 2008; Aghion et al. 2009), we investigate a previously unexplored link between adult mortality and productivity growth.[1]

Our mechanism is rooted in the premise that in preindustrial and traditional societies, a large body of useful knowledge was tacit, uncodifiable and embodied in individuals. Such embodied knowledge is to be distinguished from disembodied knowledge, which is easily replicated, usable in as many activities as desired, and lives on independently from its physical creators, and which has become the center of modeling *contemporary* economic growth (Romer 1990). Embodied knowledge is tied to its carrier, and its effective transfer between people requires extensive personal contact. In fact, we observe that most learning in traditional societies was accomplished through prolonged master–apprentice and parent–child relationships. Even today, personal contact seems very important for knowledge sharing as exemplified by conferences, academic visits, board meetings, advisor–student relationships, etc.

In a society where knowledge is embodied in individuals and personal contact is essential for transferring knowledge, the premature death of a parent/master interrupts the skill transfer to the children/apprentices;[2] this is the direct impact of mortality on productivity growth. In addition, the anticipation of the low rate of knowledge transmission across time characterizing high mortality regimes discourages investment in productivity, representing an indirect impact of mortality. Our main hypothesis is that the permanent reduction in adult mortality facilitated the takeoff in productivity growth by improving knowledge transmission across time 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 the bubonic plague and the scarlet fever epidemic that greatly depopulated the Senecas in the 1630s eliminated much of their knowledge of ceramic craftsmanship (Halverson 2007). Even though HIV, unlike preindustrial diseases and epidemics, kills people slowly, the AIDs epidemic in Sub-Saharan Africa also appears to have disrupted the transmission of indigenous farming practices to children.[3] In fact, Waterhouse (2005) documents a large negative impact of HIV/AIDS on farmers' knowledge in the identification of seed through conducting farmer surveys in several provinces of Mozambique. The negative impact of mortality on local productivity in preindustrial times in Europe is evidenced by the well-documented guilds' practice of sending their skilled craftsmen to regions that had recently experienced an epidemic. At a macro level, there are several preindustrial records of large-scale mortality episodes interfering with productivity growth spurts: for example, Black Death in the case of fourteenth century Europe, wars with the Ottoman empire in the case of the Italian Golden Age and Jurchen and Mongol invasions in the case of Sung China (Monteiro and Pereira 2007). Several more historical episodes of

[1] Acemoglu and Johnson (2007) find no evidence that increases in output per capita levels are associated with increases in life expectancy. However, as shown by Aghion et al. (2009), if one includes the level effect of lifetime expectancy in addition to its growth effect, then the same dataset and methodology yield a positive dependence of output per capita growth on life expectancy and its growth. Consistently with these findings, our model incorporates both the level and the growth effects.

[2] Our explicit modeling of mortality interfering with the process of knowledge transfer across individuals by diminishing the amount of personal contact in learning relationships, i.e. the first effect, is closely related to the idea developed in Lucas (2009). However, while in Lucas (2009), everyone gets a productivity draw, and high mortality limits the number of personal contacts with others that can help raise one's productivity, we assume that one learns from a parent/master and high mortality limits the amount of personal contact in these learning pairs.

[3] This information was collected from www.un.org and www.fao.org.

technological regress following a population decline are discussed in Aiyar et al. (2008), who also attempt to understand why high mortality episodes often lead to the loss of knowledge.[4]

We first derive the impact of adult mortality on productivity, starting with reasonable assumptions regarding the process according to which embodied ideas are transferred across time and regarding the mapping of these ideas into aggregate productivity. We obtain a simple and intuitive form of the adverse impact of adult mortality on productivity growth that captures the insight described above. We derive that the destruction rate of productivity is an increasing and convex function of adult mortality experienced by the current and past cohorts. This convexity arises due to the assumption of diminishing marginal returns to the cohort fraction of carriers of a given idea/skill. We then embed this impact of adult mortality on productivity transmission into a general equilibrium model of endogenous technological change.[5] We allow for knowledge diffusion across regions in order to moderate the adverse impact of a high realization of a region-specific 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 takeoff of output per capita, focusing on the historical period from 1680 to 1880, the midpoint of which dates the beginning of the Industrial Revolution.[6] We first parameterize the model using the method of moments and the assumption of balanced growth (BG) to capture the available economic and demographic observations for England around 1600–1700. We employ the historical data on real wage dynamics following the Black Death epidemic in England to identify the key parameter in the knowledge destruction function. The Black Death is the perfect case study, as it entailed an unexpected disease, with a short period of illness, followed by death—exactly the situation that should lead to low rates of knowledge transfer over time. Taking as given fertility rates and the time varying parameters of the mortality shock process, estimated using the actual time series data on age-specific mortality, we quantitatively study the role of gains in adult longevity in the takeoff from stagnation to growth. Figure 1 reports the actual time series of adult mortality (ages 25–50) alongside output per capita, revealing the positive relation between gains in adult longevity and growth.

Our findings suggest that, in addition to their more conventional effects, reductions in adult mortality contribute to the process of development by extending the amount of personal contact in learning relationships and thereby improving the process of knowledge transmission across generations. In our quantitative exercise, we find that the decline in adult mortality generates about 90% of the rise in the log of output per capita observed during the two hundred year period that we study. Moreover, the influence it exerted through the knowledge transmission mechanism proposed here accounts for one half of the empirical rise in the log of output per capita, working through both, the direct effect and the indirect effect on productivity investments.

[4] While we focus on the disruption of knowledge transfer as a direct consequence of an epidemic, this paper analyzes a different reason for why some knowledge may be forgotten forever. A smaller population size, and hence a lower aggregate demand, imply that it is not profitable to produce some of the varieties. These production techniques are not passed on to the future generations. Note that this mechanism is also built on a premise of embodied knowledge.

[5] Bar and Leukhina (2010a), using price data, estimates large productivity changes during 1600–1910 and finds that these changes were largely responsible for the economic transformation of England. Hence, we focus on changes in productivity, abstracting from physical capital here.

[6] It is important that we are able to use age-specific mortality data. The reason that we focus on England is that such data is not available for other countries.

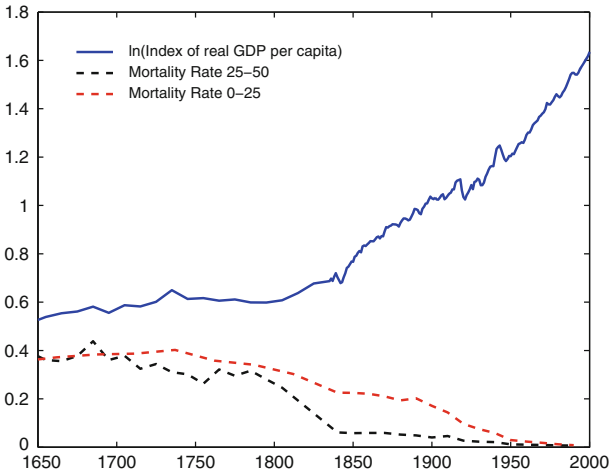

Fig. 1 Mortality rates and output per capita in England

We believe that our model, built on the notion of embodied knowledge, is suitable for examining the period up to 1880. Mokyr (2002) argues that "most practical knowledge in the eighteenth century 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 nineteenth century." 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 nineteenth century] was primarily due to the practical experience by men of little or no formal scientific training (Mitch 1998)."

Nineteenth century England is often associated with a number of developments, other than gains in adult longevity, that could potentially improve the speed and efficiency of knowledge transfers, such as the rise in informal learning networks and scientific communities, the rise in population density, which would make idea exchange among productive peers easier, the decline in transportation and post costs, and a movement away from small family farms towards larger-scale production (Mokyr 2002). In the context of our framework, these developments would (1) further reduce the negative impact of adult mortality on knowledge transfer across time and hence reinforce the mechanism examined here and (2) increase the diffusion of knowledge across regions. We explore the potential effects of increasing knowledge diffusion across regions in Sect. 3.4.1, by modeling the rate of absorption of the gap between the local and frontier technology as positively dependent on population density.

Finally, we place our work in the context of existing literature on takeoffs. By modeling the impact of mortality on knowledge transmission, we contribute to the previous work that emphasized the important role of gains in life expectancy in the process of output per capita takeoff. Such papers include Ehrlich and Lui (1991), Kalemli-Ozcan et al. (2000), Kalemli-Ozcan (2002), Boucekkine et al. (2002, 2003), Lagerlöf (2003a,b), Cervellati and Sunde (2005), Soares (2005) and Tamura (2006) among others.[7]

Two mechanisms linking declines in mortality to increases in growth are commonly used. First, a decline in child mortality and/or its uncertainty reduces parental costs of educating each surviving child, inducing greater investments in children's human capital, and in turn, leading to a takeoff in output per capita and a drop in fertility through the quantity–quality

[7] Cervellati and Sunde (2005), Lagerlöf (2003a,b), and Tamura (2006) endogenize mortality.

trade-off. The timing of reductions in adult mortality, however, is better aligned with the period of takeoff than the timing of reductions in child mortality. For the case of England, for example, output per capita accelerates around 1800, while child mortality declines at the very end of the nineteenth century (Fig. 1).

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 gains in life expectancy to growth. Gains in life expectancy increase the expected period over which investments in human capital are paid off, consequently encouraging more human capital accumulation and growth. This particular link has been recently criticized by Hazan and Zoaby (2006) and Hazan (2009). The latter paper documents that for cohorts of American men born in 1840–1970, the labor input declined despite the dramatic gains in life expectancy. The author then argues that because a rise in the lifetime labor supply is a necessary implication of the Ben-Porath type model that he examines, gains in longevity could not have caused human capital accumulation and hence growth via the Ben-Porath mechanism. Although a closer look into more general Ben-Porath type frameworks is warranted to determine whether or not they can maintain the causal link from gains in longevity to growth in the face of a declining labor supply, the model proposed here, which also incorporates the Ben-Porath channel, certainly can; i.e., it reconciles Hazan's findings with the possibility of causation from gains in longevity to growth. First, the direct impact of the decline in adult mortality on knowledge retention over time and consequently on growth, is independent of the lifetime labor supply. Second, the indirect impact of the decline in adult mortality on growth is through increasing innovation time. This effect can exist even in the presence of a declining lifetime labor supply, because falling mortality directly raises the return to human capital accumulation by improving knowledge transmission over time, i.e. by raising the productivity of time spent innovating.[8]

Because we show that the existence of a high adult mortality regime can contribute to long periods of stagnation, our mechanism is also complementary to the Malthusian mechanism, the conventional explanation for long periods of stagnation.[9] Recently, this effect has been subjected to quantitative analysis by Crafts and Mills (2007), Ashraf and Galor (2011) and Voigtlander and Voth (2009). The effect of a temporary and unexpected high mortality shock introduced in the model proposed here is consistent with a positive Malthusian check. Despite its negative impact on productivity, it raises output per capita, because it raises per capita land holdings.

Finally, because we endogenize gains in productivity, our work also complements existing work that emphasizes the importance of technological progress in driving the economic transformation (e.g. Galor and Weil 2000; Fernandez-Villaverde 2001; Jones 2001; Hansen and Prescott 2002; Greenwood and Seshadri 2002; Doepke 2004; Bar and Leukhina 2010a).

The rest of the paper is organized as follows. In Sect. 2, we discuss the link between adult mortality and intergenerational knowledge transfer, present the environment and equilibrium analysis. In Sect. 3, we present our quantitative analysis. This includes estimating the time-dependent parameters of the mortality shock distribution, calibrating the model, performing the experiments and sensitivity analysis. We conclude in Sect. 4.

[8] We should point out that there exist other mechanisms that can be adopted to raise returns to human capital accumulation even in the presence of a declining lifetime labor input, e.g. the link between population density and productivity of human capital production, explored in Lagerlöf (2003a,b) and Boucekkine et al. (2007).

[9] See Galor (2005) for the description of the Malthusian epoch.

2 The benchmark model

Time is discrete and indexed by $t = 0, 1, 2 \ldots$. There are J regions within the country, indexed by $j \in \{1, 2, \ldots, J\}$, each endowed with Λ_j acres of land, total land is given by $\Lambda = \sum_j \Lambda_j$. We model several regions to moderate the adverse impact of a region-specific epidemic. We represent total population in region j at time t by $N_{t,j}$, and the total country population by $N_t = \sum_j N_{t,j}$. In each region, there is a representative dynasty.[10] Every generation lives for at most three periods, as a child, a young adult and an old adult, indicated by the superscripts c, y and o. The populations of children, and young and old adults in region j at time t are denoted by $N^c_{t,j}$, $N^y_{t,j}$ and $N^o_{t,j}$.

2.1 Laws of motion for population and land

We intentionally abstract from modeling endogenous fertility and focus on the period prior to the fall in birth rates, the beginning of which dates to 1880. The main reason for doing so is to ensure that the model closely matches the population dynamics and hence the evolution of land per capita, which is a crucial factor behind growth in output per capita.[11]

In each period, random fractions of children and young adults in region j, denoted by $m^c_{t,j}$ and $m^y_{t,j}$, do not survive to the next period. Similar to Lagerlöf (2003a,b), we model mortality rates as functions of a log-normally distributed random variable, $\ln(\omega_t) \sim N\left(\mu^\omega_t, \sigma^\omega_t\right)$, drawn at the end of each period. More precisely, we assume that the realized mortality rates are given by

$$m^y_{t,j} = m^y\left(\omega_{t,j}\right) = \frac{\omega_{t,j}}{1+\omega_{t,j}} \quad \text{and} \quad m^c_{t,j} = \zeta_t m^y\left(\omega_{t,j}\right), \tag{1}$$

where $\omega_{t,j}$ denotes region- and time-specific realization of mortality shock ω_t, and ζ_t captures the difference in trends of child and adult mortality. Thus, the survival rate in region j is a strictly decreasing function of $\omega_{t,j}$. The implied distribution of mortality rates has the support of [0,1] and ensures that large-scale epidemics are rare events. Note that the distribution parameters for ω_t are indexed by time, and will be estimated, along with the time series for ζ_t, to match age-specific mortality data.[12] While all regions possess identical μ^ω_t, σ^ω_t, ζ_t, they will differ in mortality shock realizations.

[10] Alternatively, we could reinterpret the setup as consisting of J representative dynasties within a single location.

[11] In addition, since the probability of a 25 year old surviving to 50 was already as large as 95% by 1880, there is no reason to believe that changes in adult mortality could contribute to the fall in birth rates. We find it more plausible that the decline in birth rates in the late nineteenth century can be attributed to factors such as compulsory education and child labor reforms (Hazan and Berdugo 2002; Doepke and Zilibotti 2005) or the change in the nature of the technological progress to being more skill-biased (Galor 2005). Previous versions of this paper did incorporate endogenous fertility and the time cost of raising children, but we found that their inclusion added little to the main message of the paper.

[12] Lagerlöf (2003a,b) instead keeps the distribution parameters for ω fixed, but allows the constant in the denominator to decrease as the amount of human capital in the economy increases. In that model, only two periods of life are assumed, and hence it is the mortality in the first period of life (eliminating those who do not make decisions) that is modeled. By contrast, we model both child and young adult mortality, focusing on the causal link from young adult mortality to growth.

The following relationships are implied

$$N^y_{t+1,j} = \left(1 - m^c_{t,j}\right) N^c_{t,j}, \tag{2}$$

$$N^o_{t+1,j} = \left(1 - m^y_{t,j}\right) N^y_{t,j}, \tag{3}$$

$$N^c_{t,j} = n_t N^y_{t,j}, \tag{4}$$

where n_t denotes fertility, assumed to be exogenous and the same across regions.

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

$$N^y_{t+1,j} = \left(1 - m^c_{t,j}\right) n_{t,j} N^y_{t,j}, \tag{5}$$

$$N^o_{t+1,j} = \left(1 - m^y_{t,j}\right) N^o_{t,j} \left(1 - m^c_{t-1,j}\right) n_{t-1,j}/(1 - m^y_{t-1,j}). \tag{6}$$

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

$$\lambda^y_{t,j} = \frac{0.5\Lambda_j}{N^y_{t,j}}, \tag{7}$$

$$\lambda^o_{t,j} = \frac{0.5\Lambda_j}{N^o_{t,j}}. \tag{8}$$

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:

$$\lambda^o_{t+1,j}\left(\omega_{t,j}\right) = \frac{\lambda^y_{t,j}}{1 - m^y\left(\omega_{t,j}\right)}. \tag{9}$$

2.2 Mortality and the law of motion of total factor productivity

The production technology in region j is given by $Y_{t,j} = A_{t,j} L^{\theta}_{t,j} \Lambda^{1-\theta}_j$, where $L_{t,j}$ is the labor input, Λ_j is land input, $A_{t,j}$ is total factor productivity (TFP), and $\theta \in (0, 1)$. The purpose of this section is to motivate the form of TFP law of motion employed in our model.

We assume that for the duration of a time period, TFP can be incremented due to innovation activity and diffusion of ideas from other regions. While the form of these increments can be borrowed from the growth literature, the challenge is to motivate a reasonable form of TFP dependence on realized mortality. We aim to capture the idea that premature death interferes with the intergenerational knowledge transmission, and hence adversely impacts future TFP. Instead of simply assuming an ad-hoc form of dependence, we want to be explicit about the assumptions that can deliver the form we employ. To do so, we follow Lucas (2009) and start with the premise that all presently known ideas and techniques reside in individual heads.

Suppose that useful ideas (or production techniques) known collectively in region j at the beginning of period t are indexed by $i = 1, 2, \ldots, I_{t,j}$. An adult can embody any number of known ideas. The number of young and old adult carriers of idea i are denoted by $N^y_{t,j}(i)$ and $N^o_{t,j}(i)$. Assume that TFP level is a sum of known ideas' contributions $a_{t,j}(i)$,

$$A_{t,j} = \sum_{i=1}^{I_{t,j}} a_{t,j}(i).$$

An idea-specific contribution $a_{t,j}(i)$ depends on the fraction of young adults and the fraction of old adults that embody this idea,

$$a_{t,j}(i) = \left(\frac{N^y_{t,j}(i)}{N^y_{t,j}}\right)^{\phi}\left(\frac{N^o_{t,j}(i)}{N^o_{t,j}}\right)^{\phi},$$

where ϕ is a small number governing the degree of diminishing returns to the cohort-specific fractions of idea carriers, identification of which will be crucial for our quantitative results. Note we assume the old carriers are just as important as the younger ones: their expertise may be essential in facilitating the application of a given technique, especially in farming economies (Rosenzweig and Wolpin 1985).

Suppose that innovation and diffusion terms, which augment TFP during a given period, do so by introducing new ideas. Let these new ideas be indexed by $i = I_{t,j}+1,\ldots,I_{t,j}+\Delta_{t,j}$, so that the number of embodied ideas by the end of period t and in the beginning of period $t+1$ is $I_{t+1,j} = I_{t,j} + \Delta_{t,j}$.

Recall the mortality shock is drawn at the end of the period. By the law of large numbers, fraction $1 - m^y_{t,j}$ of young adult carriers of idea i, known at the end of period t, survive to the next period and become old adult carriers of idea i:

$$N^o_{t+1,j}(i) = \left(1 - m^y_{t,j}\right) N^y_{t,j}(i). \tag{10}$$

Further assume that a child observing a parent produce throughout the entire period (i.e. for a long enough time), will master the technique, while a child whose parent dies prematurely, will not receive the sufficient amount of personal contact needed to acquire the parent's skills. We do not model the motive behind this knowledge transfer. The simple message we aim to deliver is that whatever this motive might be, learning by watching, altruistic transfer, or beneficial exchange, if the amount of contact is cut short by death, learning is disrupted.[13] Hence, only the young adults who survive to old adulthood, i.e. fraction $1 - m^y_{t,j}$ of their cohort, pass all of their ideas to their surviving children, $\left(1 - m^c_{t,j}\right) n_{t,j}$. The law of motion of the young adult carriers of idea i is then given by

$$N^y_{t+1,j}(i) = \left(1 - m^y_{t,j}\right) N^y_{t,j}(i) \left(1 - m^c_{t,j}\right) n_{t,j}. \tag{11}$$

We allude to knowledge transmission along genetic lines strictly for illustrative purposes. As long as every adult (master) gets to train $n_{t,j}$ students, i.e. the average number of children in the economy, the same law of motion obtains.

Relations (5) and (11) imply that the young adult cohort's knowledge of idea i depends on the knowledge of this idea by their parents' cohort and their parents' survival:

$$\frac{N^y_{t+1,j}(i)}{N^y_{t+1,j}} = \frac{\left(1 - m^y_{t,j}\right) N^y_{t,j}(i) \left(1 - m^c_{t,j}\right) n_{t,j}}{\left(1 - m^c_{t,j}\right) n_{t,j} N^y_{t,j}} = \left(1 - m^y_{t,j}\right) \frac{N^y_{t,j}(i)}{N^y_{t,j}}.$$

Using (10) and (3), we also have $\frac{N^o_{t+1,j}(i)}{N^o_{t+1,j}} = \frac{\left(1-m^y_{t,j}\right)N^y_{t,j}(i)}{\left(1-m^y_{t,j}\right)N^y_{t,j}} = \frac{N^y_{t,j}(i)}{N^y_{t,j}}$, i.e. the fraction of idea i carriers in a given cohort remains the same over time as young adult mortality equally

[13] Introducing knowledge transfer as a choice in the context of our model would be potentially interesting. We expect the results to be reinforced, as the change in the mortality regime would increase the expected returns to a knowledge transfer.

affects the number of this cohort members and the number of this cohort's carriers of i. Manipulating the above further gives

$$\frac{N^o_{t+1,j}(i)}{N^o_{t+1,j}} = \frac{N^y_{t,j}(i)}{N^y_{t,j}} = \left(1-m^y_{t-1,j}\right)\frac{N^y_{t-1,j}(i)}{N^y_{t-1,j}} = \left(1-m^y_{t-1,j}\right)\frac{N^o_{t,j}(i)}{N^o_{t,j}},$$

that is, the relationship between the fraction of old adult carriers and the fraction of old adult carriers in the older (their parents') cohort is given by the survival rate of the parents $\left(1-m^y_{t-1,j}\right)$, precisely the rate at which ideas were transmitted between these two cohorts.

The above two relations then allow us to derive the law of motion of idea i contribution to TFP:

$$a_{t+1,j}(i) = \left(\frac{N^y_{t+1,j}(i)}{N^y_{t+1,j}}\right)^{\phi}\left(\frac{N^o_{t+1,j}(i)}{N^o_{t+1,j}}\right)^{\phi} = \left(1-m^y_{t,j}\right)^{\phi}\left(1-m^y_{t-1,j}\right)^{\phi}a_{t,j}(i).$$

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

$$A_{t+1,j} = \sum_{i=1}^{I_{t+1,j}} a_{t+1,j}(i) = \left(1-m^y_{t,j}\right)^{\phi}\left(1-m^y_{t-1,j}\right)^{\phi}\left[A_{t,j} + \sum_{i=I_{t,j}+1}^{I_{t,j}+\Delta_{t,j}} a_{t,j}(i)\right].$$

The second term in the brackets is TFP augmentation, which we model directly as due to the average time spent by young adults on innovation $i_{t,j}$ and diffusion of ideas from other regions: $\sum_{i=I_{t,j}+1}^{I_{t,j}+\Delta_{t,j}} a_{t,j}(i) = A_{t,j}i^{\eta}_{t,j} + d_{t,j}$, where $\eta \in (0,1)$. Note that it is assumed that innovation results from building on what is already known.[14] There is extensive evidence that in preindustrial time, landlords and tenant farmers actively engaged in the costly process of improving their methods of production: Macdonald (1979), using postal records, uncovers a great variety of methods farmers used for that purpose, such as traveling and corresponding with other farmers, sending their children to another farm for a year to master a more advanced technique, and attempting new techniques on their own farms. In fact, Macdonald (1979) reports that at the end of the eighteenth century, farmers were willing to pay as much as 60 pounds, or roughly twice the annual farm wages, to "intern" on a more progressive farm.

To summarize, we obtain the following law of motion of TFP

$$A_{t+1,j} = \left(1-m^y_{t,j}\right)^{\phi}\left(1-m^y_{t-1,j}\right)^{\phi}\left[A_{t,j} + A_{t,j}i^{\eta}_{t,j} + d_{t,j}\right]. \tag{12}$$

We see that the fraction of potential TFP lost due to premature death interfering with knowledge transmission across time, i.e. $1-\left(1-m^y_{t,j}\right)^{\phi}\left(1-m^y_{t-1,j}\right)^{\phi}$, is an increasing and convex function of time $t-1$ and time t adult mortality rates. We refer to the influence of mortality through factors $\left(1-m^y_{t,j}\right)^{\phi}\left(1-m^y_{t-1,j}\right)^{\phi}$ in the law of motion (12) as its *direct effect* on knowledge transmission across time.

There are several important messages to take. First, for reasonable assumptions on idea aggregation and the process of their intergenerational transfer, we derived the adverse effect of adult mortality on the evolution of TFP: the knowledge transmission factor $\left(1-m^y_{t,j}\right)^{\phi}\left(1-m^y_{t-1,j}\right)^{\phi}$ is less than one and decreases in both mortality rates. Second,

[14] The motivation for this assumption can be found in papers by Weitzman (1998), Olsson (2000) and Lucas (2009) that model the process of new idea creation.

the adverse impact of adult mortality on knowledge transmission is due to the premature death of the parent (master) disrupting the technique transfer to the next generation. Finally, the above derivations explicitly reveal the assumptions that give rise to the convexity of the knowledge destruction function. It arises due to the assumption of diminishing marginal returns to cohort-specific fractions of idea carriers. For example, if time t adult mortality rate is low, a young adult survivor has a very low marginal contribution to next period's TFP, as ideas this person passed to the next generation were also passed by many other adult survivors to their children. However, if mortality is high, the marginal contribution of an adult survivor is large, as idea carriers among the young adults are scarce in the next period. The negative impact of one death on TFP thus increases in adult mortality.

2.3 Preferences, constraints, young adults' problem

The time t decisions are made prior to the end-of-period realization of mortality shock ω_t. For a given realization $\omega_{t,j}$, 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

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

where $c_{t,j}^{y}$ and $c_{t+1,j}^{o}\left(\omega_{t,j}\right)$ denote consumption when young and old, respectively. Parameter $\sigma>0$ governs substitutability of the two consumption 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 TFP evolution constraints. We further assume that the utility of the young adult not surviving to old adulthood is given by $U_{t,j}^{NS}=\frac{\left(c_{t,j}^{y}\right)^{1-\sigma}}{1-\sigma}$. The young adult's expected utility associated with a particular realization $\omega_{t,j}$ is therefore

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

When young, adults produce, consume and innovate. When old, adults produce and consume. We abstract from physical capital accumulation in order to maintain tractability and focus on understanding the TFP acceleration, which characterized the economic transformation of England (see Bar and Leukhina 2010a).

The young adult cohort, taking $A_{t,j}$, $\lambda_{t,j}^{y}$, $m_{t-1,j}^{y}$ and frontier productivity $\bar{A}_t:=\max\left\{A_{t,j}\right\}_{j\in\{1,\ldots,J\}}$ as given, chooses consumption $c_{t,j}^{y}$, time allocated to production $l_{t,j}$ and innovation $i_{t,j}$ to solve

$$\begin{aligned}\max_{c_{t,j}^{y},l_{t,j},i_{t,j}}\ E_{\omega_{t,j}}E\left(U_{t,j}|\omega_{t,j}\right)&=\frac{\left(c_{t,j}^{y}\right)^{1-\sigma}}{1-\sigma}+\frac{\beta}{1-\sigma}E_{\omega_{t,j}}\\&\times\left[\left(1-m^{y}\left(\omega_{t,j}\right)\right)\left(c_{t+1,j}^{o}\left(\omega_{t,j}\right)\right)^{1-\sigma}\right]\ \text{s.t.}\ \ c_{t,j}^{y}=A_{t,j}l_{t,j}^{\theta}\left(\lambda_{t,j}^{y}\right)^{1-\theta},\end{aligned}\tag{13}$$

$$A_{t+1,j}\left(\omega_{t,j}\right)=\left(1-m^{y}\left(\omega_{t-1,j}\right)\right)^{\phi}\left(1-m^{y}\left(\omega_{t,j}\right)\right)^{\phi}\left[A_{t,j}+A_{t,j}i_{t,j}^{\eta}+\tau\left(\bar{A}_t-A_{t,j}\right)\right],\tag{14}$$

$$l_{t,j} + i_{t,j} = 1, \tag{15}$$

$$c^o_{t+1,j}\left(\omega_{t,j}\right) = A_{t+1,j}\left(\omega_{t,j}\right)\left(\frac{\lambda^y_{t,j}}{1 - m^y\left(\omega_{t,j}\right)}\right)^{1-\theta}, \tag{16}$$

where in the last constraint we substituted for $\lambda^o_{t+1,j}\left(\omega_{t,j}\right)$ from (9) and used the optimal choice to allocate the entire unit of time to production when old. Constraint (13) states that all output produced when young is used as consumption. Constraint (14) is the law of motion of TFP (12) derived in Sect. 2.2, in which we further assumed, following Acemoglu (2009), that $d_{t,j} = \tau\left(\bar{A}_t - A_{t,j}\right)$, i.e. the TFP increment due to diffusion of knowledge from other regions increases in the gap between the frontier and local technology. The rate of absorption of this gap is denoted by $\tau \in (0, 1)$; it dictates the ease of knowledge flow across regions. It is possible that this parameter is time-dependent, increasing in population density and availability of factors emphasized in Mokyr (2002), such as accessibility to communication devices (railroads, post) and the extent of the availability of scientific communities.[15] We explore the implications of a rising τ in Sect. 3.4.1. Note that, in order to avoid dealing with externalities, we assume the time allocation problem is solved efficiently within the young adult cohort.

2.4 Optimal time allocation

After substituting from the constraints into the objective function and dividing by $A_{t,j}$, the problem simplifies to

$$\max_{i_{t,j}} \frac{1}{1-\sigma}\left[(1-i_{t,j})^{\theta}(\lambda^y_{t,j})^{1-\theta}\right]^{1-\sigma} + \frac{\beta}{1-\sigma}\cdot E_{\omega_{t,j}} \times\left[(1-m^y(\omega_{t,j}))\left((1-m^y(\omega_{t-1,j}))^{\phi}(1-m^y(\omega_{t,j}))^{\phi}[1+x_{t,j}+i^{\eta}_{t,j}]\left(\frac{\lambda^y_{t,j}}{1-m^y(\omega_{t,j})}\right)^{1-\theta}\right)^{(1-\sigma)}\right], \tag{17}$$

where $x_{t,j} = \tau\left(\bar{A}_t - A_{t,j}\right)/A_{t,j}$ denotes the diffusion term. Simplifying further gives

$$\max_{i_{t,j}} \frac{(1-i_{t,j})^{\theta(1-\sigma)}}{1-\sigma} + B_t\frac{\left[1+x_{t,j}+i^{\eta}_{t,j}\right]^{(1-\sigma)}}{1-\sigma}, \tag{18}$$

where $B_t = \beta E_{\omega_{t,j}}\left[\left(1-m^y\left(\omega_{t,j}\right)\right)\left(\left(1-m^y\left(\omega_{t-1,j}\right)\right)\left(1-m^y\left(\omega_{t,j}\right)\right)\right)^{\phi(1-\sigma)}\left(1-m^y\left(\omega_{t,j}\right)\right)^{(\theta-1)(1-\sigma)}\right]$ is the relative weight on future utility, which may differ across regions only due to time $t-1$ mortality realizations. To the extent that mortality histories affect TFP in region j as well as the identity of the frontier region, the diffusion term $x_{t,j}$ is the summary of all histories sufficient for the decision-making. It is instructive to point out that our channel introduces non-linearity to the effective discount rate, as opposed to more standard effects of mortality captured by the linear discount rate $\left(1-m^y\left(\omega_{t,j}\right)\right)$. This non-linearity could potentially change the quantitative impact of gains in longevity on time allocation.

[15] In fact, diffusion of information about new technology is at the heart of the literature focusing on explaining the speed of new technology adoption, in particular, the epidemic models of technology diffusion (Mansfield 1961).

Because labor is supplied inelastically in the last period of life, the decisions of the time t cohort are independent of the time t realization of the mortality shock. Time t parameters of the mortality shock distribution affect the time t decision making only through the expectation term $E_{\omega_{t,j}}[\cdot]$ that enters B_t. Past realizations of mortality, however, affect the time t decision-making through its influence on $x_{t,j}$. Note that the state variables $\lambda^y_{t,j}$ and $N^y_{t,j}$ have no bearing on the optimal time allocation.

$E_{\omega_{t,j}}[\cdot]$ appearing inside B_t is taken over the product of three factors. Formulation (17) elucidates their origin. The first, $D(\omega_{t,j}) := \left(1 - m^y\left(\omega_{t,j}\right)\right)$, represents the discount factor on future consumption due to lifespan uncertainty. As mortality declines and people adjust their expectations accordingly, the weight on future utility increases, which encourages more innovation. The influence of the mortality decline exerted through this factor can be interpreted as the standard Ben-Porath channel, in our formulation appearing as independent of the elasticity of substitution between current and future consumption. The second factor is the knowledge transmission factor $K(\omega_{t-1,j}, \omega_{t,j}) := \left(\left(1 - m^y\left(\omega_{t-1,j}\right)\right)\left(1 - m^y\left(\omega_{t,j}\right)\right)\right)^{\phi(1-\sigma)}$, a function of the rate of TFP transmission across time. Certainly, $\left(1 - m^y\left(\omega_{t-1,j}\right)\right)^{\phi(1-\sigma)}$ can be taken out of the expectation term, as $\omega_{t-1,j}$ is known at time t. We keep it inside so we can talk of the rate of TFP transmission. Declining mortality raises TFP transmission factor, and hence increases the returns to innovation. The impact of declining mortality on time allocation through the knowledge channel represents the *indirect effect* of the knowledge transmission mechanism proposed here. It depends on the substitutability of the two consumption 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 towards current consumption and away from future consumption, the sector experiencing productivity gains. The last factor in the expectation term is the resource dilution channel $M(\omega_{t,j}) := \left(1 - m^y\left(\omega_{t,j}\right)\right)^{(\theta-1)(1-\sigma)}$, appearing due to the assumed land redistribution among the surviving cohort members, and branded Malthusian because of the active role that land per capita played in determining living standards (see Malthus 1798). As mortality declines, people adjust to expect less augmentation of their land holding over the lifecycle in case of survival, which lowers the returns to innovation and, if the consumption goods are substitutes, encourages labor reallocation away from innovating.

The objective function in (18) is strictly concave, which guarantees uniqueness of the solution. The solution is interior, as setting the labor input to zero in either sector implies infinitely large gains from time reallocation.[16] The first order condition characterizing the optimal $i_{t,j}$ is given by

$$\theta\left(1 - i_{t,j}\right)^{\theta(1-\sigma)-1} = B_t \frac{\eta i_{t,j}^{\eta-1}}{\left[1 + x_{t,j} + i_{t,j}^{\eta}\right]^{\sigma}}. \tag{19}$$

The first term gives the marginal benefit from decreasing $i_{t,j}$, i.e. marginal utility from the resulting increase in consumption when young. The second term is the marginal cost of doing so, i.e. the loss, in terms of utility, of future productivity. In the Appendix (Lemma 2), we prove, by examining (19), that $i_{t,j}$ strictly decreases in $x_{t,j}$, that is, the amount of knowledge inflow from other regions discourages innovation.

[16] See the Appendix for a formal proof of uniqueness and interiority (Lemma 1).

2.5 Equilibrium dynamics

Definition 1 For given parameter values, sequence $\{\zeta_t\}$, time-dependent parameters of the mortality shock distribution $\left\{\mu_t^{\omega}, \sigma_t^{\omega}\right\}$, mortality shock realizations $\left\{\omega_{t,j}\right\}$ that imply a history of mortality rates according to (1), birth rates $\left\{n_{t,j}\right\}$ and initial conditions $\left\{N_{0,j}^{y}, N_{0,j}^{0}, A_{0,j}\right\}$, an equilibrium consists of sequences $\left\{N_{t+1,j}^{y}, N_{t+1,j}^{o}, A_{t+1,j}, \lambda_{t+1,j}^{y}, \lambda_{t+1,j}^{o}, c_{t,j}^{y}, c_{t,j}^{o}, l_{t,j}, i_{t,j}\right\}$ such that $\forall t, j$, the young adult's problem is solved and the laws of motion (2), (3), (9) and (14) hold.

The economy starts off with the initial conditions $\left\{N_{0,j}^{y}, N_{0,j}^{0}, A_{0,j}\right\}$. 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 $\left\{N_{t,j}^{y}, N_{t,j}^{o}, A_{t,j}\right\}$, Eqs. 7 and 8 give $\lambda_{t,j}^{y}, \lambda_{t,j}^{o}$. Equation 16 gives consumption of the current old adults $\{c_{t,j}^{o}\}$. Then (4) pins down $N_{t,j}^{c}$. The optimization problem is solved for $c_{t,j}^{y}$ and $i_{t,j}$ in each region. The mortality shocks $\left\{\omega_{t,j}\right\}$ are realized, so the region-specific mortality rates $m^{c}\left(\omega_{t,j}\right)$ and $m^{y}\left(\omega_{t,j}\right)$ are determined for each j according to (1). Given the realizations $\left\{\omega_{t,j}\right\}$, Eqs. 2, 3 and 14 determine the state variables in period $t+1$: $\{N_{t+1,j}^{y}, N_{t+1,j}^{o}$ and $A_{t+1,j}\}$.

2.6 Balanced growth

Consider a deterministic version of our model economy in which $\zeta, n, \mu^{\omega}, \sigma^{\omega}$ are fixed, and mortality rates are identical across time and regions and given by m^{y}. Assuming further that all regions are identical in their initial conditions implies the lack of knowledge diffusion across regions, and hence regions remain identical. It suffices to drop the subscript j. It is useful to discuss a BG equilibrium of this deterministic economy, as it represents the predicted path of our model under fixed $\mu^{\omega}, \sigma^{\omega}$ and $\tau = 0$, as long as m^{y} is related to $\mu^{\omega}, \sigma^{\omega}$ in a way that $(1-m^{y})^{2\phi} = E_{\omega_t}\left[(1-m^{y}(\omega_t))^{2\phi}\right]$ (see Bar and Leukhina 2010b for a formal proof). We employ a BG equilibrium in the calibration exercise below.

Proposition 1 *Denote the solution to Eqs. 20 and 21 by* i^{BG} *and* γ_A^{BG} :

$$\gamma_A = \left(1-m^{y}\right)^{2\phi}\left[1+i^{\eta}\right], \tag{20}$$

$$\theta(1-i)^{\theta(1-\sigma)-1} = \frac{\beta\left(1-m^{y}\right)^{\phi(1-\sigma)} E\eta i^{\eta-1}}{\left[1+i^{\eta}\right]^{\sigma}}, \tag{21}$$

where $E = E_{\omega_t}\left[(1-m^{y}(\omega_t))^{1+(\phi+\theta-1)(1-\sigma)}\right]$. *Assume an arbitrary number* J *of regions. Suppose the initial conditions in each region are identical and satisfy* $N_0^{c} = n^{BG} N_0^{y}$, $\frac{N_0^{o}}{N_0^{y}} = \frac{(1-m^{y})}{(1-\zeta m^{y})n}$, $\lambda_0^{y} = \frac{0.5A}{N_0^{y}}$, $\lambda_0^{y} = \frac{0.5A}{N_0^{y}}$. *Then the equilibrium paths exhibit BG behavior from period* 0 *and onward, that is,* $i_t = i^{BG}$ *and* $\frac{A_{t+1}}{A_t} = \gamma_A^{BG}, \forall t$. *Moreover, the growth rates of all population subgroups, output per capita, and land holdings are given by*

$$\gamma_{pop}^{BG} = \frac{N_{t+1}^{y}}{N_t^{y}} = \frac{N_{t+1}^{o}}{N_t^{o}} = \frac{N_{t+1}^{c}}{N_t^{c}} = \left(1-\zeta m^{y}\right)n, \tag{22}$$

$$\gamma_{y}^{BG} = \gamma_A^{BG}\left(\frac{1}{\gamma_{pop}^{BG}}\right)^{1-\theta}, \tag{23}$$

$$\gamma_{\lambda^y}^{BG} = \gamma_{\lambda^o}^{BG} = \frac{1}{\gamma_{pop}^{BG}}. \tag{24}$$

Proof See the Appendix. □

As seen from (20), we obtain that the rate of TFP growth positively depends on the rate of knowledge transmission across generations, $(1 - m^y)^{2\phi}$, and the time spent on innovation.[17] Relation (23) elucidates that TFP growth positively contributes to the growth of output per capita, while population growth inhibits it, the latter effect capturing the negative impact due to dilution of the non-reproducible factor. Population growth is given by the product of the child survival rate and fertility.

2.7 Effects of temporary and permanent mortality changes

Next we analyze the behavior of the economy with diffusion. To simplify the discussion, suppose the mortality regime is fixed, i.e. μ^ω, σ^ω are constant across time and regions.

The purpose of incorporating the possibility of knowledge diffusion across regions into the model was to moderate the adverse impact of a mortality shock on TFP. Indeed, with no diffusion ($\tau = 0$), the optimal time allocation is fixed by the solution to (19) with $x_{t,j}$ set to zero. So, time allocation is independent of the particular realizations of mortality, the history of mortality shocks affecting only the law of motion of TFP and land. Hence, a region-specific large-scale epidemic, i.e. a relatively high mortality shock, would imply a permanent negative impact on this region's TFP in all future periods.

In a model with diffusion, the impact of a region-specific large-scale epidemic on the level of TFP is only temporary. To be precise, suppose two regions experience identical mortality realizations equal to m^y, except that at some point, region 2 is hit with a one-time (unexpected) higher mortality. Prior to the epidemic, both regions are on a BG characterized in Proposition 1, and no knowledge flows across regions. Immediately after the epidemic, TFP in region 2 is lower, and hence knowledge inflow into region 2 becomes positive. The effect of the presence of knowledge inflow is to discourage innovation (relative to the BG/leader innovation time), but this indirect negative impact less than offsets additional TFP growth due to the knowledge inflow (Lemmas 2 and 3 in the Appendix). Although the follower's (region 2) TFP always remains behind that of the leader, its TFP grows faster, in the long run converging, in relative terms, to the leader's TFP (Proposition 2 in the Appendix). Hence, allowing for diffusion in our model changes the negative impact of a large-scale, region-specific epidemic on TFP from permanent to temporary. This simple example well captures a more general intuition that differential mortality histories tend to increase differences in TFP while the presence of diffusion works to reduce them.

It also follows from the above discussion that in the long run, when differences in TFP, and hence labor inputs, disappear, standards of living will be unambiguously higher in region 2 due to the so called positive Malthusian check of per capita land holdings (higher $(\lambda^y)^{1-\theta}$). Our calibration (discussed below), which will employ real wage response to the Black Death to identify ϕ, will imply that for moderately high shocks, even in the short run, the positive Malthusian check dominates the negative impact on TFP growth.[18] Hence, unexpected

[17] Note the growth rate of productivity is independent of the population size, i.e. there are no scale effects, criticized in Jones (1995).

[18] This is true for shocks implying $m_{t,j}^y < \frac{(1-\theta)\zeta-\phi}{(1-\theta)\zeta-\phi\zeta} = 0.83$. Because of the convexity of knowledge destruction, for sufficiently high shocks, the direct impact on TFP would dominate in the short run. For details, see Bar and Leukhina (2010b).

epidemics would increase output per capita growth, and would appear as a "positive check" to an observer like Thomas Malthus.

Next we generalize our discussion first to J regions and then to J regions and uncertain future mortality shocks (in a footnote). In an economy with J identical regions, different mortality shock realizations across regions immediately lead to TFP differences. We examine how the time t distribution of TFP levels $\left\{A_{t,j}\right\}_{j=1}^{J}$ evolves if, after time t, all regions experience mortality rates fixed across time and regions, $m_{t+k,j}^{y} = m^{y}\ \forall k = 1, 2$. We show that for any region pair i, j, if $A_{t,i} > A_{t,j}$, then $A_{t+k,i} > A_{t+k,j}$, $\frac{A_{t+k+1,i}}{A_{t+k+1,j}} < \frac{A_{t+k,i}}{A_{t+k,j}}$ and $\lim_{k\to\infty} \frac{A_{t+k,i}}{\bar{A}_{t+k}} = 1$ (Proposition 2). The first result establishes that the ranking of TFP levels, and in particular, the leader's identity, remain unchanged. In fact, the leader's TFP remains on a BG path given by $\bar{A}_{t+k} = (1-m^{y})^{2k\phi}\left[1+\bar{\imath}^{\eta}\right]^{k}\bar{A}_{t}$, where $\bar{\imath}$ is defined as the solution to (19) in which $x_{t,j}$ is set to zero (Corollary 3). The second and third results establish that all TFP levels converge, in relative terms, to the frontier TFP, with regions farther behind growing faster.[19] It follows that time t frontier TFP path determines the long run TFP level for all regions. Note that it follows from our discussion that a large-scale epidemic will have a permanent effect on long-run TFP levels only if it is widespread, affecting all regions, and hence, frontier TFP.

One result we established is that for a fixed mortality regime, our calibration implies that an unexpected moderately high temporary mortality shock would increase output per capita growth. The opposite would be true for a surprise low mortality rate. A permanent expected rise in mortality, however, works very differently. Upon the realization of the shock, TFP growth suffers more relative to the case of the temporary/unexpected rise in mortality, because in addition to the decline in the knowledge transmission factor (direct effect), less time is allocated to innovation (indirect effect). Although the positive Malthusian check is just as large, the negative impact on TFP growth may be strong enough to reverse it. We will show in Sect. 3.3 that permanent declines in mortality, by increasing knowledge transmission and by encouraging innovation, worked to increase output per capita growth, despite their influence on land holdings.

Finally, we obtain an intuitive result that increasing the number of regions in the economy with diffusion, further moderates the negative impact due to mortality. It increases expected leader's TFP, and therefore, expected TFP in any region, at any given time in the future. To be precise, consider the effects of adding one region on $A_{t,j}$. Intuitively, adding one history of mortality draws to any J histories can be at worst ineffective (if the history is such that the additional region is never a leader up to time t). However, the additional region may possess the frontier TFP at some point before t, which would positively impact $A_{t,j}$ for all j.

[19] With uncertain future mortality shocks, we examine how *expected* future TFP level distribution depends on the current distribution $\left\{A_{t,j}\right\}_{j=1}^{J}$, obtaining the results that closely correspond to the results applying to the deterministic case discussed above. In Bar and Leukhina (2010b), we formally establish the effect of the initial TFP distribution on k-period ahead predicted TFP distribution, holding the $t-1$ mortality fixed across locations. We establish that TFP ranking is predicted to remain, that TFP differences shrink over time, and finally, that in the long run, the predicted TFP levels of all locations converge in relative terms to the predicted value of frontier TFP, the evolution of the latter given by $E\left(\bar{A}_{t+k}|\left\{A_{t,j}\right\}_{j=1}^{J}, m_{t-1}^{y}\right) = E\left[\left(1-m_{t}^{y}\right)^{\phi}\right] E\left[\left(1-m_{t}^{y}\right)^{2\phi}\right]^{k-1}\left(1-m_{t-1}^{y}\right)^{\phi}\left[1+\bar{\imath}^{\eta}\right]^{k}\bar{A}_{t}$ for $k = 1, 2, \ldots$

3 Quantitative analysis

3.1 Estimating the parameters of the mortality shock distribution

Because the main question that we ask is how changes in the frequency and severity of epidemics affected the process of output takeoff in England, we need to estimate $\left\{\mu_t^{\omega}, \sigma_t^{\omega}, \zeta_t\right\}_{t=1600}^{2000}$. We take a stand that each period in the model is 25 years. We gather the data on country-level mortality rates for the age groups 0–25 and 25–50 from Wrigley et al. (1997) and the Human Mortality Database for more recent years.[20] The data is available at 10-year frequency. We need to make the appropriate adjustments to accommodate the 25-year period length in the model. We first create a time series $\ln \omega_{1600-09}, \ln \omega_{1610-19}, \ldots$ based on the 10-year frequency data of 25–50 mortality according to (1) i.e. $\ln \omega_t = \ln\left(\frac{m_t}{1-m_t}\right)$. Next, for each $t = 1600, 1625, \ldots$, we estimate μ_t^{ω}, i.e. the mean of $\ln \omega_t$, as the sample (weighted) average over the relevant two and a half decades.[21] We find that estimated μ_t^{ω} fell from -0.40 to -4.98, the trajectory of the fall closely resembling the trajectory of m_t^{25-50} plotted in Fig. 1, with the sharpest fall occurring during 1780–1840.

To obtain the estimates for σ_t^{ω}, standard deviation of the log mortality shock, $\ln \omega_t$, we first derive variance estimates of $\left(\sigma_{1600-24}^{\omega}\right)^2$ and $\left(\sigma_{1975-99}^{\omega}\right)^2$ based on the data for 1600–1650 and 1950–2000 respectively. To estimate $\left(\sigma_{1600-24}^{\omega}\right)^2$, for example, we record the variance of $\ln \omega_{1600-9}, \ldots, \ln \omega_{1650-59}$ and multiply it by $10J/25$ in order to adjust for the number of regions and convert 10 year frequency into 25 year frequency. In other words, because we are using time variation of country-wide average mortality, our estimate of the region-specific variance positively depends on the number of regions we choose to model. To obtain the remaining estimates of $\left(\sigma_t^{\omega}\right)^2$, we linearly interpolate between the end-point variance estimates. We obtain that the ratio of its standard deviation to the absolute value of the mean (i.e. coefficient of variation) declined by 90%.

We estimate $\{\zeta_t\}$ by taking the ratio of the observed mortality rates for the age groups 0–25 and 25–50, reported in Fig. 1, obtaining a hump-shaped series. The ratio starts at 0.82, grows initially because the mortality rate of the age group 0–25 is roughly fixed in the data, while the adult mortality declines. It eventually decreases because in this period child mortality falls faster than the adult mortality.

3.2 Calibration of the benchmark model

One goal of this study is to quantitatively assess the importance of our hypothesis. In order to accomplish this goal, we parameterize the model to match several key features of preindustrial England assuming that the behavior of the seventeenth century England can be accurately captured by the BG behavior of the model economy. We thus employ equations from Proposition 1, where we relate m^y to the seventeenth century estimates of $\mu^{\omega}, \sigma^{\omega}$ obtained in Sect. 3.1 so that $(1-m^y)^{2\phi} = E_{\omega_t}\left[(1-m^y(\omega_t))^{2\phi}\right]$. This ensures that BG equilibrium of the deterministic economy represents the predicted path of our model under fixed $\mu^{\omega}, \sigma^{\omega}$ and $\tau = 0$ (see Bar and Leukhina 2010b for a formal proof). Because all the regions are identical ex-ante and hit with an identical shock, there will be no knowledge flows across regions, and the regions will remain identical. It suffices to ignore the subscript j.

[20] For the list of all data sources used in this paper, refer to the Appendix.

[21] For example, we estimate $\mu_{1600-25}^{\omega} = 0.4 \ln \omega_{1600-09} + 0.4 \ln \omega_{1610-19} + 0.2 \ln \omega_{1620-29}$.

Thus, from the estimation of the mortality process parameters, we have $m^y = 0.4$ and $\zeta = 0.82$. We also construct the general fertility rate (GFR)[22] series using the demographic data in Wrigley et al. (1997), Mitchell (1975) and the Human Mortality Database. GFR is around 134 in the seventeenth century England, which implies $n = (134/1{,}000)\,25/2 = 1.675$, where the adjustment is made to the 25-year frequency and the fact that our model is unisex. Then, according to (22), the model implies population growth to be $\gamma_{pop} = 1.125$, or 0.47% annual growth rate, which is roughly consistent with the population growth estimates in Wrigley et al. (1997).

Clark (2001b) reports that the land share in total income during 1600–1700 is around 0.3, so we set $\theta = 0.7$. From Clark (2001b), output per capita in the 17th century rose at about 0.176% per year, which implies the growth rate of 1.045 per 25 years. Then (23) implies the model must predict TFP growth along the calibrated BG path to be $\gamma_A^{BG} = 1.082$, or 0.3% annual growth.

Wrigley et al. (1997) documents the total population size around 1600 to be 4,109,981. Hence, $N_0 = 4{,}109{,}981/J$ per region. Since we do not have the data on the population age composition in the seventeenth century, we employ the BG implication of the constant age structure, i.e. $N_0^c/N_0^y = n$ and $N_0^o/N_0^y = \frac{1-m^y}{(1-\zeta m^y)n}$. We already determined the mortality and fertility rates, which allows us to determine the population structure along the calibrated BG. We must set $N_0^y = 4{,}109{,}981/J/(1 + n + \frac{1-m^y}{(1-\zeta m^y)n}) = 4{,}109{,}981/J/(1 + 1.675 + \frac{1-0.4}{(1-0.7(0.4))1.675}) = 1{,}281{,}223/J$, which implies 31% of the total population are of ages 25–50. It also follows that $N_0^o = 682{,}710/J$ per region.

Next we calibrate ϕ, an important parameter that governs the rate of knowledge destruction due to mortality. Along the calibrated BG, realizations of mortality shocks equal their expected value every period, and marginal (and average) products of labor of the young and the old grow at the rate of $\left(A_{t+1}\lambda_{t+1}^{1-\theta}\right) / \left(A_t\lambda_t^{1-\theta}\right) = \gamma_A^{BG}\left((1-\zeta m^y)\,n\right)^{\theta-1} = 1.045$. However, a surprise mortality shock, higher than its expected value, will cause the marginal (and average) products of labor to rise, as it will raise the per capita land holdings among 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 the case of perfect knowledge transmission across time. How much less depends on ϕ. To identify this parameter, we use the documented response of real agricultural wages to the Black Death epidemic in England. The Black Death epidemic appears to be the most useful case study for this purpose. The epidemic was not expected; people of prime age were infected and died quickly; physical capital remained intact.[23]

Suppose a period t^* (unexpected) mortality rate, m_{t^*}, affects children and young adults indiscriminately. Clark (2001a) estimates that 25 years after the Black Death, population of England was reduced by approximately 45%. We compute the (unexpected) mortality rate m_{t^*}, that would imply a comparable destruction of the population in the calibrated model economy, by solving

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

[22] The general fertility rate is the number of births in a given year per 1,000 females of ages 15–44.

[23] Wars would not be as appropriate for identifying the key parameter of the knowledge transmission mechanism, as they tend to destroy physical capital too, which we do not model here. Epidemics that involved a long and perhaps not as unpleasant a period of infection before death, would not be as appropriate either, because there would be time for the infected parent to expedite their skill transmission to the children or make suitable alternative learning arrangements.

where we used the population age structure implied by BG. The above implies $m_{t^*} = 0.678$. Because of the unexpected nature of the shock and because the shock equally affects all regions (so there is no knowledge diffusion), the optimal time allocation is unaffected, i.e. $i_{t^*} = i_{t^*+1} = i^{BG}$. It follows that, in response to m_{t^*}, the calibrated model implies the following rise in the marginal (and average) product of labor of the young adults:

$$\frac{A_{t^*+1}\lambda_{t^*+1}^{y1-\theta}}{A_{t^*}\lambda_{t^*}^{y1-\theta}} = \left(\frac{1-m_{t^*}}{1-m^y}\right)^{\phi} \gamma_A^{BG} \left(\frac{N_{t^*+1}^y}{N_{t^*}^y}\right)^{\theta-1} = \left(\frac{1-m_{t^*}}{1-m^y}\right)^{\phi} \gamma_A^{BG} \left((1-m_{t^*})\, n\right)^{\theta-1}, \tag{25}$$

where we used (20). Clark (2001a) presents data on real agricultural wages during the 14th c., revealing that 25 years after the Black Death, real wages were 20% higher than their preplague level. Setting (25) to 1.2 gives $\phi = 0.132$.

Our procedure of identifying ϕ implies a very low rate of knowledge destruction: even under the high mortality regime of the seventeenth century, $m^y = 0.4$, only 12.6% of productivity fails to be transmitted from one 25 year generation to the next. Even the Black Death episode, that destroyed nearly half the population, implied only a 19.5% level of productivity destruction. The smaller the value of ϕ, the smaller will be the effect of gains in life expectancy. What we aim to identify is the influence of these quantitatively small gains in knowledge transmission on the course of development.

It is difficult to map and measure in the data the time spent on innovation i^{BG}. We set it to a low value of 0.05 for the seventeenth century England. Then the BG path Eq. 20 can be solved to determine its remaining unknown: $\eta = 0.48$. This estimate is consistent with most available estimates of labor elasticities in human capital production functions. Erosa et al. (2009), for example, estimate it to be 0.52.

Setting $\beta = \left(\frac{1}{1.02}\right)^{25} = 0.6$, we can solve the final BG Eq. 21 for σ :

$$\theta l^{\theta(1-\sigma)-1} = \frac{\beta\,(1-m^y)^{\phi(1-\sigma)}\, E\eta i^{\eta-1}}{[1+i^{\eta}]^{\sigma}}.$$

Note that E depends on σ and, given the distributional assumptions on ω_t, E does not have a closed form solution. We solve the above equation numerically and obtain $\sigma = 0.6$ as a unique solution.[24]

Finally, we normalize $A_{0,j} = 1$ and $\Lambda = 10{,}000$.[25] Table 1 summarizes the calibrated parameters.

[24] A word of caution is needed to warn the reader about comparing the implied intertemporal elasticity of substitution (IES) $\frac{1}{\sigma}$ to the estimates found in the real business cycles literature. First, the existing estimates of the IES are mixed, although most papers tend to use $\frac{1}{\sigma} < 1$. Second, the existing macro estimates of the IES are based on an Euler equation from the neoclassical growth model, which is obviously very different from our model, in the setup and length of period. Finally, Gruber (2006) 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 the IES implied by our calibration. We explore the implications of setting $\sigma > 1$ in Sect. 3.4.2.

[25] The normalization to a large number here is to ensure that land per capita is not too small a number.

Table 1 Calibrated parameter values

Consumption technology	$A_0 = 1, \Lambda = 10{,}000, \theta = 0.70$
Idea production technology	$\eta = 0.48, \phi = 0.132$
Preferences	$\sigma = 0.6, \beta = 0.6$
Mortality	$m^y = 0.4, \zeta = 0.82$

3.3 Results

We now use the calibrated model to examine the impact of the adult mortality decline on output per capita takeoff in England, and in particular, the impact due to the link proposed here. As discussed in Sect. 2.7, before the onset of permanent mortality changes, a surprise high mortality shock would increase output per capita growth, and would appear as "a positive check" to an observer like Thomas Malthus. The opposite would be true for a surprise low mortality rate. However, as we will show next, permanent declines in mortality, by increasing knowledge transmission across time through factor $\left(1 - m^y\left(\omega_{t-1,j}\right)\right)^{\phi}\left(1 - m^y\left(\omega_{t,j}\right)\right)^{\phi}$ and by encouraging innovation through its influence on $K\left(\omega_{t-1,j}, \omega_{t,j}\right)$, worked to increase output per capita growth. This qualitative change is not built into the model; it arises due to our calibration.

3.3.1 The overall effect of mortality decline on output takeoff

In this main experiment, we simulate the model subject to the exogenously changing parameters of the mortality shock process, obtained in Sect. 3.1, and GFR. In other words, we vary $\left\{\mu_t^{\omega}, \sigma_t^{\omega}, \zeta_t, n_t\right\}$ according to our estimation, simulate location-specific mortality shock histories, solve for the equilibrium dynamics and track the magnitude of the resulting change in per capita output and other quantities.[26] While these inputs are identical across regions, regions will differ in their history of mortality shock realizations.

Although we solve the model for the time period corresponding to [1600, 2000], we report the results by zooming in on the two-hundred year period 1680–1880, the midpoint of which dates the beginning of the industrial revolution and the endpoint of which dates the beginning of the fertility transition.

In the calibrated deterministic version of the model, all regions were identical, and hence no knowledge diffused across regions. With differential mortality shocks across regions, introduced by this simulation, regions become different within one period, and hence knowledge begins to flow across regions. There are two effects that allowing for knowledge diffusion introduces in our model economy. First, it moderates the negative impact of a high mortality episode on productivity, making it only short-lived as discussed in Sect. 2.7. Second, because a knowledge inflow increases productivity of labor when old, it lowers the marginal utility of future consumption, consequently discouraging innovation.

To simulate the economy, we need to choose J and τ. We find that the volatility of the growth rate of total population in the country is decreasing with the number of regions assumed. We choose $J = 5$ to match the coefficient of variation of population growth of 0.048 in the seventeenth century. We set the absorption rate to a low level of 0.3, which

[26] We set $n_t = 25GFR_t/2{,}000$ to adjust for the 25-year frequency and the fact that the model is unisex.

allows to have a low level of variation of region-specific time investments in innovation.[27] In Sect. 3.4.1, we will investigate the quantitative implications of a rising rate of absorption. As discussed in Sect. 2.7, in the presence of diffusion, expected output growth increases in the number of regions. Our quantitative simulations, however, reveal that the model with 5 regions and diffusion still closely captures the seventeenth century output dynamics.

Figure 2 reports the results from one such simulation. Panel a plots the log of output per capita index in each of the five regions, together with the log of output per capita index in the data. The presence of differential time paths of per capita output across regions indicates the influence of differential histories of mortality shock realizations.

Column 3 of Table 2 summarizes the results from averaging over 1,000 simulations. The main experiment generates about 90% of the increase in the log of output per capita observed in the data ($0.8337/0.9172 = 0.9$). Panel c of Fig. 2 plots the population dynamics in the data and its model counterpart. Since we directly fed in our estimates of GFR and we modeled mortality dynamics in an attempt to reflect the actual mortality experience, it is not surprising that the model population dynamics matches its empirical counterpart very well. Nonetheless, it validates our estimation of the mortality shock process.

Panel b of Fig. 2 reports the time spent on innovation in each region. There is little difference in innovation activity across regions because of the low rates of knowledge diffusion. Panel b reveals that the mortality decline generates a substantial rise in the time spent on innovation (from around 5% to 10%) due to the rise in the expectation term, reflected in the plot of the relative weight on future utility B_t reported in panel d. Recall the expectation term inside B_t consists of three factors that we referred to as $D\left(\omega_{t,j}\right)$, $K\left(\omega_{t-1,j}, \omega_{t,j}\right)$ and $M\left(\omega_{t,j}\right)$. 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 overall rise in the weight on future utility.

Panels e and f reveal the forces behind the takeoff in per capita output produced by this experiment. All the quantities reported are averages over the regions. Panel e reports the growth rates of the components of output per young adult y_t^y, that is, the growth rates of A_t, l_t^θ and $\left(\lambda_t^y\right)^{1-\theta}$. Recall that the old adults enjoy the same productivity A_t, while the growth rate of their labor input is 1 by construction, and the growth rate of λ_t^o closely resembles the growth rate of λ_t^y, and therefore omitted from the graph. This panel reveals that gains in TFP were the main force behind the output per capita takeoff, while the dilution of land among the growing population worked in the opposite direction. Column 3 of Table 2 decomposes the change in output per young adult during the 200-year period into changes of its components: resource dilution implied that in the absence of other changes, output per young adult should shrink by 40% ($(\lambda_{1880}/\lambda_{1680})^{1-\theta} = 0.5949$), the declining labor input implied a further 4% drop in y_t^y, while TFP grew by the factor of 3.96, more than offsetting the negative impact. Together these changes implied that y_t^y rose by a factor of 2.26.

Column 3 of Table 2 further decomposes the growth rate of A_t into $\left(1-m_{t-1}^y\right)^\phi \left(1-m_t^y\right)^\phi$ and $[1+i_t^\eta+\tau\left(\bar{A}_t/A_t-1\right)]$, the two terms inside the brackets reported separately. The growth rate of A_t increased from 10% to 32%. This increase came about due to both, the increase in the rate of knowledge transmission, i.e. the first component, from 0.88 to 0.9855 (a 12% rise), and an increase in the innovation increment i_t^η from 0.24 to 0.33 (a 40% rise), with the innovation increment exerting a larger influence. Panel f of Fig. 2 plots the changes of these two components. The contribution of the knowledge inflow from the frontier region to TFP

[27] Keeping the leader productivity fixed, $\tau = 0.3$ implies that even after 100 years of knowledge diffusion, the gap between the leader and any follower still remains at 24% of the original difference.

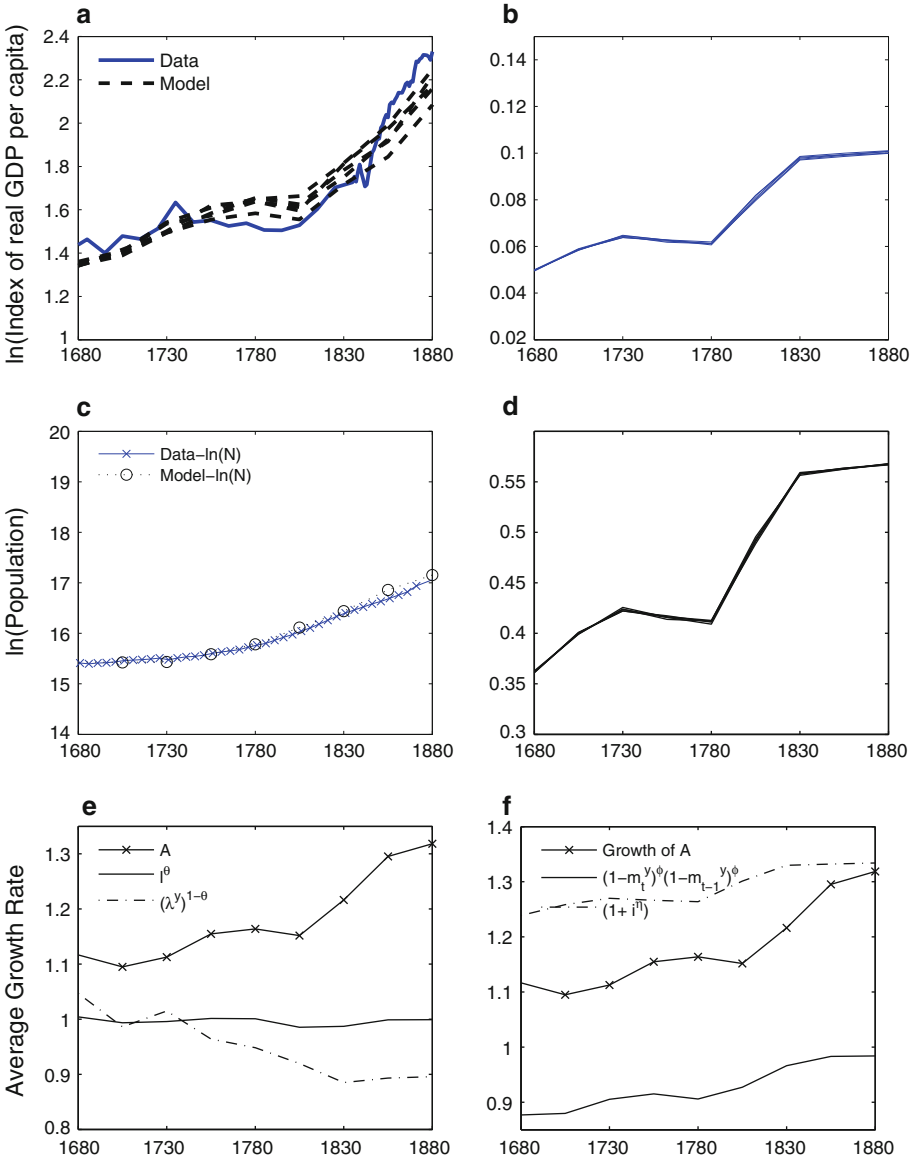

Fig. 2 Main experiment. **a** Main, **b** fraction of time spent innovating, **c** population dynamics, **d** relative weight on future utility, **e** sources of growth of young adults' output, and **f** sources of growth of A

growth was small in 1680 at 0.01 and was further reduced to 0.006 in 1880 as the relative TFP gap declined.[28]

[28] Even though the contribution of the diffusion term $\tau\left(\bar{A}_t/A_t - 1\right)$ to TFP growth is small, allowing for knowledge diffusion across locations is not quantitatively irrelevant. To see this, consider subjecting all regions to identical mortality histories, as implied by the evolution of the mean mortality, $\left\{E_{\omega_t} m^y(\omega_t)\right\}$, in accordance with our estimates of the mortality process. In this case, there is no knowledge diffusion across locations,

Table 2 Results

	Data	Main exp.	Main & fixed destruction of A	Main & increasing τ
$\ln y_{1880} - \ln y_{1680}$	0.9172	0.8337	0.3430	0.8473
$\frac{y^y_{1880}}{y^y_{1680}}$		2.2569	1.3853	2.2891
$\frac{A_{1880}}{A_{1680}}$		3.9625	2.4147	4.0100
$\left(\frac{l_{1880}}{l_{1680}}\right)^{\theta}$		0.9623	0.9673	0.9623
$\left(\frac{\lambda_{1880}}{\lambda_{1680}}\right)^{1-\theta}$		0.5949	0.5956	0.5961
$\frac{A_{1705}}{A_{1680}}$		1.0996	1.0824	1.1011
$\left(1-m^y_{1655}\right)^{\phi}\left(1-m^y_{1680}\right)^{\phi}$		0.8800	0.8733	0.8908
$1+i^{\eta}_{1680}$		1.2394	1.2395	1.2392
$\tau\left(\bar{A}_{1680}/A_{1680}-1\right)$		0.0092	0	0.0107
$\frac{A_{1905}}{A_{1880}}$		1.3203	1.1560	1.3215
$\left(1-m^y_{1855}\right)^{\phi}\left(1-m^y_{1880}\right)^{\phi}$		0.9855	0.8733	0.9843
$1+i^{\eta}_{1880}$		1.3345	1.3237	1.3343
$\tau\left(\bar{A}_{1880}/A_{1880}-1\right)$		0.0060	0	0.0076

3.3.2 The importance of the knowledge transmission mechanism

In order to assess the importance of our hypothesis, that is, to single out the impact of adult mortality on takeoff through the knowledge transmission mechanism proposed here, we perform an additional experiment. The impact of gains in adult longevity on takeoff through the knowledge transmission mechanism transpires through two different channels: its direct effect through the knowledge transmission factor $\left(1-m^y_{t-1}\right)^{\phi}\left(1-m^y_t\right)^{\phi}$ in the law of motion of TFP (14) and its indirect effect on time allocation through its influence on factor $K\left(\omega_{t-1,j},\omega_{t,j}\right)$ in the expectation term. How much of the total influence of gains in adult mortality established by the main experiment transpired through these two channels? To answer this question, we simulate the economy again, except this time we fix the channels in question at their 1600 level. We then examine how much of the overall influence of mortality on takeoff established by the main experiment is eliminated.

Operationally, we vary $\left\{\mu^{\omega}_t,\sigma^{\omega}_t,\zeta_t,n_t\right\}$ according to our estimation, but we do not allow changes in $\mu^{\omega}_t,\sigma^{\omega}_t$ to affect the factor $K\left(\omega_{t-1,j},\omega_{t,j}\right)$ in the expectation term and the factor $\left(1-m^y\left(\omega_{t-1,j}\right)\right)^{\phi}\cdot\left(1-m^y\left(\omega_{t,j}\right)\right)^{\phi}$ in the law of motion of TFP. Instead, we set these factors to their levels implied by the mortality rate at its expected value in 1600. The remaining effects of mortality changes are those on the expectation terms through factors $D\left(\omega_{t,j}\right)$ and $M\left(\omega_{t,j}\right)$ and on population (and hence land) dynamics.

Footnote 28 continued
and locations remain identical. This variant of the main experiment, which shuts down heterogeneity across locations and hence the role of knowledge diffusion, yields an increase in the log of output per capita of 0.78, accounting for a slightly smaller part (85%) of the empirical rise in the log of per capita output. Thus, allowing for knowledge diffusion across regions, which introduces spillover effects from the leader onto the followers' growth, does help explain an extra 5% of the rise in per capita output.

The difference in mortality histories across regions, generated in this simulation, translate into differences in population, and therefore, land dynamics, but not into differences in TFP dynamics. The reason for the identical TFP dynamics across regions is the following. Initially, TFP is identical across regions, so none of the regions experience a knowledge inflow. Time allocation is identical across regions as it can differ only due to differences in E or knowledge inflows. Since the direct effect of mortality on TFP evolution is fixed at the same level for all regions, TFP remains identical to the next period, and so on.

The results for one arbitrary simulation are reported in Fig. 3, while column 4 of Table 2 reports important statistics obtained by averaging over 1,000 simulations. Panel a of Fig. 3 reveals that the influence of mortality changes on output takeoff is significantly weakened. The rise in the log of output per capita obtained in this experiment accounts for only 38% of its rise in the data, indicating that the influence exerted through our knowledge transmission mechanism accounts for approximately half (52%) of the empirical takeoff.

Panel d of Fig. 3 reveals that, relative to the main experiment, the rise in the relative weight on future utility is only slightly smaller; this is because ϕ (and hence the influence of mortality through factor $K\left(\omega_{t-1,j}, \omega_{t,j}\right)$ present in the main experiment but eliminated in this experiment) is very small in our calibration. Consequently, the rise in the time spent innovating is only slightly smaller: i_t increases from 5 to 9% (as opposed to 5–10% rise observed in the main experiment).

Note that because the effect of mortality changes on population dynamics is identical to that in the main experiment, population and land plot dynamics are unchanged: resource dilution again implies that output should shrink by 40%. Again, it is the gains in TFP growth that drive the output per capita takeoff (panel e). The output takeoff obtained in this experiment is weaker essentially because TFP rises by a smaller factor (2.4 instead of its 3.96 counterpart in the main experiment), its growth rate rising only to 1.16 (instead of 1.32 observed in the main experiment). In turn, the growth rate of A exhibits a smaller rise because the direct contribution of the knowledge transmission channel is entirely shut down (direct effect) and because the innovation increment rises less (by 35% instead of 40% obtained in the main experiment) as we shut down the influence of mortality changes on $K\left(\omega_{t-1,j}, \omega_{t,j}\right)$ (indirect effect).[29]

In other words, we found that shutting down the influence of mortality through the two knowledge transmission channels substantially inhibits potential economic growth. Much of the growth generated in the main experiment is lost, and the loss is almost entirely due to inhibited growth of TFP. In turn, the TFP acceleration was weakened mainly due to the lack of influence of mortality reductions through their direct effect.

Overall, we learn that despite our calibration procedure identifying a very low level of ϕ, gains in adult mortality transpiring through the knowledge transmission mechanism account for approximately one half of the rise in the log of output per capita index. These findings suggest that, in addition to their more conventional effects, reductions in adult mortality substantially contribute to the process of development by improving the process of knowledge transfer across generations.

[29] None of the TFP growth is due to the knowledge diffusion across regions, because TFP dynamics is identical across regions, as explained above.

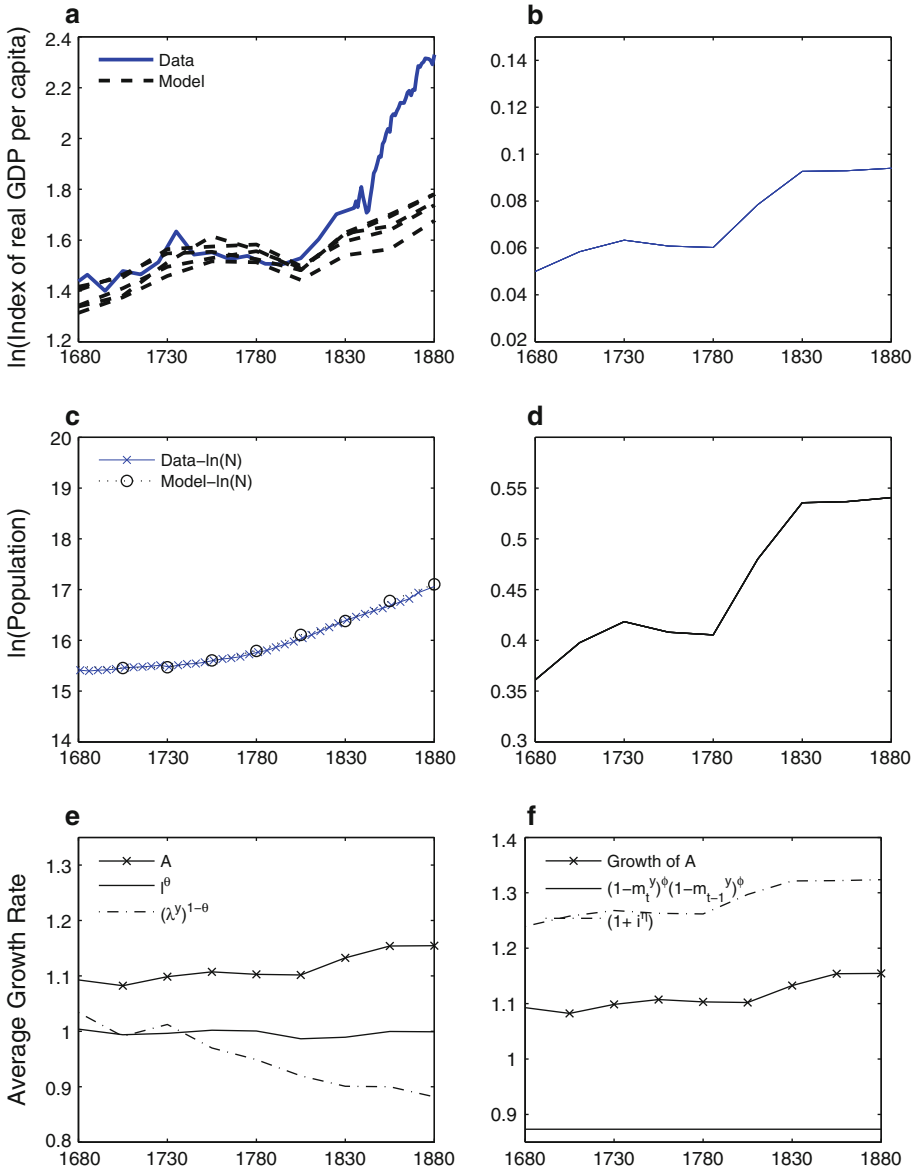

Fig. 3 Main experiment with knowledge destruction of 1600. **a** Main & knowledge destruction 1600, **b** fraction of time spent innovating, **c** population dynamics, **d** relative weight on future utility, **e** sources of growth of young adults' output, **f** sources of growth of *A*

3.4 Robustness analysis

3.4.1 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 TFP. This potentially important influence of mortality

decline is through its influence on population density and consequently the rate of knowledge absorption across regions. To capture the idea that higher population density facilitates idea exchange across regions, we set

$$\tau_t := \frac{\left(N_t^y + N_t^o\right)/\Lambda}{\psi + \left(N_t^y + N_t^o\right)/\Lambda}, \tag{26}$$

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

To maintain $\tau_{1600} = 0.3$, given the initial population age structure, we set $\psi = 500$. Next, we repeat the main experiment, except that now we allow for τ_t to evolve according to (26). Over the time period examined, as population density rises, τ_t increases from 0.3 to 0.76. The last column of Table 2 reveals that this large change in the rate of absorption, however, has an almost negligible effect on the dynamics generated by the model, attributing only a slightly larger role to changes in mortality. Intuitively, the rate of knowledge absorption increases the rate of catching up to the frontier TFP and hence reduces the productivity gap across regions, thus lowering potential gains from increasing the rate of absorption in the future. Moreover, mortality shocks also become less severe over time, also offsetting the potential gains from increasing the rate of knowledge absorption.[30]

Note that because the second experiment performed in Sect. 3.3.2 implied identical TFP across regions, this change in τ would not alter the results of that experiment. Thus, our conclusions regarding the importance of the knowledge transmission mechanism would remain the same.

We also chose to investigate the implications of changing J. For a higher J, as discussed in Sect. 2.7, the path of leader TFP is higher and grows faster. Changing the mortality regime in a setting with a steeper leader path has a larger influence on other locations. We repeated our simulation for $J = 1$, obtaining that the main experiment generates slightly less (85% instead of 90%) of the increase in GDP per capita growth. Because the second experiment is unchanged, we attribute a slightly smaller (47% instead of 52%) to the direct influence of the knowledge transmission mechanism. For a larger number of locations, our findings would be reinforced.

3.4.2 Sensitivity analysis with respect to σ

We already discussed how the choice variables are affected by σ. Next, we study how our quantitative results are affected if we adjust the calibration procedure to deliver $\sigma > 1$. If we set the time spent innovating to 0.0385 instead of 0.05 as in the benchmark model, then our calibration procedure yields $\eta = 0.44$ (η must drop to produce a larger σ) and $\sigma = 1.3$. Increasing σ actually slightly (but almost negligibly) increases the rise in the expectation term implied by declining mortality for the following reasons. In the benchmark model, the influence of gains in longevity exerted through the knowledge transmission channel $K(\omega_{t-1,j}, \omega_{t,j}) := \left(1 - m^y\left(\omega_{t-1,j}\right)\right)^{\phi(1-\sigma)}\left(1 - m^y\left(\omega_{t,j}\right)\right)^{\phi(1-\sigma)}$ was positive, while the influence exerted through the resource dilution channel $M(\omega_{t,j}) := \left(1 - m^y\left(\omega_{t,j}\right)\right)^{(\theta-1)(1-\sigma)}$ was negative. As the present and future consumption goods become complements, both effects are reversed,

[30] We do find that this experiment generates a lower variation in experiences of different regions.

Table 3 Results, $\sigma = 1.3$, $\eta = 0.55$

	Data	Main exp.	Main & fixed destruction of A
$\ln y_{1880} - \ln y_{1680}$	0.9172	0.7865	0.3382
$\frac{y_{1880}}{y_{1680}}$		2.1655	1.3828
$\frac{A_{1880}}{A_{1680}}$		3.7409	2.3947
$\left(\frac{l_{1880}}{l_{1680}}\right)^{\theta}$		0.9759	0.9734
$\left(\frac{\lambda_{1880}}{\lambda_{1680}}\right)^{1-\theta}$		0.5962	0.5959
$\frac{A_{1705}}{A_{1680}}$		1.0986	1.0824
$\left(1 - m^{y}_{1655}\right)^{\phi} \left(1 - m^{y}_{1680}\right)^{\phi}$		0.8806	0.8733
$1 + i^{\eta}_{1680}$		1.2377	1.2395
$\tau \left(\bar{A}_{1680}/A_{1680} - 1\right)$		0.0090	0
$\frac{A_{1905}}{A_{1880}}$		1.2989	1.1531
$\left(1 - m^{y}_{1855}\right)^{\phi} \left(1 - m^{y}_{1880}\right)^{\phi}$		0.9846	0.8733
$1 + i^{\eta}_{1880}$		1.3129	1.3204
$\tau \left(\bar{A}_{1880}/A_{1880} - 1\right)$		0.0060	0

but the overall rise in E is slightly greater, because the influence of mortality through the resource dilution channel is larger in magnitude. Intuitively, an increase (decrease) in returns to innovation due to the increase in knowledge transmission (resource dilution) raises (lowers) the demand for both present and future consumption, and consequently induces labor reallocation towards (away) current consumption and away (towards) from innovation.

Table 3 reports the results. Even though the main experiment generates approximately the same increase in the relative weight on future utility, the innovation increment i^{η} rises to a slightly lower level: to 0.31 (instead of 0.33 implied by the benchmark model). The reason for this is the lower labor elasticity in the innovation sector.

Overall, the main experiment generates a slightly smaller rise in the log of output per capita (0.7865), accounting for 86% of its actual change. Shutting down the knowledge transmission channel reduces this quantity to 37%. This calibration thus implies that gains in adult longevity working through the knowledge transmission channels accounted for 49% of the takeoff. This number is only slightly smaller than 52% implied by the benchmark model.

Essentially, with $\sigma > 1$, the decline in mortality positively influences growth only through the direct effect, while the direction of its influence through the indirect channel is reversed. However, because its influence through the direct channel is overwhelmingly stronger (just as was the case for the benchmark model), the impact of the decline in adult mortality on development through the knowledge transmission mechanism remains important.

4 Conclusions

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

For reasonable assumptions on idea aggregation and their intergenerational transfer, we derived a simple and intuitive form of the adverse impact of adult mortality on knowledge transmission across generations. The destruction rate of TFP we obtained is an increasing and convex function of adult mortality of current and past cohorts. We then incorporated this form into an overlapping generations model of endogenous growth, which allowed us to identify its key parameter by taking advantage of the available historical and the Black Death epidemic data for England.

In order to assess the quantitative relevance of the proposed link between adult mortality and growth, we applied the model to investigate the long-run growth experience of England. When we vary the time-dependent parameters of the mortality shock distribution and fertility rates according to our estimates, so that the model accurately captures the population and land plot dynamics, the parameterized model generates both an early stagnation and the later rise in output per capita, explaining 90% of the increase in log output per capita between 1680 and 1880. Moreover, the influence of the decline in adult mortality, exerted through the two knowledge transmission channels (by directly improving knowledge transmission across time and by indirectly encouraging investment in productivity), accounts for one half of the empirical rise in the log of output per capita.

These findings suggest that, in addition to their more conventional effects, reductions in adult mortality substantially contribute to the process of development by extending the amount of personal contact in learning relationships and thereby improving the process of knowledge transmission across generations. The causal link between mortality and growth, proposed here, thus deserves further theoretical and empirical investigation.

Acknowledgments We thank Michele Boldrin, Oded Galor, Lutz Hendricks, Larry Jones, Fidel Pérez-Sebastián, John Seater and David Weil for their comments and the seminar participants at the Federal Reserve Bank of Minneapolis, University of North Carolina, NC State University, 2008 ASSA meetings, and Growth Workshop hosted by Duke University in March 2008

Appendix

Data sources [31]

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, 1975 (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).

[31] 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.

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

Lemma 1 (Uniqueness and interiority) *If* $0 < \theta, \eta < 1$ *and* $\sigma > 0$, *the maximization problem (18) has a unique and interior* ($0 < i < 1$) *solution.*

Proof To show uniqueness, it suffices to prove that the objective function in (18) is strictly concave. In particular, if both terms in the objective function are strictly concave, their sum is also strictly concave. The first term's second derivative, $\theta\left[\theta(1-\sigma)-1\right](1-i)^{\theta(1-\sigma)-2}$ is negative when $\theta(1-\sigma) < 1$, which is guaranteed by $0 < \theta < 1$ and $\sigma > 1$. The second term is a composite of two functions, the one in the squared brackets, and $B\left[\cdot\right]^{1-\sigma}/(1-\sigma)$. The first is strictly concave if $\eta < 1$, and the latter is strictly concave when $\sigma > 0$. Thus, the composite is strictly concave. To show interiority, we examine the slope of the objective function,

$$g'(i) \equiv -\theta(1-i)^{\theta(1-\sigma)-1} + B\left[1 + x + i^{\eta}\right]^{-\sigma} \eta i^{\eta-1}.$$

Notice that $\lim_{i \searrow 0} g'(i) = +\infty$, as the limit of the first term is a constant and the limit of the second term is ∞. Hence, the optimal $i > 0$. Moreover, $\lim_{i \nearrow 1} g'(i) = -\infty$, as the limit of the first term is $-\infty$, while the limit of the second term is a constant. Hence, the optimal $i < 1$. □

Proposition 1 See text.

Proof to Proposition 1 We want to show that the BG equilibrium described in Proposition 1, i.e. $\{A_t = \left(\gamma_A^{BG}\right)^t A_0, N_t^y = \left(n\left(1-\zeta m^y\right)\right)^t N_o^y, N_t^c = \left(n\left(1-\zeta m^y\right)\right)^t N_o^c, N_t^o = \left(n\left(1-\zeta m^y\right)\right)^t N_o^o, l_t = l^{BG}, \lambda_t^y = \lambda_o^y/\left(n\left(1-\zeta m^y\right)\right)^t, \lambda_t^o = \lambda_o^o/\left(n\left(1-\zeta m^y\right)\right)^t\}$, 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 $N_0^c = nN_0^y$, $\frac{N_0^o}{N_0^y} = \frac{(1-m^y)}{(1-\zeta m^y)n}$ and (22) and hence it satisfies the equilibrium relationships (4)–(6). The candidate solution satisfies $\lambda_0^y = \frac{0.5A}{N_0^y}$, $\lambda_0^y = \frac{0.5A}{N_0^y}$ and (24), and hence it satisfies the equilibrium law of motion for land holdings (7) and (8). Because the candidate solution satisfies (20) and because all regions 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 (14). Finally, since the candidate solution satisfies (21), it also satisfies the first order condition (19). Since there is a unique equilibrium solution, the candidate BG solution is the equilibrium solution. □

Lemma 2 *Optimal innovation time of the follower decreases in diffusion,* $i'\left(x_{t,j}\right) < 0$.

Proof We examine Eq. 19. The left hand side strictly increases in $i_{t,j}$. The right hand side strictly decreases in $i_{t,j}$. These curves intersect at a single point since the solution is unique by Lemma 1. Notice that a higher $x_{t,j}$ lowers the value of the RHS, and hence lowers the equilibrium level of innovation. □

Corollary 1 $\sup\left(i\left(A_{t,j}\right)\right) = i\left(\bar{A}_t\right)$, *i.e. the innovation time of the leader is the upper bound on the innovation time of the followers.*

Lemma 3 *Locations which are farther behind the leader in term of TFP, experience a higher rate of TFP augmentation. Formally,* $x_{t,j} > x_{t,i} \Rightarrow \left[i\left(x_{t,j}\right)^{\eta} + x_{t,j} > i\left(x_{t,i}\right)^{\eta} + x_{t,i}\right]$.

Proof We want to show that $i\,(x)^{\eta}+x$ increases in $x, \forall x>0$, i.e. $\eta i^{\eta-1}\frac{\partial i}{\partial x}>-1$. Since $\partial i/\partial x<0$ by Lemma 2, it suffices to show that $\left|\eta i^{\eta-1}\frac{\partial i}{\partial x}\right|<1$. Define a shorthand $g=1+x+i^{\eta}$. Applying the implicit function theorem to (19) gives

$$\frac{\partial i}{\partial x}=-\frac{-\sigma Bg^{-\sigma-1}\eta i^{\eta-1}}{\theta\left[\theta(1-\sigma)-1\right](1-i)^{\theta(1-\sigma)-1}-\sigma Bg^{-\sigma-1}\left(\eta i^{\eta-1}\right)^{2}+Bg^{-\sigma}\eta\left(\eta-1\right)i^{\eta-2}},$$

which implies that $\left|\eta i^{\eta-1}\frac{\partial i}{\partial x}\right|=\frac{\sigma Bg^{-\sigma-1}\left(\eta i^{\eta-1}\right)^{2}}{\theta[1-\theta(1-\sigma)](1-i)^{\theta(1-\sigma)-1}+\sigma Bg^{-\sigma-1}\left(\eta i^{\eta-1}\right)^{2}+Bg^{-\sigma}\eta(1-\eta)i^{\eta-2}}<1$. □

Corollary 2 *If two regions experience the same mortality rate, the region which is farther behind the leader will experience a higher growth rate of TFP.*

Proof Follows directly from $\frac{A_{t+1,j}}{A_{t,j}}=\left(1-m_{t,j}^{y}\right)^{\phi}\left(1-m_{t-1,j}^{y}\right)^{\phi}\left[1+i\left(x_{t,j}\right)^{\eta}+x_{t,j}\right]$ and Lemma 3. □

Proposition 2 *Suppose that after time t mortality shocks are fixed across time and across regions: $m_{t+s,j}^{y}=m^{y}\ \forall s=0,1,2,\ldots$ and $\forall j=1,\ldots,J$. For any pair of regions i, j and for $k=1,2,\ldots$, if $A_{t,i}>A_{t,j}$ then*

$$(i):A_{t+k,i}>A_{t+k,j}$$
$$(ii):\frac{A_{t+k+1,i}}{A_{t+k+1,j}}<\frac{A_{t+k,i}}{A_{t+k,j}}$$
$$(iii):\lim_{k\to\infty}\frac{A_{t+k,i}}{\bar{A}_{t+k}}=1$$

Proof (i) For this part, it is convenient to write the law of motion of TFP as follows:

$$A_{t+1,i}=\left(1-m^{y}\right)^{2\phi}\left[\tau\bar{A}_{t}+\left(1-\tau+i_{t,i}^{\eta}\right)A_{t,i}\right]$$

From Lemma 2, we know that innovation is increasing in TFP, and therefore $A_{t+1,i}$ is an increasing function of $A_{t,i}$. Therefore, $A_{t,i}>A_{t,j}\Rightarrow A_{t+1,i}>A_{t+1,j}$, and $A_{t+1,i}>A_{t+1,j}\Rightarrow A_{t+2,i}>A_{t+2,j}$,.... This implies that $A_{t+k,i}$ is an increasing function of $A_{t,i}$ (since $A_{t+k,i}$ is obtained via composition of increasing functions), and therefore

$$A_{t,i}>A_{t,j}\Rightarrow A_{t+k,i}>A_{t+k,j}$$

(ii) For this part, it is convenient to write the law of motion of TFP as follows:

$$A_{t+1,i}=\left(1-m^{y}\right)^{2\phi}\left[1+i\left(x_{t,i}\right)^{\eta}+x_{t,i}\right]A_{t,i}$$
$$\text{where } x_{t,i}=\tau\left(\bar{A}_{t}-A_{t,i}\right)/A_{t,i}$$

By Lemma 3, we have $x_{t,j}>x_{t,i}\Rightarrow\left[i\left(x_{t,j}\right)^{\eta}+x_{t,j}>i\left(x_{t,i}\right)^{\eta}+x_{t,i}\right]$. Thus, for all t, if region j is lagging behind region i, this implies that $A_{t,j}$ is growing faster than $A_{t,i}$.

(iii) Define a sequence of real numbers: $a_{k}=A_{t+k,i}/\bar{A}_{t+k}$. By definition of a leader, we have $a_{k}<1\forall k=1,2,\ldots$. Thus, the sequence $\{a_{k}\}$ is bounded above by 1, i.e. $\sup_{k}\{a_{k}\}=1$. Also, by part (ii) of this proposition, we know that $A_{t+k,i}$ is growing faster than the leader for all k, and therefore the sequence $\{a_{k}\}$ is monotone increasing. Then, by the monotone convergence theorem we have $\lim_{k\to\infty}a_{k}=1$. □

Corollary 3 *The time t leader remains a leader with TFP path given by* $\bar{A}_{t+k} = (1-m^y)^{2k\phi}\,[1+\bar{\imath}^{\eta}]^k\,\bar{A}_t$, *where* $\bar{\imath}$ *is defined as the solution to (19) in which* $x_{t,j}$ *is set to zero.*

References

Acemoglu, D. (2009). *Introduction to modern economic growth*. Princeton: Princeton University Press.

Acemoglu, D., & Johnson, S. (2007). Disease and development: The effect of life expectancy on economic growth. *Journal of Political Economy, 115*(6), 925–985.

Aghion, P., Howitt, P., & Murtin, F. (2009). The relationship between health and growth: When Lucas meets Nelson–Phelps. Working Manuscript, Harvard University.

Aiyar, S., Dalgaard, C.-J., & Moav, O. (2008). Technological progress and regress in pre-industrial times. *Journal of Economic Growth, 13*, 125–144.

Ashraf, Q., & Galor, O. (2011). Dyanmics and stagnation in the Malthusian Epoch. *American Economic Review, 101*.

Bar, M., & Leukhina, O. (2010a). Demographic transition and industrial revolution: A macroeconomic investigation. *Review of Economic Dynamics, 13*(2), 424–451.

Bar, M., & Leukhina, O. (2010b). Technical Appendix to "The role of mortality in the transmission of knowledge". http://www.bss.sfsu.edu/mbar/Research/KnowledgeSupplement.pdf.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy, 75*, 352–365.

Boucekkine, R., de la Croix, D., & Licandro, O. (2002). Vintage human capital, demographic trends, and endogenous growth. *Journal of Economic Theory, 104*, 340–375.

Boucekkine, R., de la Croix, D., & Licandro, O. (2003). Early mortality decline at the dawn of modern growth. *Scandinavian Journal of Economics, 105*, 401–418.

Boucekkine, R., de la Croix, D., & Peeters, D. (2007). Early literacy achievements, population density and the transition to modern growth. *Journal of the European Economic Association, 5*, 183–226.

Cervellati, M., & Sunde, U. (2005). Human capital formation, life expectancy and process of economic development. *American Economic Review, 95*(5), 1653–1672.

Clark, G. (2001a). Microbes and markets: Was the Black Death an economic revolution? Working Manuscript, UC Davis.

Clark, G. (2001b). The secret history of the industrial revolution. Working Manuscipt, UC Davis.

Crafts, N., & Mills, T. C. (2007). From Malthus to Solow: How did the Malthusian economy really evolve? *Journal of Macroeconomics, 31*(1), 68–93.

Doepke, M. (2004). Accounting for fertility decline during the transition to growth. *Journal of Economic Growth, 9*(3), 347–383.

Doepke, M., & Zilibotti, F. (2005). The macroeconomics of child labor regulation. *American Economic Review, 95*(5), 1492–1524.

Ehrlich, I., & Lui, F. T. (1991). Intergenerational trade, longevity and economic growth. *Journal of Political Economy, 99*(5), 1029–1059.

Erosa, A., Koreshkova, T., & Restuccia, D. (2009). How important is human capital? A quantitative theory of world income distribution. Working Manuscript, Concordia University.

Fernandez-Villaverde, J. (2001). Was Malthus right? Economic growth and population dynamics. Working Manuscript.

Galor, O. (2005). From stagnation to growth: Unified growth theory. In P. Aghion & S. N. Durlauf (Eds.), *Handbook of economic growth, Vol. 1B* (pp. 1113–1180). Amsterdam: Elsevier.

Galor, O., & Weil, D. N. (2000). Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond. *American Economic Review, 90*(4), 806–828.

Greenwood, J., & Seshadri, A. (2002). The U.S. demographic transition. *American Economic Review, 92*(2), 153–159.

Gruber, J. (2006). A tax-based estimate of the elasticity of intertemporal substitution. NBER Working paper No. 11945.

Halverson, M. S. (2007). Native American beliefs and medical treatments during the Smallpox epidemics: An evolution. *The Early America Review, VII*(4), 6–12.

Hansen, G. D., & Prescott, E. C. (2002). Malthus to Solow. *American Economic Review, 92*(4), 1205–1217.

Hazan, M. (2009). Longevity and lifetime labor input: Evidence and implications. *Econometrica, 77*, 1829–1863.

Hazan, M., & Berdugo, B. (2002). Child labor, fertility, and economic growth. *Economic Journal, 112*(482), 810–828.
Hazan, M., & Zoaby, H. (2006). Does longevity cause growth? A theoretical critique. *Journal of Economic Growth, 11*, 363–376.
Human Mortality Database. http://www.mortality.org.
Jones, C. (1995). R&D-based models of economic growth. *Journal of Political Economy, 103*(4), 759–784.
Jones, C. (2001). Was an industrial revolution inevitable? Economic growth over the very long run. *Advances in Macroeconomics, 1*(2), 1028–1048.
Kalemli-Ozcan, S. (2002). Does mortality decline promote economic growth? *Journal of Economic Growth, 7*(4), 411–439.
Kalemli-Ozcan, S., Ryder, H. E., & Weil, D. N. (2000). Mortality decline, human capital investment and economic growth. *Journal of Development Economics, 62*(1), 1–23.
Kelley, A. C., & Schmidt, R. M. (1995). Aggregate population and economic growth. *Demography, 32* (4), 543–555.
Lagerlöf, N.-P. (2003a). From Malthus to modern growth: Can epidemics explain the three regimes? *International Economic Review, 44*(2), 755–777.
Lagerlöf, N.-P. (2003b). Mortality and early growth in England, Sweden, and France. *Scandinavian Journal of Economics, 105*, 419–439.
Lorentzen, P., McMillan, J., & Wacziarg, R. (2008). Death and development. *Journal of Economic Growth, 13*(2), 81–124.
Lucas, R. E. (2009). Ideas and growth. *Economica, London School of Economics and Political Science, 76*(301), 1–19.
Macdonald, S. (1979). The diffusion of knowledge among Northumberland Farmers, 1780–1815. *The Agricultural History Review, 27*(1), 30–39.
Maddison, A. (1995). *Monitoring the world economy: 1820/1992*. Washington: OECD Development Center.
Malthus, T. R. (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.
Mansfield, E. (1961). Technical change and the rate of imitation. *Econometrica, 29*, 741–766.
Mitch, D. (1998). In J. Mokyr (Ed.), *The British industrial revolution: An economic perspective* (2nd ed.). Boulder: Westview Press.
Mitchell, B. R. (1975). *European historical statistics, 1750–1970*. New York: Columbia University Press.
Mokyr, J. (2002). *The gifts of Athena: Historical origins of the knowledge economy*. Princeton: Princeton University Press.
Monteiro, G., & Pereira, A. S. (2007). From growth spurts to sustained growth: The nature of growth and unified growth theory. Working manuscript, University of York.
Olsson, O. (2000). Knowledge as a set in idea space: An epistemological view on growth. *Journal of Economic Growth, 5*, 253–275.
Romer, P. (1990). Endogenous technological change. *Journal of Political Economy, 98*(5), S71–S102.
Rosenzweig, M. R., & Wolpin, K. I. (1985). Specific experience, household structure, and intergenerational transfers: Farm family land and labor arrangements in developing countries. *The Quarterly Journal of Economics, 100*(Supplement).
Shastry, G. K., & Weil, D. N. (2003). How much of cross-country income variation is explained by health? *Journal of the European Economic Association, 1*, 387–396.
Soares, R. R. (2005). Mortality reductions, educational attainment, and fertility choice. *American Economic Review, 95*(3), 580–601.
Tamura, R. F. (2006). Human capital and economic development. *Journal of Development Economics, 79*(1), 26–72.
Voigtlander, N., & Voth, H.-J. (2009). The three horsemen of riches: Plague, war, and urbanization in Early Modern Europe. Working Manuscipt, UCLA.
Weitzman, M. L. (1998). Recombinant growth. *The Quarterly Journal of Economics, 113*(2), 331–360.
Wrigley, E. A., Davies, R. S., Oeppen, J. E., & Schofield, R. S. (1997). *English population history from family reconstitution 1580–1837*. Cambridge: Cambridge University Press.