

This article was downloaded by: [University of Verona]

On: 5 February 2009

Access details: Access Details: [subscription number 788844018]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Mathematical Population Studies

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713644738>

A Theory of Medical Effectiveness, Differential Mortality, Income Inequality and Growth for Pre-Industrial England

DAVID DE LA CROIX ^a; ALESSANDRO SOMMACAL ^b

^a Department of Economics and CORE, Université Catholique de Louvain, Belgium ^b Department of Economics, Université Catholique de Louvain, Belgium, and L. Bocconi University, Milan, Italy

Online Publication Date: 01 January 2009

To cite this Article CROIX, DAVID DE LA and SOMMACAL, ALESSANDRO(2009)'A Theory of Medical Effectiveness, Differential Mortality, Income Inequality and Growth for Pre-Industrial England',Mathematical Population Studies,16:1,2 — 35

To link to this Article: DOI: 10.1080/08898480802619538

URL: <http://dx.doi.org/10.1080/08898480802619538>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

A Theory of Medical Effectiveness, Differential Mortality, Income Inequality and Growth for Pre-Industrial England

David de la Croix

Department of Economics and CORE, Université Catholique de Louvain, Belgium

Alessandro Sommacal

Department of Economics, Université Catholique de Louvain, Belgium, and L. Bocconi University, Milan, Italy

The interactions between mortality reductions and income growth are studied, with a special attention at their relationship prior to the Industrial Revolution, when income per head was stagnant. The choice of individual medical spending is modelled, giving a rationale for individual health expenditures even when medicine is not effective in postponing death. The rise of effective medicine is then explained by a learning process function of expenditure on health. The rise in effective medicine is linked to the economic growth of the eighteenth century through life expectancy increases which foster capital accumulation. The rise of effective medicine has also had an effect on the relationship between growth and inequality and on the intergenerational persistence of differences in income. These channels are operative through differential mortality induced by medical effectiveness that turns out to determine a differential in the propensity to save among income groups.

Keywords: differential mortality; health expenditure; life expectancy; propensity to save

1. INTRODUCTION

In the last few centuries, after years of Malthusian stagnation, income per head has shown a remarkable increase. Better material conditions and important advances in medical science have changed the quality

Alessandro Sommacal is now affiliated with the University of Verona and Econpubblica, L. Bocconi University.

Address correspondence to David de la Croix, Université Catholique de Louvain, Department of Economics, Place Montesquieu 3, B-1348 Louvain-la-Neuve, Belgium. E-mail: delacroix@core.ucl.ac.be

and the length of life for millions of people. A question naturally emerges, concerning the relationship between medical knowledge and income. The causality between these two variables seems to be bidirectional: higher income allows higher health expenditure, which contributes to building medical knowledge; advances in medical knowledge increase longevity, which by modifying the incentives to invest in physical and human capital, has an effect on the growth rate and on the distribution of income. The path from growth to longevity is reviewed for example in Fogel (1994). The effect of longevity on growth is quantitatively assessed by Boucekkine et al. (2003) and Nicolini (2004) who show that small improvements in adult life expectancy in the eighteenth century caused big changes in economic decisions, leading to an acceleration in income growth.

In order to disentangle the causal links between medicine and income, it is enlightening to look at the period prior to the Industrial Revolution, when income was stagnant but medical knowledge was improving.

From the available data for England, discussed in Section 3, we decompose the evolution of medical knowledge and life expectancy into three periods. Before the seventeenth century the effectiveness of medicine was very low. In fact, medicine was probably not effective at all. Given the low effectiveness of medical services, physicians hardly managed to increase the life expectancy of their patients. As a consequence, differential mortality between the rich (who could afford the services of a physician) and the common people was very low, or even non-existent: income was of little importance in determining life expectancy. Next came a time when medicine became more effective. The rich were the first to benefit from these improvements, and their life expectancy rose, as witnessed in the study of aristocratic British families carried out by Hollingsworth (1977) on the basis of genealogical data. Differential mortality started to increase, because the improvements in longevity still did not benefit the whole population. In the third period there was a global improvement in health. Medicine became more and more effective, and the advantage of the upper classes in terms of longevity declined. Differential mortality decreased.

We first build a model of medical spending that is suitable for a period, such as pre-industrial Europe, when medical effectiveness was very low. We give a rationale for individual health expenditures even when medicine was not effective in postponing death. Then we combine our result with a simple learning process based on health expenditures to model the rise of effective medicine: the continuous demand for health services allows knowledge to accumulate, paving the way for advances in medical science. We show that these ingredients are sufficient to build a consistent explanation of the development of differential mortality in England.

Then we use our model economy to look at the implications of the evolution of medical science on income. In particular we discuss the effect on the growth rate and we also study how differential mortality, inducing a differential in the propensity to save and bequeath across income groups, plays a role in shaping the path of inequality and the relationship between inequality and growth.

In Section 2 we present a brief review of the literature linking health and growth. In Section 3 we show mortality data per social class in order to assess the role of income in shaping health. We also provide a brief discussion and some references concerning the evolution of medical science. In Section 4, we present the model. Section 5 is focused on the solution to the household optimization problem. Section 6 is devoted to the study of dynamic properties of the model. In Section 7, we comment upon the structure of the model.

2. LITERATURE REVIEW

The relationship between longevity and economic variables is receiving growing attention. Specific applications in development economics and in the literature, following Galor and Weil (1999, 2000), explain the growth of European economies during the eighteenth and the nineteenth centuries after years of Malthusian stagnation. A general and informal discussion on the possible links between health (and in particular life expectancy) and economic variables is, for example, presented in Sala-I-Martin (2002).

La Croix and Licandro (1999) explored the possibility that exogenous improvements in longevity have positive effects on education decisions. Analogous ideas were used by Boucekkine et al. (2003) and Nicolini (2004) to shed light on the historically observed relationship between rising life expectancy and the acceleration of growth in the period before the Industrial Revolution. The effects of an exogenous reduction of mortality on growth have also been shown to be shaped by social security systems, in a framework where fertility decisions are endogenous and parents face a trade-off between the quality and the quantity of their children (Zhang et al., 2001).

Not only does life expectancy have an effect on economic variables and in particular on growth; the opposite is also true. For these reasons other authors have endogenized life expectancy to study the bidirectional link between longevity and growth. Morand (2004) argues that investment in health plays a crucial role in the transition from a growth regime fuelled by physical capital to a sustained growth regime fuelled by human capital. In Chakraborty (2004) low income and high mortality reinforce each other, making possible the existence

of poverty traps. A similar result is found in Cervellati and Sunde (2005) who model an endogenous transition from a long period of stagnant growth to a period of sustained growth through a process that can be interpreted as an industrial revolution. Kalemli-Ozcan (2002), Blackburn and Cipriani (2002), and Lagerloef (2003) combine endogenous life expectancy and endogenous fertility decisions to give a complete account of the inter-relationship between demographic and economic variables. Castello-Climent and Domenech (2006) and Chakraborty and Das (2005) show that the endogenous determination of life expectancy could significantly affect the intergenerational transmission of inequality, contributing to explaining the persistent disparities in individual wealth and income levels.

Usually growth models that endogenize life expectancy relate it either directly or indirectly to medical knowledge. A direct link can be established by the explicit introduction of investment in health. An indirect link is often modelled by linking life expectancy to human capital. This is also true for those models (Morand, 2004; Cervellati and Sunde, 2005; Lagerloef, 2003) that are more directly involved in the explanation of the growth of European economies during the eighteenth and nineteenth centuries: life expectancy increase as a consequence of higher investment in health (in Morand, 2004) or a higher level of human capital, that could be interpreted as greater medical knowledge (in Cervellati and Sunde, 2005; Lagerloef, 2003). Galor and Moav (2005) develop a complementary theory (that does not rely on the role of medical technology) of the time path of life expectancy over time in the ten thousand years between the agricultural revolution in the Neolithic period and the Industrial Revolution at the end of the eighteenth century. The theory relies on the idea that the mortality risk associated with environmental factors rose in the transition from hunter-gatherer tribes to sedentary agricultural communities and ultimately to urban societies, triggering a process of natural selection that produced a population with a greater genetic potential for longer life expectancy.

However, to the best of our knowledge, though many aspects of the relationship between longevity, growth, and inequality have been explored in the literature, no research has focused on the specific role of the effectiveness of medical science, explaining its evolution and the related implications for growth and inequality.

3. INCOME AND MORTALITY IN PRE-MODERN EUROPE

It is difficult to find data on mortality rates or life expectancy by income groups for the pre-industrial period. We gather here some evidence for England as a whole, combining sources from Hollingsworth (1977) and

Wrigley et al. (1997), and for two cities in continental Europe, based on surveys by Perrenoud (1975) and Bardet (1983).

Starting with England, Figure 1 compares life expectancy at birth of the average person (from parish records in Wrigley et al., 1997) with that of the English aristocracy (from genealogical data in Hollingsworth, 1977). Before 1700, there was not much difference in adult mortality across social classes. Surprisingly the elites had lower life expectancy; this unexpected result is attributed by Johansson (1999) to an urban penalty paid by the aristocracy for the pleasures and opportunities of city life. The quasi-egalitarian mortality regime began to change by 1700. Life expectancy rose for all groups but faster and further for the elite.

Two other data sets can give further hints about mortality by social class before the Industrial Revolution. They cover the population of two cities in continental Europe, Geneva (Perrenoud, 1975) and Rouen (Bardet, 1983). Age-specific survival probabilities are presented in Table 1 for three social classes. In both data sets, infant mortality rates were much lower in the elite groups. This reflects their better living conditions. We also find that in Geneva (seventeenth century) and in Rouen (eighteenth century) there was not much difference in adult mortality across social classes. In Rouen, there was no significant difference in the survival probabilities from ages 15 to 30 or from

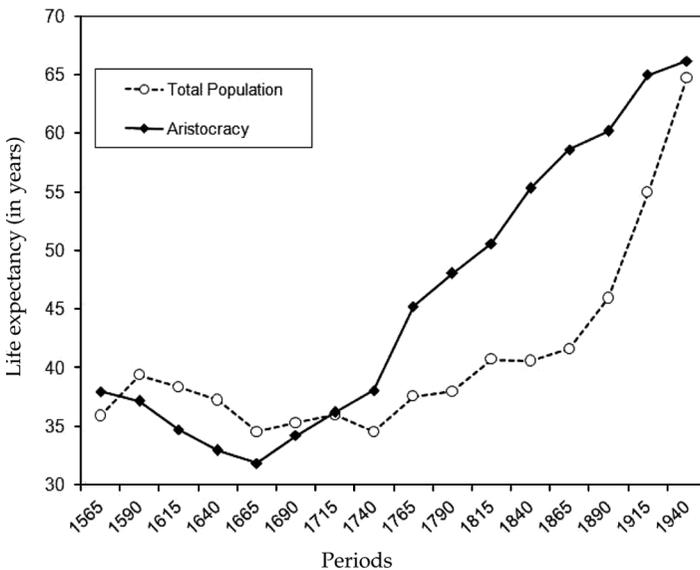


FIGURE 1 Life expectancy at birth in England.

TABLE 1 Survival Rates in Geneva and Rouen

		Survival probabilities		
		0 → 15	15 → 30	30 → 45
Geneva XVII	workers	0.34	0.80	0.70
	merchants	0.45	0.84	0.74
	nobility	0.61	0.89	0.81
Rouen XVIII	workers	0.33	0.85	0.87
	merchants	0.49	0.87	0.86
	nobility	0.47	0.86	0.84

ages 30 to 45 of notables and simple workers. This is in line with the English data. In Geneva, the upper social class had a slight advantage.

The determinants of the evolution of life expectancy during the past centuries have been the subject of lively discussions in medical history and mortality history. One of the issue concerns the role of medical science.

Ancient ideas persisted a long time in modern Europe and the confidence of consumers in medicine was low. Popular proverbs endorsed this distrust: “one physician makes work for another” (for the history of medicine: Siraisi (1990), Lindemann (1999), and Porter (1995)). As a consequence some authors claim that the rise of life expectancy in early modern Europe relied more on changes in immunology or improvement in the climate than on human factors such as medical advances.

Johansson (1999) argues against the traditional therapeutic nihilism that tends to deny that medicine had any effectiveness before the end of the nineteenth century, and suggested an increase in medical effectiveness as a possible explanation for the change, documented in Figure 1, from an egalitarian mortality regime to a regime characterized by a mortality differential between rich and poor people.

In the period 1500–1800, medicine showed an increasingly experimental attitude: no improvement was effected on the grounds of the disease theory (which was still mainly based on traditional ideas), but significant advances were made based on practice and empirical observations. For example, although the theoretical understanding of how drugs work only came progressively in the nineteenth century with the development of chemistry (Weatherall, 1996), the effectiveness of the treatment of some important diseases was improved thanks to the practical use of new drugs coming from the New World.¹ Advances in the

¹Ipecacuanha was used for severe dysentery, guaiacum for syphilis, and the bark of the cinchona tree for malaria. Citrus fruits started to be used in the prevention and treatment of scurvy.

TABLE 2 Total Number of Books on Health Published in England, 1600–1800

Period	Number of books
1600–24	9
1625–49	16
1650–74	17
1675–99	25
1700–24	28
1725–49	34
1750–74	53
1775–1800	81

treatment of syphilis were made due to improvements in the development of condoms. For a long period, mainly due to their high cost, these new medical advances were only available to rich people.²

Moreover Table 2, which shows the total number of books containing lifestyle advice written in the period 1600–1800 by sub-periods, provides some indirect evidence of the fact that lifestyle advice (concerning, e.g., personal and domestic cleanliness) became popular among upper class readers.³

As suggested by Johansson (1999), the cumulative effects of these improvements could have produced a net increase in the efficacy of medicine in the eighteenth century. “As early as 1829 Dr. F.B. Hawkins wrote a book entitled *Elements of Medical Statistics*, in which he described what could be called an early modern epidemiological transition. Several centuries before his own time leprosy, plague, sweating sickness, ague, typhus, smallpox, syphilis, and scurvy had been leading causes of death. Now all of these diseases had disappeared, could be cured, or treated effectively. At the present scarlet fever, consumption, gout, dropsy, palsy, apoplexy (including heart

²In the eighteenth century bubonic plague also disappeared from England. The reasons for this are still hotly debated and some authors explain it in exogenous terms. In any case, independently of this debate, it should be noticed that by the seventeenth century a “warning system” was developed in London. Weekly bills of mortality, containing the cause of death, were published: when the total number of plague deaths reached a worrying value, foreshadowing the outbreak of an epidemic, people with enough money left the city. This system, though not effective in the treatment of the plague itself, can be interpreted as a public health measure, which had an effect on the health of rich people.

³We are aware that, to draw conclusions about the increasing diffusion of lifestyle advice, it would be better to look at the ratios of medical books to the total number of books. However, to the best of our knowledge, data about the total number of books published in the period is not available.

attacks and strokes), mania, and diseases of the brain were the most prevalent causes of death. The last six of the eight diseases listed were not contagious; they were chronic diseases most likely to strike older adults” (Johansson, 1999, p. 48).

We shall focus on the role of medicine (including lifestyle advice). This does not mean that the increase in life expectancy is explained only by advances in medical science. Many factors may have contributed to the decline of mortality since the eighteenth century (nutrition, medicine, immunology) and assigning a specific weight to each of them is very difficult (Fogel, 2004). The reason that we restrict our analysis to medical effectiveness is that we focus more on the differences in the increase of life expectancy between the rich and the poor than on the variation of life expectancy in absolute terms. Access to health care seems to be a particularly determinant of life expectancy relevant for explaining this differential mortality between social classes.

4. A MODEL OF MEDICAL SPENDING

The model is set up in discrete time, from 0 to infinity. In each period one physical good is produced using labor, capital, and land. All households are endowed with one unit of labor. The total supply of land is fixed and it is normalized to 1. We assume as the initial condition that the ownership of land is equally distributed among some households (called “landlords” or “aristocrats” which we will denote with the exponent “L”) while other households (called “commoners” which we will denote with the exponent “W”) do not own any land. Moreover we also assume for simplicity that there is no market for land and, as a consequence, this ownership structure is unchanged over time. In this sense we have in the economy two social classes or groups. The size N^i of the newborn generation in each population group is assumed constant over time (any other stationary dynamic process for the population would qualitatively give the same result, without adding any crucial element).

Households can live for two periods, adulthood and old age. They are alive with certainty during the first period, and at the end of adulthood they die with a probability $\beta(h_t^i)$ which depends on their health status h_t^i .

4.1. Individual Budget Constraints

In the first period of life, agents receive a bequest b_t^i from their parents and obtain income from land and/or labor:

$$y_t^L = b_t^L + w_t + \rho_t x_t \quad (1)$$

$$y_t^W = b_t^W + w_t \quad (2)$$

where y_t^i denotes income, ρ_t is the return from holding one unit of land, $x_t = 1/N^L$ is the amount of land per person of group L , and w_t is the wage per person that is the same for both groups.

Income finances consumption c_t^i , investment in health d_t^i and savings s_t^i . If an agent survives to the second period he or she does not consume, and any savings are used to finance the bequest left to his or her offspring. The budget constraints of the first and second period are, respectively:

$$y_t^i = c_t^i + d_t^i + s_t^i \quad (3)$$

$$R_{t+1}s_t^i = b_{t+1}^i. \quad (4)$$

The inter-temporal budget constraint is

$$y_t^i = c_t^i + d_t^i + \frac{b_{t+1}^i}{R_{t+1}}. \quad (5)$$

Here we do not assume the existence of a perfect annuities market. If agents die before entering old age their resources ($R_{t+1}s_t^i$) pass to offspring. As a consequence the bequests received by agents at the beginning of adulthood do not depend on the survival of parents.

4.2. Medicine as an Experience Good

We distinguish the stock of health h_t^i from the investment in health d_t^i . The reason is that in the model health status is related to health expenditure in a non-deterministic way: indeed we explicitly take into account that medical treatment may turn bad, in particular if medicine is inefficient. When allocating an amount d_t^i of resources to health expenditure, the household can reach a high level of health $h^+(d_t^i)$ with probability p_t , and a worsened level of health $h^-(d_t^i)$ with probability $1 - p_t$; in particular:

$$h(d_t^i) = \begin{cases} h^+(d_t^i) = 1 + v(d_t^i) & \text{with probability } p_t \\ h^-(d_t^i) = 1 - v(d_t^i) & \text{with probability } 1 - p_t \end{cases} \quad (6)$$

where $v(0) = 0$; $v'(d_t^i) > 0$. The idea underling this formulation is that medical expenditure can be good or bad for health and that agents cannot know the quality of medical care before experiencing it, but they do know the probability p_t that can be interpreted as a measure of the effectiveness of medicine.

The possibility that medicine might harm patients could seem at first glance strange given today's standards, but it has always been recognized by physicians. One of the most basic precepts all medical students have been taught is *primum non nocere* ("first of all, do not harm"). This aphorism is meant to remind them that every medical intervention has the possibility of damaging patients.⁴ When building economic models for the demand for *modern* health services, it is considered reasonable for analytical simplicity to omit this feature of medical treatments, as the probability of medicine for damaging the health of patients is very low.⁵ However, when analyzing the demand for medical services in *pre-modern* Europe, we think that it is crucial to explicitly take into account the potential damage that medical services might cause: a theory that abstracts from this feature of health expenditure should be considered of doubtful applicability in a period where the development of medicine and its effectiveness were very far from the present levels.

The other assumption implicit in Eq. (6) is that medicine is an experience good, that is people do not know the quality of medical services before having bought them. Once again this is true nowadays and it was even more true in the past.⁶

We stress that by using these assumptions we put ourselves in the worst position for justifying our theory of the rise of effective medicine, which relies on the learning by spending mechanism described below. Indeed, if we use a standard formulation in which medicine is always a good it is easy to explain why the demand for medical services was positive in early modern Europe. This explanation is much less obvious under our formulation which allows the probability of medicine of damaging patients to be equal to 1/2.

4.3. Preferences

The preferences of an individual born at time t and belonging to class i are defined over bundles $(c_t^i, b_{t+1}^i, h_t^i)$. Preferences are represented by a

⁴This aphorism is usually considered a Latin paraphrase by Galen of an Hippocratic aphorism, but it seems more plausible to attribute it to the English physician Thomas Sydenham (1624–1689). (Smith (2005) provides a history of this aphorism and a discussion of its applicability as an ethical guide for modern medicine.)

⁵However, saying that this probability is perhaps negligible from an economic point of view, does not mean that it can be disregarded from the concrete point of view of medical science and its patients. In medical literature the problem of adverse reactions to drugs for example is widely discussed.

⁶Mokyr (1993) describes the uncertainty a consumer faced in the selection of a good medical service in the past and the way medical technology diffused among households.

utility function $U(c_t^i, b_{t+1}^i, h_t^i)$ which is increasing in the three arguments.

Health affects the utility through two channels: it influences the probability of surviving in old age and it has a direct effect on the utility function when old. Assuming time separability, $U(c_t^i, b, h_t^i)$ can be written as:

$$u(c_t^i) + \beta(h_t^i)\tilde{u}(b_{t+1}^i, h_t^i) \quad (7)$$

where u and \tilde{u} are well-behaved increasing instantaneous utility functions. The survival probability $\beta(h_t^i)$ takes the form:

$$\beta(h^+(d_t^i)) = \beta + f(d_t^i) \quad (8)$$

$$\beta(h^-(d_t^i)) = \beta - f(d_t^i) \quad (9)$$

where $f'(\cdot) > 0$, $f(0)=0$ and $f(\infty) \leq \beta \leq 1 - f(\infty)$. From Eq. (6) we have the survival probability determined by Eq. (8) and (9) respectively with probability p_t and $1 - p_t$.

If $\tilde{u}(b_{t+1}^i, h_t^i) = u(b_{t+1}^i)h_t^i$, utility can be written in expected terms as:

$$\begin{aligned} \mathbb{E} U(c_t^i, b_{t+1}^i, h_t^i) &= u(c_t^i) + p_t(\beta + f(d_t^i))u(b_{t+1}^i)(1 + v(d_t^i)) \\ &\quad + (1 - p_t)(\beta - f(d_t^i))u(b_{t+1}^i)(1 - v(d_t^i)) \end{aligned} \quad (10)$$

$$\begin{aligned} &= u(c_t^i) + u(b_{t+1}^i)(\beta + f(d_t^i)v(d_t^i) \\ &\quad + (2p_t - 1)(f(d_t^i) + \beta v(d_t^i))) \\ &= u(c_t^i) + \phi(d_t^i, p_t)u(b_{t+1}^i) \end{aligned} \quad (11)$$

$$\text{with } \phi(d_t^i, p_t) \equiv \beta + f(d_t^i)v(d_t^i) + (2p_t - 1)(f(d_t^i) + \beta v(d_t^i)). \quad (12)$$

This utility function is formally similar to the one used by Chakraborty and Das (2005). The difference is that in our case the function $\phi(d, p)$ is not a survival probability: it is a more complex expression that depends on the survival probability (and thus on the probability p_t) and on the utility of health, and is derived from our previous assumptions.

4.4. Medical Technology

The probability p_t of getting benefits from medicine depends on the stock of empirical evidence (or medical experience) M_t currently available:

$$p_t = G(M_t). \quad (13)$$

However, we assume that there is a threshold \bar{M} below which more empirical evidence does not lead to improvements in the effectiveness of medicine:

$$G(M_t) = \begin{cases} \frac{1}{2} & \text{if } M_t < \bar{M} \\ \frac{1}{2} + g(M_t) & \text{if } M_t \geq \bar{M} \end{cases} \quad (14)$$

where $g(M_t) > 0$ and $g'(M_t) > 0$. We assume that M_t accumulates with expenditure in health:

$$M_t = (1 - \delta)M_{t-1} + D_t \quad (15)$$

where aggregate expenditure in health is given by

$$D_t = N^W d_t^W + N^L d_t^L \quad (16)$$

and δ is a depreciation rate, possibly zero. The view behind this formulation is that medical effectiveness varies through a learning process based on experiments with treatments. It turns out to be related to the amount of medical expenditure.

While standard models of learning by doing (explaining improvements in labor productivity) are continuous, Eq. (14) assumes a threshold effect: empirical observations have a direct and sizable effect on medical effectiveness only when they have reached a sufficiently high level.

This assumption is justified by taking into account the fact that “physicians slowly learned how to learn from empirical observations.” The idea is that it took many observations (and thus the accumulation over time of a sufficiently high level of health expenditure) before the importance of observations themselves was fully understood and more effective methods for drawing conclusions from experience were developed. In other terms, the *discontinuity* in Eq. (14) depends on the existence of a discontinuity in the way in which knowledge has been related to empirical observation. This discontinuity is part of the “scientific *revolution*” of the sixteenth and seventeenth centuries. As mentioned in Section 3, it was only after the seventeenth century that medicine really started to become an empirical science, which relies on systematic and organized methods of observation.⁷ Before that period the link between empirical observations and medicine was much less systematic and tight (for simplicity, in Eq. (14) we

⁷The philosopher Francis Bacon (1561–1626) (known for his theory of induction and “The history of life and death with observations natural and experimental for the prolonging of life”) and the physician Thomas Sydenhan (1642–1689) (known as the “father of clinical medicine”).

assumed that, at an aggregate level, the effect of empirical evidence on the effectiveness of medicine was negligible).

We did not consider any exogenous technical progress, either in the medical sector or in the final good sector. This does not mean that we believe that there cannot be improvements in medical knowledge without experimentation (after all, fundamental research may play a role) but it does imply that our results do not rely on the presence of such exogenous factors.

4.5. Final Good Technology

The final good is produced from a technology combining land, labor, and capital. We assume that capital fully depreciates after one period. The production function is of the AK type, and includes a positive externality of the aggregate stock of physical capital \bar{K}_t in such a way that social marginal returns to capital are constant:

$$Y_t = AL_t^{1-\alpha-\theta} K_t^\alpha X_t^\theta \bar{K}_t^{1-\alpha} \quad (17)$$

with A denoting total factor productivity, L_t labor input, K_t capital input, and X_t land. The profits of the representative firm are given by

$$Y_t - w_t L_t - R_t K_t - \rho_t X_t. \quad (18)$$

4.6. The Equilibrium

Given an initial stock of medical knowledge M_0 , an initial stock of capital K_0 , and an initial asset distribution (s_{-1}^W, s_{-1}^L) such that $K_0 = N^W s_{-1}^W + N^L s_{-1}^L$, an equilibrium is

- a vector of individual variables maximizing utility subject to the budget constraint:

$$\begin{aligned} (c_t^i, d_t^i, b_{t+1}^i) &= \arg \max u(c_t^i) + \phi(d_t^i, p_t) u(b_{t+1}^i) \\ &\text{subject to } y_t^i = c_t^i + d_t^i + \frac{b_{t+1}^i}{R_{t+1}} \end{aligned} \quad (19)$$

with income given by (1)–(2) and with savings s_t^i related to bequests b_{t+1}^i through Eq. (4);

- a vector of factor input maximizing profits

$$(L_t, K_t, X_t) = \arg \max AL_t^{1-\alpha-\theta} K_t^\alpha X_t^\theta \bar{K}_t^{1-\alpha} - w_t L_t - R_t K_t - \rho_t X_t \quad (20)$$

- a vector of medical effectiveness (p_t) and medical knowledge (M_t) satisfying Eq. (13), (14) and

$$M_t = (1 - \delta)M_{t-1} + N^L d_t^L + N^W d_t^W \quad (21)$$

- a vector of prices (w_t, R_t, ρ_t) such that all markets clear

$$L_t = N^W + N^L \quad (22)$$

$$K_{t+1} = N^W s_t^W + N^L s_t^L \quad (23)$$

$$X_t = 1, \quad (24)$$

- an aggregate stock of capital $\bar{K}_t = K_t$.

5. THE CONSUMER PROBLEM

We now study the solution to the consumer optimization problem. A specific objective is to prove that, although medicine is not effective in the sense that $p = 1/2$, people with sufficiently high income levels have an incentive to invest in health. To alleviate notation, we abstract for time and class indices in this section.

In a first step (Section 5.1) we present analytical results under some simplifying assumptions. In a second step (Section 5.2) we conduct numerical experiments to study the sensitivity of these results in the general case, which is then used in Section 6.

Analytical results for a simplified optimization problem and the analysis of the analogies and the discrepancies with respect to the general case should strengthen the understanding of the intuition behind the main mechanisms at work in our model economy.

5.1. A Simplified Model

We consider a simple version of the model with no consumption in the first period; this implies that the only arbitrage is between leaving a bequest and spending on health. We also assume the specific following functional forms:

$$u(b) = \frac{\sqrt{b}}{2}, \quad v(d) = \frac{d}{1+d} \quad f(d) = \tau \frac{d}{1+d} \quad (25)$$

where we take $\tau \leq \beta \leq 1 - \tau$ to give a well defined life expectancy. This condition is necessary and sufficient to ensure, according to Eq. (8) and (9), $0 \leq \beta(h(d)) \leq 1$ when d tends to infinity.

The individual optimization becomes:

$$\max_{d,b} \frac{\sqrt{b}}{2} \left(\beta + \tau \frac{d^2}{(1+d)^2} + (2p-1)(\beta + \tau) \frac{d}{1+d} \right) \quad (26)$$

s.t.

$$y = \frac{b}{R} + d \quad (27)$$

$$d, b \geq 0. \quad (28)$$

Substituting the budget constraint Eq. (27) into the objective function, the problem is stated as:

$$\max_{d \in [0,y]} W(d,y,p) \quad (29)$$

where:

$$W(d,y,p) = \frac{\sqrt{R(y-d)}}{2} \left(\beta + \tau \frac{d^2}{(1+d)^2} + (2p-1)(\beta + \tau) \frac{d}{1+d} \right). \quad (30)$$

We define the set of optimal levels of d as:

$$D^* = D^*(y,p) = \left\{ d : d = \arg \max_{d \in [0,y]} W(d,y,p) \right\}. \quad (31)$$

D^* is not empty because $W(d,y,p)$ is a continuous function and $[0,y]$ is a closed and bounded set. An element of D^* is denoted by $d^* = d^*(y,p)$. In general D^* could contain interior solutions ($d^* \in (0,y)$), corner solutions ($d^* = 0$ or $d^* = y$), or both.

To study the properties of D^* we compute:

$$\begin{aligned} W_d(d,y,p) = & \frac{\sqrt{R}}{2} \left(\sqrt{(y-d)} \left(\tau \frac{2d}{(1+d)^3} + (2p-1)(\beta + \tau) \frac{1}{(1+d)^2} \right) \right. \\ & \left. - \frac{1}{2\sqrt{(y-d)}} \left(\beta + \tau \frac{d^2}{(1+d)^2} + (2p-1)(\beta + \tau) \frac{d}{1+d} \right) \right). \end{aligned} \quad (32)$$

We define:

$$\Gamma(d,y,p) \equiv W_d(d,y,p) \quad (33)$$

and list its main properties:

P1 $\Gamma(d, y, p)$ is continuous in its three arguments.

P2 When $d = 0$ we have:

$$\Gamma(0, y, p) = \frac{\sqrt{R}}{2} \left(\sqrt{y}(2p - 1)(\beta + \tau) - \frac{\beta}{2\sqrt{y}} \right). \quad (34)$$

In particular if $p > 1/2$ and y is sufficiently high then $\Gamma(0, y, p) > 0$; if $p = 1/2$ then $\Gamma(0, y, p) < 0$.

P3 When $d = y$ we have

$$\Gamma(y, y, p) = \Gamma(d, d, p) = -\infty. \quad (35)$$

P4 The effect of income is given by:

$$\Gamma_y(d, y, p) > 0. \quad (36)$$

P5 The value $\Gamma(d, y, p)$ as y goes to ∞ depends on the value of p and on the fact that $d = 0$ or $d \in (0, y)$:

$$\begin{aligned} \Gamma(d, \infty, p) &= \infty \quad \forall d \in [0, y) \quad \text{if } p > 1/2 \\ \Gamma(d, \infty, p) &= \infty \quad \forall d \in (0, y) \quad \text{if } p = 1/2 \\ \Gamma(0, \infty, p) &= 0 \quad \text{if } p = 1/2. \end{aligned} \quad (37)$$

P6

$$\Gamma_p(d, y, p) > 0 \quad \text{for } d \leq d^*(y, p). \quad (38)$$

See Appendix A.1 for the proof.

P3 implies $d^* \neq y$: leaving nothing for a bequest is never optimal; hence $d^* \in [0, y)$. In Proposition 1 we prove that the solutions can be corner ($d^* = 0$) or interior ($d^* \in (0, y)$) depending on the value of y . For this purpose we define the minimum income such that we can have an interior maximum:

$$\hat{y} = \hat{y}(p) = \min\{y > 0 : \exists \text{ at least one } d \in (0, y) \text{ such that } W(d, y, p) - W(0, y, p) \geq 0\}.$$

Proposition 1.

- (i) $\hat{y}(p)$ is finite $\forall p$;
- (ii) $\forall y > \hat{y}(p)$ we have $d^* \in (0, y)$, that is the optimal choices of d are all interior;
- (iii) $\forall y < \hat{y}(p)$ we have $d^* = 0$, that is $d = 0$ is the unique optimal choice.

Proof.

- (i) A necessary and sufficient condition for a $d^* = 0$ is

$$W(d, y, p) - W(0, y, p) = \int_0^d \Gamma(s, y, p) ds \leq 0 \quad (40)$$

$\forall d \in (0, y)$. According to P5, when income tends to infinity, $\Gamma(s, y, p)$ is ≥ 0 for $s = 0$ and it tends to infinity for $s > 0$. By continuity P1, this also holds true when income is finite but large. Hence, $W(d, y, p) - W(0, y, p)$ is positive for a sufficiently large income.

- (ii) From definition (39) we know that $W(d, \hat{y}, p) - W(0, \hat{y}, p) = \int_0^d \Gamma(s, \hat{y}, p) ds \geq 0$ for at least one $d \in (0, y)$. From P4 we know that $\Gamma_y(s, \hat{y}, p) > 0$. Thus if $y > \hat{y}$ then $W(d, y, p) - W(0, \hat{y}, p) > 0$, implying that $d = 0$ is not an optimal choice.
- (iii) This follows directly from Eq. (39).

Proposition 1 implies that people with a sufficiently high level of income invest in health even when medicine is not effective, that is $p = 1/2$. This is our first result and it is the basic ingredient of our explanation of the rise of effective medicine.

To understand the idea behind Proposition 1, we start with the simple remark that d^* is chosen by comparing the marginal cost of health investment (represented by the value of bequests foregone), with its marginal gain. This comparison is affected by the income level: when income increases the marginal cost of d shrinks and the marginal benefit (if it is positive) rises. Thus as long as the marginal gain from investing in health is positive there will exist a y such that d^* is greater than zero.

What is less obvious is the fact that the gain from investing in health is positive even when medicine is equally likely to improve health status or to harm it. The intuition relies on the fact that, according to Eq. (7), health status has an effect both on the length and on the quality of life. The interaction between these two effects is the crucial element: if medicine is a good, the positive effect on the quality of life is enjoyed for a longer period, while if medicine is a bad, the negative effect on the quality of life is enjoyed for a shorter period.

It is useful to explicitly write the expected gain from health investment. From Eq. (10), we can see that in general the effect of choosing a positive amount of d can be written as:

$$pu(b)(\beta v(d) + f(d) + f(d)v(d)) \quad (41)$$

$$+(1-p)u(b)(-\beta v(d) - f(d) + f(d)v(d)) \quad (42)$$

where Eq. (41) is p times the utility derived from investing in health when medicine is a good and Eq. (42) is $1-p$ times the disutility derived from investing in health when medicine is a bad. In Eq. (41) and (42) the utility (or disutility) of health investment was decomposed: the effects of medicine on the quality of life and on the length of life as well as their interaction appear clearly. In particular, both in Eq. (41) and (42) the first term in brackets represents the quality of life effect: given the initial length of life β , the quality of life increases (or decreases) due to a change in health status equal to $v(d)$. The second term is the life expectancy effect: given the initial quality of life equal to 1, the length of life increases (or decreases) by an amount equal to $f(d)$. The third one is the cross effect, that is the positive (or negative) quality of life effect $v(d)$ applied to the increase (or the reduction) $f(d)$ in life expectancy. The first two terms may be positive or negative, depending on whether medicine is a good or a bad. However the third term, which represents the interaction between the effect on the quality of life and the effect on the length of life, is always positive: indeed if medicine is a good, this term represents an increase in the total number of years during which a better quality of life is enjoyed; if medicine is bad, it represents a reduction in the total number of years during which a lower quality of life is experienced. The presence of this cross effect implies that the quantity in Eq. (37) is greater than the quantity in Eq. (38) even when $p = 1/2$.

Given d^* , the optimal level of bequest b^* is determined using the budget constraint Eq. (27). As a consequence we have $b^* \in (0, Ry)$ and in particular when $\forall y \geq \hat{y}(p)$ we have $b^* \in (0, Ry)$, whereas when $\forall y < \hat{y}(p)$ we have $b^* = Ry$.

We now study the effects on d^* of changes in y and prove that health expenditure is a normal good.

Proposition 2. *For $y \geq \hat{y}(p)$, d^* is strictly increasing in y .*

Proof. The proof is a simple application of supermodularity to our one-dimensional choice problem (Edlin and Shannon, 1998). Let us consider $d_1^* = d^*(y_1, p)$. If y_1 increases to y_2 then $d_2^* = d^*(y_2, p)$ cannot be below d_1^* . Indeed, by definition of d_1^* and d_2^* , we always have $W(d_1^*, y_1, p) \geq W(d_2^*, y_1, p)$ and $W(d_1^*, y_2, p) \leq W(d_2^*, y_2, p)$ and as a consequence $W(d_1^*, y_2, p) - W(d_1^*, y_1, p) \leq W(d_2^*, y_2, p) - W(d_2^*, y_1, p)$

(i.e., the effect of y on $W(d, y, p)$ is higher at d_2^* than at d_1^*). If $d_1^* > d_2^*$ the last inequality contradicts P4, which is equivalent to $W_{yd}(d, y, p) > 0$. Therefore we have $d_1^* \leq d_2^*$. Moreover, as $W_d(d_1^*, y_1, p) = 0$ for the first order condition, P4 also implies that $W_d(d_1^*, y_2, p) > 0$ and thus we conclude that $d_1^* < d_2^*$.

Next, we focus on the behavior of the propensity to save when y changes. The propensity to save is defined as:

$$\frac{s^*}{y} = \frac{b^*}{Ry} = 1 - \frac{d^*}{y}. \quad (43)$$

Proposition 3.

- (i) If $y < \hat{y}(p)$ then $\frac{s^*}{y} = 1$;
- (ii) If $y \geq \hat{y}(p)$ then $\frac{s^*}{y} < 1$;
- (iii) $\lim_{y \rightarrow \infty} \frac{s^*}{y} \rightarrow 1$.

Proof. Using Proposition (1), we have that if $y < \hat{y}(p)$ then $\frac{s^*}{y} = 1$ and if $y \geq \hat{y}(p)$ then $\frac{s^*}{y} < 1$. Then we define $\eta = d^*/y$ and write the first order condition of the optimization problem:

$$(1 - \eta) \left(\tau \frac{2d^*}{(1 + d^*)^2} + (2p - 1)(\beta + \tau) \frac{1}{1 + d^*} \right) - \frac{1}{2} \left(\frac{1 + d^*}{d^*} \beta + \tau \frac{d^*}{1 + d^*} + (2p - 1)(\beta + \tau) \right) \eta = 0. \quad (44)$$

By Proposition 2, d^* is an increasing function of income y . As $y \rightarrow \infty$, d^* tends either to some limit $\bar{d} < \infty$, or to $+\infty$. If it goes to \bar{d} , η goes to zero and s^*/y goes to 1. If d^* goes to $+\infty$ Eq. (44) becomes

$$2p(\beta + \tau)\eta = 0 \quad (45)$$

which implies $\eta = 0$ and $s^*/y = 1$.

Proposition 3 shows that the relationship between propensity to save and income has a U -shape: very rich people and poor people have a higher propensity to save than people with an average income.

Finally, we analyze the effects of changes in p , which is a measure of medical effectiveness.

Proposition 4.

- (i) For $y \geq \hat{y}(p)$:
 d^* is strictly increasing in p
 b^* and s^*/y are strictly decreasing in p ;
- (ii) $\hat{y}(p)$ is strictly decreasing in p .

Proof. See Appendix A.2.

According to Proposition 4 an increase in medical effectiveness has a positive effect on investment in health and a negative one on bequests and propensity to save out of income. Moreover, the threshold level above which people start investing in health decreases. This means that poor agents do not invest in health when medicine is not efficient, but may decide to invest in health once medical effectiveness increases.

As we explain in the next section, Proposition 3 and part (i) of 4 turn out to be different in the general model when first period consumption is present.

5.2. The Model with First Period Consumption

In the simplified model of Section 5.1, we prove a result important for modelling the rise of effective medicine: people with a sufficiently high income invest in medicine even when it is not effective (that is $p = 1/2$). However, there are two properties of the model in Section 5.1 which are undesirable. First, the propensity to save does not increase monotonically with income. Second, better medicine can increase the propensity to save only by reducing the share of income spent on health. These two properties are a direct consequence of having abstracted from first period consumption. In such a case, savings are negatively related to health expenditure through budget constraint Eq. (27). This is no longer true if first period consumption is introduced, because the relationship between savings and health expenditure depends on consumption in the first period through budget constraint Eq. (3).

This last consideration and more generally the need to assess the robustness of the results derived in Section 5.1, motivate us to include consumption in the first period. We perform such a task by way of numerical simulations.

With regard to the utility of bequests $u(b)$ and the effect of health investment on health status $v(d)$ and survival probability $f(d)$, we keep assumption (25). The functional form for the utility of consumption in the first period is the same as that for bequests: $u(c) = \sqrt{c}/2$.

Agents maximize the utility function with respect to b and d :

$$W(b, d) = \frac{\sqrt{y - \frac{b}{R} - d}}{2} + \frac{\sqrt{b}}{2} \left(\underbrace{\beta + \tau \frac{d^2}{(1+d)^2} + (2p-1)(\tau + \beta) \frac{d}{1+d}}_{\equiv \phi(d,p)} \right) \quad (46)$$

where we used the budget constraints Eq. (5) to express consumption c in the first period as a function of b and d . For the numerical simulation of this section, we put $\beta = 0.25$, $\tau = 0.1$ (this is the parameterization which we use in the simulation of the model in Section 6) and $R = 1$.

Figure 2 displays $\hat{y}(p)$. The function $\hat{y}(p)$ exists and is finite for all levels of p and in particular for $p = 1/2$. Moreover, $\hat{y}(p)$ decreases with p . Proposition 1 and Proposition 4 (ii) still hold true.

Simulations also show that investment in health increases with income and thus Proposition 2 is robust to the inclusion of consumption in the first period. Moreover, bequests and consumption in the first period turn out to be normal goods.

We look at health expenditure as a share of income and at the propensity to save and to consume. Figure 3 shows d^*/y and s^*/y as a function of y . c^*/y is presented in Figure 4. The behavior of d^*/y is similar to that described in Section 5.1: the share of income devoted

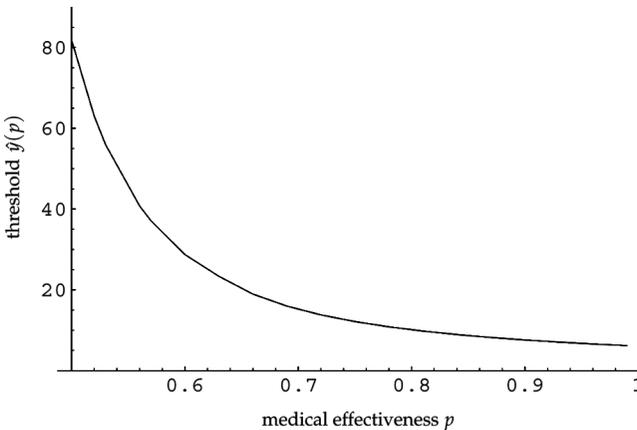


FIGURE 2 Income threshold $\hat{y}(p)$ which households invest in health as a function of medical effectiveness (p).

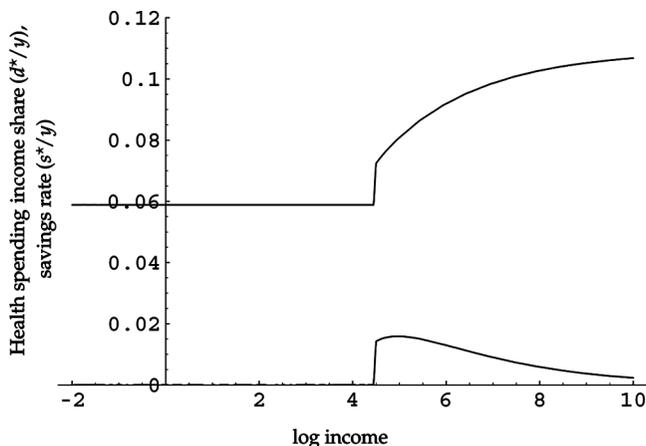


FIGURE 3 Health spending income share (d^*/y) and savings rate (s^*/y) as a function of income y in logarithmic scale.

to health expenditure is zero for people with an income below $\hat{y}(p)$, and then follows an inverse U -shape. This does not imply that b^*/y behaves analogously as in Section 5.1. The propensity to save is a constant below 1 for income levels up to $\hat{y}(p)$ and then it increases. This is due to the fact that c^*/y decreases for $y > \hat{y}(p)$. This can be understood intuitively by taking into account that d^* depends positively on y and

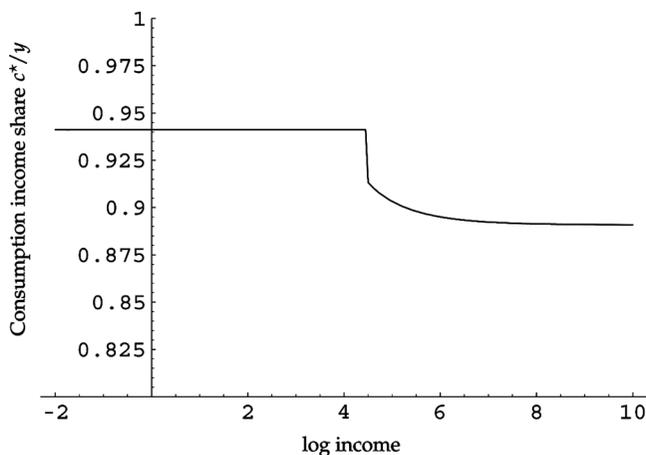


FIGURE 4 Consumption income share (c^*/y) as a function of income y in logarithmic scale.

that a higher d^* induces agents to increase the weight they put on the utility of the second period.

The fact that the propensity to save firstly increases introduces a mechanism of intergenerational transmission of inequality: rich people devote a larger fraction of their income to bequests than poor people; secondly determines a positive relationship between growth and inequality: the higher the share of income distributed to the rich, the higher the total amount of saving in the economy. This kind of relationship between growth and inequality is an assumption which is sometimes made in the growth literature for which we have now proposed a foundation based on the role of medicine (Galor and Moav (2004) who present a general discussion of all possible channels through which inequality could be related to growth).

Figure 5 displays the effect of an increase in medical effectiveness (that is a change from p to $p' > p$) on d^*/y , s^*/y and c^*/y . For income levels below $y(p')$, this effect is zero. For income levels above $y(p')$ and contrary to what is claimed in Proposition 4, both the share of income devoted to health expenditure and the propensity to save increase in response to an increase in medical effectiveness. Once again this is now possible because a rise in p causes c^*/y to drop for $y > y(p')$.

6. DYNAMICS

We presented the building blocks of our model economy and discussed in details the features of the individual optimization problem. In this section we discuss the dynamic implications of the results.

First we show that the model is a consistent explanation of the empirical evidence on life expectancy presented in Figure 1. Then we look at the implication of the model for growth, inequality, and the relationship between them. For this purpose we study the behavior over time of the propensity to save. *Ceteris paribus*, an increase in the propensity to save of at least one of the two classes implies an increase in aggregate savings and as a consequence an increase in the growth rate of income. As already mentioned, the higher the propensity to save of the rich compared to that of the poor, the higher inequality and the stronger the positive relationship between inequality and growth.

We do not aim to calibrate our two-period overlapping generation model to actual data. We only give a numerical example of the dynamics generated by the model. We retain the assumptions about the utility function used in Section 5.2 ($\beta = 0.25, \tau = 0.1$). One period is

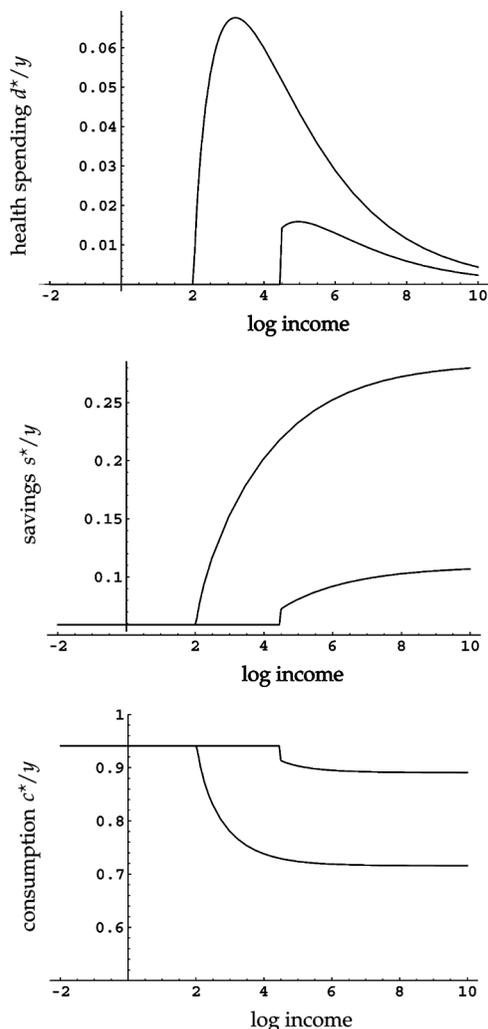


FIGURE 5 Effects of an increase in medical effectiveness p on income shares devoted to health spending (d^*/y), savings (s^*/y) and consumption (c^*/y) for different levels of income y in logarithmic scale.

30 years. The values for the parameters of the production function are $A = 0.525$, $\alpha = 1/3$, $\theta = 1/10$. As far as the evolution of medical knowledge is concerned, we initially want $p = 1/2$ and thus according to Eq. (15), we choose $M_0 < \bar{M}$. In particular we put $M_0 = 0$, $\bar{M} = 15$ and $\delta = 0.12$. For the function $g(M_t)$ we choose $0.3 \frac{M_t/10}{1+M_t/10}$. Thus

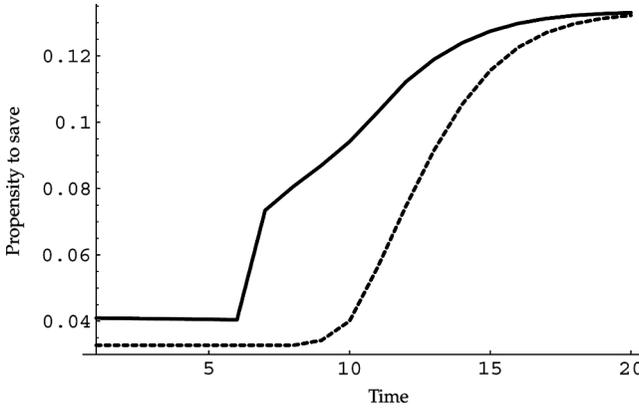


FIGURE 6 Propensity to save of landlords (solid) and commoners (dots).

$g'(M_t) > 0$, $g''(M_t) < 0$ and $\lim_{M_t \rightarrow \infty} g(M_t) = 0.3$. Finally we set the initial condition for the income of the two social classes: in period zero the income of the landlords is assumed to be higher than $\hat{y}(1/2)$, while the income of the commoners is fixed below $\hat{y}(1/2)$.

Figures 6 and 7 show the evolution over time of the propensity to save and of life expectancy for the landlords and the commoners. The behavior of inequality (defined here as the ratio of the income of the landlords to the income of the commoners) and of the growth rate of aggregate income are shown in Figures 8 and 9. Looking at these figures, we can distinguish three phases.

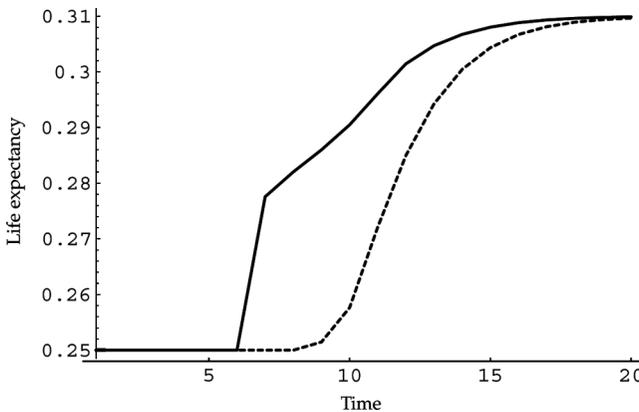


FIGURE 7 Life expectancy of landlords (solid) and commoners (dots).

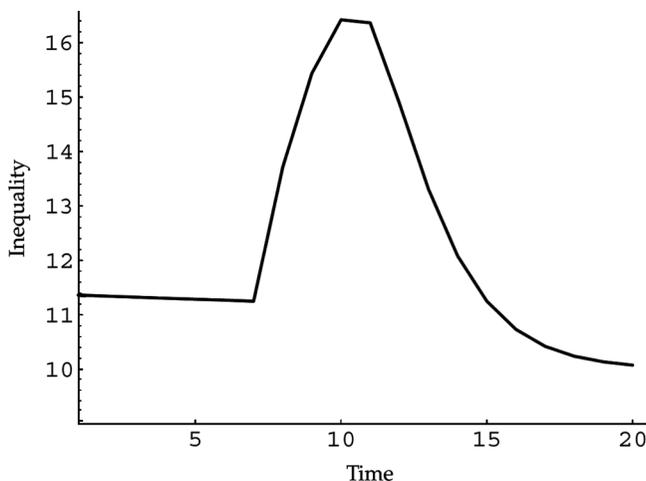


FIGURE 8 Inequality.

In the first phase, medicine is ineffective ($p = 1/2$) and has no effect in terms of life expectancy. Only the landlords invest in health. Their propensity to save is slightly higher than that of the commoners.

According to Eq. (15), medical experimental knowledge accumulates. At some point, medical experiments are numerous enough to generate an improvement in medical effectiveness. A *second phase* starts. The landlords go on investing in health, while the commoners

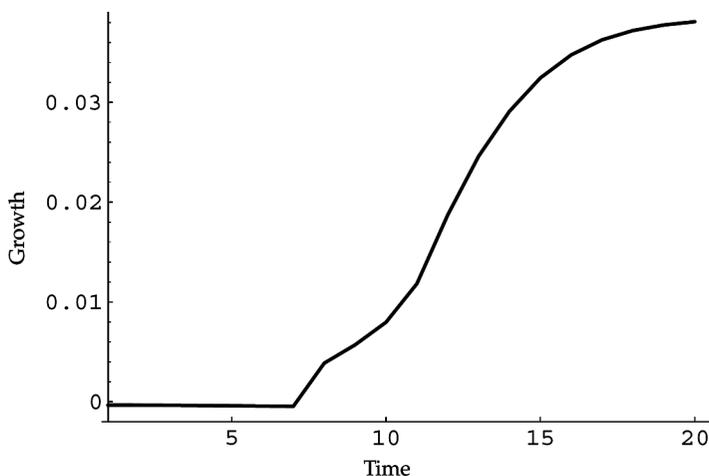


FIGURE 9 Growth.

still choose not to buy medical services. Life expectancy increases for the landlords and the mortality differential widens. Stimulated by longer lives, the propensity to save of landlords strongly increases over time. This boosts the growth rate of income and inequality. Moreover, the positive relationship between inequality and growth is strengthened.

Thanks to the growth in capital and the increase in medical effectiveness, the income of the commoners eventually reaches a threshold above which they also start to invest in health. In this *third phase* their life expectancy increases. Medical knowledge, stimulated by the growth in health expenditures, progresses rapidly. Asymptotically, the life expectancy of the poor catches up that of the rich, and they converge to a constant value. The propensity to save of the two groups follows the same path. As a consequence, inequality decreases over time and the positive relationship between growth and inequality disappears. In the long run, the growth rate of income tends to its balanced growth value.

We stress the importance for the inequality of the behavior of the propensities to save or bequeath. Inequality is always present, because the landlords are endowed with an additional factor of production, but it may be magnified depending on the differential in the propensity to bequeath across income groups. The fact that the propensity of the rich to bequeath is greater than that of the poor means that the former not only leave more in absolute terms to their offspring, but they also leave a larger *fraction* of their income. This channel of intergenerational transmission of inequality is absent when the differential in the propensity to bequeath is zero, because the rich and the poor leave the same *fraction* of their income to their offspring: in such a situation the only reason for the persistence of inequality *in the long run* is the difference in the ownership of land. In our dynamic simulation, the differential in the propensity to save is initially positive but low. Then it increases, giving rise to the jump in inequality. Finally it shrinks to the point at which it disappears, producing a level of inequality which is positive but lower than the initial one. The reason for the fall in inequality is due to the progressive disappearance of the channel of intergenerational transmission of inequality that passes through the differences in the propensity to bequeath.

From the simulation we draw two main conclusions concerning the implication for economic variables of our model economy. First, the evolution of medical effectiveness may have contributed to the acceleration in the growth rate in England during the eighteenth century. It may also have played an important role in shaping the evolution of inequality and of the relation between inequality and growth, both of which follow an inverted *U*. Initially, when the development process

starts, inequality increases and the positive relationship between inequality and growth is strong. After a further increase in output, there is a trend toward a more equal distribution of income and the positive relationship between inequality and growth disappears. This relationship between the level of development and inequality qualitatively resembles the well known Kuznets curve, for which our framework provides a theoretical justification.⁸ As far as the empirical evidence is concerned, the existence of the Kuznets curve in England between the eighteenth and nineteenth centuries seems to be corroborated by the evidence on income distribution (Lindert, 2000).⁹

7. ROBUSTNESS AND EXTENSIONS

We assume that there is no consumption in the *second* period:¹⁰ the only reason people save is to leave bequests to their offspring. The elderly, after having transferred their resources to their heirs at the beginning of their old age, do not make any more active choices. Their consumption is incorporated in the consumption of their adult children. This assumption is made for simplicity, but it is also a sufficiently good approximation of the historical period we are studying, because the young and the old often lived together in the same house and were part of the same decisional unit. This assumption is not crucial. The mechanism behind the choice of health investments would be unaffected by the introduction of consumption in the second period. The reason is that, from the point of view of the decision maker, consumption in the second period is of the same nature as bequests: it gives utility in the second period and is financed through savings.¹¹ Both the trade-offs between consumption in the first period and bequests and consumption in the first and in the second period are enhanced by being healthy in old age: people with a higher level of health investment (the rich) are

⁸Acemoglu and Robinson (2002) and the references therein give other explanations of the Kuznets curve.

⁹However, given the limited reliability of available data, to best guess the evolution of income inequality, the available information on income distribution should be integrated with an analysis of movements in factor-price ratios and in inequalities of wealth and of earnings (Lindert, 2000).

¹⁰As we explained at the beginning of Section 5, the removal of *first* period consumption is only done in Section 5.1 to provide intuitive explanations; in Section 5.2 consumption in the first period is re-introduced.

¹¹The utility function in this case is $E U(c_t^i, b_{t+1}^i, c_{t+1}^{i,o}, h_t^i) = u(c_t^i) + \phi(d_t^i, p_t)u(b_{t+1}^i, c_{t+1}^{i,o})$ where $c_{t+1}^{i,o}$ is old-age consumption. The intertemporal budget constraint is $y_t^i = c_t^i + d_t^i + \frac{b_{t+1}^i + c_{t+1}^{i,o}}{R_{t+1}}$.

more likely to substitute consumption in the first period with bequests and consumption in the second period, because they have a higher value of $\phi(d_t^i, p_t)$. Thus the propensity to save (and also the propensity to bequeath and to consume in the second period) should still be increasing with income, provided that an annuity market does not exist.¹² Furthermore there is still a threshold level of income $\hat{y}(p)$ such that people invest in health even when $p = 1/2$. As for Proposition 1 in Section 5.1, the key condition to have this result is that the expected gain from investing in health is positive. This condition still holds true after the introduction of consumption in the second period.¹³ Modifications in the opportunity costs of the investment in health do not affect the existence of $\hat{y}(p)$ but simply its level.

Another assumption concerns the exogeneity of fertility. Although it goes beyond the scope of our study, an explicit treatment of fertility choice could be an interesting extension, because it could give rise to differences in the fertility rate of the rich and the poor, allowing a simultaneous analysis of differential mortality and differential fertility. The effect of differential fertility (defined as the difference between the fertility of the rich and the poor) on inequality would depend on its sign. If the sign is positive, that is landlords have more children than commoners, differential fertility leads to a dilution of the mechanism of intergenerational transmission of inequality, while the opposite is true if the sign is negative. The sign of differential fertility depends on the specific way in which endogenous fertility is introduced into the model. In particular, if the cost of having children is mostly a time cost (as posited by La Croix and Doepke (2003)), differential fertility will be negative. If it is a cost in terms of goods, differential fertility will be positive (standard income effect). Another interesting issue appearing in a framework with endogenous fertility is the study of the trade-off between investment in own health against children. This trade-off is particularly relevant if children provide parental care in old age. Parents could decide to increase their own future well-being

¹²The role of the absence of an annuity market is explained by Chakraborty and Das (2005). Using this assumption, they present a model with second period consumption and bequests in which the propensity to bequeath and to consume in the second period increases with income. Again, the crucial difference with our work is the way in which we model the relationship between health investment and health status (the notion of medicine as an experience good) in order to study the issue of medical effectiveness. As mentioned for Eq. (11), this difference is captured by the specific functional form behind the term $\phi(d_t^i, p_t)$.

¹³With consumption in the second period, the only change to Eq. (41) and (42) concerns the replacement of $u(b)$ with $u(b, c^o)$. This change does not affect the reasoning presented there.

by investing in health or by having children. In turn, this choice may be affected by the comparison between the survival probability of infants and the effectiveness of medicine in reducing old age mortality.

Our model is related to Doepke and Zilibotti (2005), who studies the demise of the aristocracy and the emergence of the bourgeoisie as a new economic elite. One of our results is that aristocrats have a higher propensity to save than commoners. This result may seem at odds with Doepke and Zilibotti's (2005) model, which relies on the argument that the aristocracy have a lower propensity to save than the bourgeoisie, because the degree of individual patience is positively affected by the kinds of profession typically chosen by the bourgeoisie. However our model is not an alternative to Doepke and Zilibotti's (2005) but complementary to it. We are well aware that society is more complex than a simple two class representation; at the very least we have the aristocracy, the bourgeoisie, and the proletariat. We do not distinguish between the two latter social classes; the *only* reason for this choice is that we do not have data in Figure 1 for the specific life expectancy of the bourgeoisie. However our results still hold if we characterize the upper class as the one with the highest wage, without any need to refer to the ownership of land. The model can be reformulated by replacing the landlords with the bourgeoisie, modelled as people with a higher level of skill than the proletariat. In other terms, the purpose of our model is to compare the rich (who could be aristocrats or bourgeoisie) and the poor. An important element affecting the propensity to save of these two groups is the investment in health. At the same time, in order to study the dynamic evolution of the elite, it is useful to consider other factors affecting the propensity to save. While both the aristocracy and the bourgeoisie were able to buy expensive medical services, the bourgeoisie had a greater propensity to save due to a greater level of patience as suggested by Doepke and Zilibotti (2005).

8. CONCLUSION

No previous model has formally discussed the specific role played by the evolution of effective medical science, which has been viewed by some historians (see Section 3) as an important factor for understanding the evolution of life expectancy in England in the eighteenth and the nineteenth centuries.

Our first contribution is to build a model of medical spending suitable for a period, such as pre-industrial Europe, where medical effectiveness is very low. We modelled medicine as an experience good: medicine can be a good with probability p_t and a bad with probability $1 - p_t$, and agents know this probability distribution.

We give a rationale for individual health spending even when $p_t = 1/2$, that is when medicine was on average not effective. Combining this result with a simple learning process based on health expenditure, we give a consistent explanation of the path of life expectancy in England for the different income groups: initially there were no differences in life expectancy between the rich and the poor; then the life expectancy of the rich increased while the life expectancy of the poor remained stagnant; finally the life expectancy of the poor started increasing and eventually caught up with that of the rich.

We use our model economy to look at the implication for economic variables. First the rise of effective medicine, producing an increase in life expectancy and in savings, positively affects the growth rate of income. Secondly we show that the evolution over time of medical effectiveness plays a role in shaping inequality and the strength of the relationship between inequality and growth; inequality follows the same inverted U shape as the differential in mortality between the rich and the poor.

ACKNOWLEDGEMENT

We are grateful to Raouf Boucekkine and two anonymous referees for useful comments and suggestions. All the usual disclaimers apply. We acknowledge financial support from the Belgian French speaking community (Grant ARC 03/08-235 “New macroeconomic approaches to the development problem”) and the Belgian Federal Government (Grant PAI P5/21, “Equilibrium theory and optimization for public policy and industry regulation”).

REFERENCES

- Acemoglu, D. and Robinson, J.A. (2002). The political economy of the Kuznets curve. *Review of Development Economics*, 6(2): 183–203.
- Bardet, J.P. (1983). *Rouen au XVII^e et XVIII^e siècles*. Paris: SEDES.
- Blackburn, K. and Cipriani, G. (2002). A model of longevity, fertility and growth. *Journal of Economic Dynamics and Control*, 26(1): 187–204.
- Boucekkine, R., La Croix (de), D., and Licandro, O. (2003). Early mortality declines at the dawn of modern growth. *Scandinavian Journal of Economics*, 105(3): 401–418.
- Castello-Climent, A. and Domenech, R. (2008). Human capital inequality, life expectancy and economic growth. *The Economic Journal*, 118(528): 653–677.
- Cervellati, M. and Sunde, U. (2005). Human capital formation, life expectancy and the process of development. *American Economic Review*, 95(5): 1653–1672.
- Chakraborty, S. (2004). Endogenous lifetime and economic growth. *Journal of Economic Theory*, 116(1): 119–137.
- Chakraborty, S. and Das, M. (2005). Mortality, human capital and persistent inequality. *Journal of Economic Growth*, 10(2): 159–192.

- Doepke, M. and Zilibotti, F. (2005). Social class and the spirit of capitalism. *Journal of the European Economic Association*, 3: 516–524.
- Edlin, A.S. and Shannon, C. (1998). Strict monotonicity in comparative statics. *Journal of Economic Theory*, 81(1): 201–219.
- Fogel, R. (1994). Economic growth, population theory and physiology: The bearing of long-term processes on the making of economic policy. *American Economic Review*, 84(3): 369–395.
- Fogel, R. (2004). *The Escape from Hunger and Premature Death, 1700–2100*. Cambridge, UK: Cambridge University Press.
- Galor, O. and Moav, O. (2004). From physical to human capital accumulation: Inequality in the process of development. *Review of Economic Studies*, 71(4): 1001–1026.
- Galor, O. and Moav, O. (2005). Natural selection and the evolution of life expectancy. *CEPR Discussion Paper*. No. 5373.
- Galor, O. and Weil, D. (1999). From Malthusian stagnation to modern growth. *American Economic Review*, 89(2): 150–154.
- Galor, O. and Weil, D. (2000). Population, technology, and growth: From the Malthusian regime to the demographic transition and beyond. *American Economic Review*, 90(4): 806–828.
- Hollingsworth, T. (1977). Mortality in the British peerage families since 1600. *Population*, 32: 323–352.
- Johansson, R. (1999). Death and the doctors: Medicine and elite mortality in Britain from 1500 to 1800. *Cambridge Group for the History of Population and Social Structure Working Paper Series*, No. 7.
- Kalemli-Ozcan, S. (2002). Does mortality decline promote economic growth? *Journal of Economic Growth*, 7(4): 411–439.
- La Croix (de), D. and Doepke, M. (2003). Inequality and growth: Why differential fertility matters. *American Economic Review*, 93(4): 1091–1113.
- La Croix (de), D. and Licandro, O. (1999). Life expectancy and endogenous growth. *Economics Letters*, 65(2): 255–263.
- Lagerloef, N.P. (2003). From Malthus to modern growth: Can epidemics explain the three regimes? *International Economic Review*, 44(2): 755–777.
- Lindemann, M. (1999). *Medicine and Society in Early Modern Europe*. Cambridge, UK: Cambridge University Press.
- Lindert, P.H. (2000). Three centuries of inequality in Britain and America. In A.B. Atkinson and F. Bourguignon (Eds.), *Handbook of Income Distribution* (Vol. 1). Amsterdam: Elsevier Science, 167–216.
- Mokyr, J. (1993). Technological progress and the decline of European mortality. *American Economic Review Papers and Proceedings*, 83(2): 324–330.
- Morand, O. (2004). Economic growth, longevity and the epidemiological transition. *European Journal of Health Economics*, 5(2): 166–174.
- Nicolini, E. (2004). Mortality, interest rates, investment, and agricultural production in 18th century England. *Explorations in Economic History*, 41(2): 130–155.
- Perrenoud, A. (1975). L'inégalité sociale devant la mort à Genève au XVII^e siècle. *Population*, 30: 221–243.
- Porter, R. (1995). *Disease, Medicine, and Society in England, 1550–1860. Vol. 3, New Studies in Economic and Social History*. Cambridge, UK: Cambridge University Press.
- Sala-I-Martin, X. (2002). Unhealthy people are poor people and vice versa. Keynote address at the European Conference on Health Economics of the International Health Economics Organization, Paris, July 7.
- Siraisi, N. (1990). *Medieval and Early Renaissance Medicine: An Introduction to Knowledge and Practice*. Chicago: University of Chicago Press.

- Smith, C. (2005). Origin and uses of *primum non nocere* – above all, do no harm! *The Journal of Clinical Pharmacology*, 45: 371–377.
- Weatherall, M. (1996). Drug treatment and the rise of pharmacology. In Cambridge Illustrated History: Medicine. Cambridge, UK: Cambridge University Press.
- Wrigley, E.A., Davies, R.S., Oeppen, J.E., and Schofield, R.S. (1997). *English Population History from Family Reconstitution: 1580–1837*. Cambridge, UK: Cambridge University Press.
- Zhang, J., Zhang, J., and Lee, R. (2001). Mortality decline and long-run economic growth. *Journal of Public Economics*, 80(3): 485–507.

A APPENDIX

A.1. Proof of Property P6

From Eq. (32) we compute

$$\Gamma_p(d, y, p) \frac{\sqrt{R}}{2} (\beta + \tau) \left(2\sqrt{(y-d)} \frac{1}{(1+d)^2} - \frac{1}{\sqrt{(y-d)}} \frac{d}{(1+d)} \right). \quad (47)$$

Thus

$$\Gamma_p(d, y, p) > 0 \Leftrightarrow 2(y-d) - d(1+d) > 0. \quad (48)$$

Using the first order condition derived from Eq. (32), inequality (48) at $d = d^*$ is written as:

$$\begin{aligned} \Gamma_p(d^*, y, p) > 0 &\Leftrightarrow \frac{\beta + \tau \frac{d^{*2}}{(1+d^*)^2} + (2p-1)(\beta + \tau) \frac{d^*}{1+d^*}}{\tau \frac{2d^*}{(1+d^*)^3} + (2p-1)(\beta + \tau) \frac{1}{(1+d^*)^2}} - d^*(1+d^*) > 0 \\ &\Leftrightarrow (\beta - \tau)d^2 + 2\beta d + \beta > 0 \end{aligned} \quad (49)$$

which always holds true because we have assumed $\beta \geq \tau$. Moreover from Eq. (47) $\Gamma_p(d, y, p)$ is decreasing in d . Thus $\Gamma_p(d, y, p) > 0 \forall d < d^*$.

A.2. Proof of Proposition 4

- (i) The proof is similar to the one used for Proposition 2. Let us consider $d_1^* = d^*(y, p_1)$. If p_1 increases to p_2 , then $d_2^* = d^*(y, p_2)$ cannot be below d_1^* . Indeed, by definition of d_1^* and d_2^* , we always have $W(d_1^*, y, p_1) \geq W(d_2^*, y, p_1)$ and $W(d_1^*, y, p_2) \leq W(d_2^*, y, p_2)$ and as a consequence $W(d_1^*, y, p_2) - W(d_1^*, y, p_1) \leq W(d_2^*, y, p_2) - W(d_2^*, y, p_1)$ (the effect of p on $W(d, y, p)$ is higher at d_2^* than at d_1^*). If $d_1^* \geq d_2^*$ the last inequality contradicts P6, which is equivalent to $W_{pd}(d, y, p_1) > 0 \forall d < d_1^*$. Therefore, $d_1^* \leq d_2^*$. Moreover, as $W_d(d_1^*, y, p_1) = 0$ for the first order condition, P6 also implies that

$W_d(d_1^*, y, p_2) > 0$ and thus $d_1^* < d_2^*$. Using the budget constraint (27) and Eq. (43), we also determine how b^* and $\frac{s^*}{y}$ react to a change in p .

- (ii) By definition (39), $W(d, \hat{y}(p), p) - W(0, \hat{y}(p), p) \geq 0$ for at least one $d \in (0, y)$.

Using Eq. (30) to compute

$$\begin{aligned}
 W(d, y, p) - W(0, y, p) &= \frac{\sqrt{R(y-d)}}{2} \\
 &\times \left(\beta + \tau \frac{d^2}{(1+d)^2} + (2p-1)(\beta + \tau) \frac{d}{1+d} \right) \\
 &- \frac{\sqrt{Ry}}{2} \beta
 \end{aligned} \tag{50}$$

we can see that $W(d, y, p) - W(0, y, p)$ is strictly increasing in y and p .

Thus if p increases to p' then $W(d, \hat{y}(p), p') - W(0, \hat{y}(p), p') > 0$. As $W(d, y, p)$ is continuous in y , we choose $y < \hat{y}(p)$ and have $W(d, y, p') - W(0, y, p') \geq 0$. As a consequence, $\hat{y}(p') < \hat{y}(p)$.