THE ECONOMIC JOURNAL

The Economic Journal, **129** (*January*), 338–374 DOI: 10.1111/ecoj.12577
Advance Access Publication Date: 19 May 2018

A THEORY OF SOCIO-ECONOMIC DISPARITIES IN HEALTH OVER THE LIFE CYCLE*

Titus J. Galama and Hans van Kippersluis

Motivated by the observation that medical care explains only a relatively small part of the socio-economic status (SES)-health gradient, we present a life-cycle model that incorporates several additional behaviours that potentially explain (jointly) a large part of observed disparities. As a result, the model provides not only a conceptual framework for the SES-health gradient but more generally an improved framework for the production of health. We derive novel predictions from the theory by performing comparative dynamic analyses. More generally, our comparative dynamic method can be applied to models of similar form, e.g. human capital, health deficits, firm investment, to name a few.

Disparities in health across socio-economic status (SES) groups – often called the SES-health gradient – are substantial. For example, Case and Deaton (2005) show how in the US, a 20-year-old low-income (bottom quartile of family income) male, on average, reports to be in similar health as a 60-year-old high-income (top quartile) male. In cross sectional data, the disparity in health between low and high SES groups appears to increase over the life cycle until ages 50–60, after which it narrows. These patterns exist across a wide range of measures of SES, such as education and wealth, and across all indicators of health, including the onset of chronic diseases, disability and mortality (Adler *et al.*, 1994; Marmot, 1999). The pattern is also remarkably similar between countries with relatively low levels of protection from loss of work and health risks, such as the US, and those with stronger welfare systems, such as the Netherlands (Case and Deaton, 2005; Smith, 2007; Van Kippersluis *et al.*, 2010).

* Corresponding author: Titus J. Galama, University of Southern California, Dornsife Center for Economic and Social Research (CESR), USA. Email: galama@usc.edu.

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Numbers K02AG042452, R01AG037398, R01AG030824, RF1AG055654, P30AG012815 and P01AG022481. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Van Kippersluis thanks Netspar (Health and Income, Work and Care across the Life cycle II), and the Netherlands Organization of Scientific Research (NWO Veni grant 016.145.082) for support. Galama is grateful to the School of Economics of Erasmus University Rotterdam for a Visiting Professorship in the Economics of Human Capital. We thank the Editor, two anonymous reviewers, Mauricio Avendano, Megan Beckett, Dan Belsky, Jan Brinkhuis, Eileen Crimmins, Michael Grossman, Bas Jacobs, Arie Kapteyn, Valter Longo, Joao Pedro Magalhaes, Erik Meijer, Owen O'Donnell, Rosalie Pacula, Bastian Ravesteijn, Peter Savelyev, Eddy van Doorslaer, Tom Van Ourti, Arthur van Soest and participants in the Health Inequality Network (HINet) meeting at the University of Chicago, the 8th iHEA World Congress on Health Economics in Toronto, the 12th Viennese Workshop on Optimal Control, Dynamic Games and Nonlinear Dynamics, the 9th Annual SHESG Conference at the University of Memphis and the 2011 RAND Conference on Health, Aging and Human Capital, as well as seminar participants at Erasmus University, the Harvard Center for Population and Development Studies, the London School of Economics, the RAND Corporation, the University of Chicago, the University of Lausanne, the University of Southern California and the University of Tilburg, for useful discussions and comments. We thank Timothy Halliday and Hui He for generously sharing their health investment estimates from the Medical Expenditure Panel Survey.

Recent significant contributions to the understanding of socio-economic disparities in health have concentrated on the identification of causal effects, but have stopped short of uncovering the underlying mechanisms that produce the causal relationships. For example, education is found to have a causal protective effect on mortality (Lleras-Muney, 2005) but it is not known exactly how the more educated achieve their health advantage (Cutler and Lleras-Muney, 2010).

Case and Deaton (2005) argue that it is extremely difficult to understand the relationships between health, education, income, wealth and labour-force status without some guiding theoretical framework. Integrating the roles of proposed mechanisms and their long-term effect into a theoretical framework allows researchers to disentangle the differential patterns of causality and assess the interaction between mechanisms. Such understanding is essential in designing effective policies to reduce disparities (Deaton, 2002). It is no surprise then that several authors (Case and Deaton, 2005; Cutler *et al.*, 2011) have pointed to the absence of a theory of SES and health over the life cycle and have emphasised the importance of developing one.

This article develops a conceptual framework for health and SES in which the SES-health gradient is the outcome of rational (constrained) individual choices made over the life cycle. The article makes two main contributions. The first main contribution is of a fundamental nature and consists of extending the canonical human-capital model for the demand for health (Grossman, 1972*a*,*b*) in two ways. First, we employ a relatively straightforward extension by allowing for decreasing returns to scale in health investment of the health-production process (see Galama and Kapteyn, 2011; Galama and van Kippersluis, 2013; Galama, 2015, for the reasoning behind this extension). Second, motivated by the observation that differences in medical care usage explain only a small part of the health gradient (Adler *et al.*, 1993), we conduct an extensive review of the literature from multiple disciplines to identify the most important mechanisms through which socio-economic characteristics such as wealth, earnings and education, interact with health. We then include several additional decisions regarding health (besides health investment), such as choices regarding lifestyle (exercise, healthy/unhealthy consumption), working conditions, labour-force participation and longevity, as mechanisms generating disparities in health. In doing so, we develop a comprehensive theory of the SES-health gradient, by integrating the most important interactions between health, longevity, health behaviour and SES (wealth, education and earnings) during adulthood.[1]

[1] Several papers contain components of our generalised theory of health. The Grossman model (Grossman, 1972*a*,*b*) contains health and health investment and interactions with earnings, and wealth but lacks other health behaviours. Ehrlich and Chuma (1990) were the first to introduce endogenous longevity, but their model does not include other decisions besides health investment. Forster (2001) models the relation between health, longevity, healthy consumption and unhealthy consumption but does not model health investment, wealth accumulation or job conditions. Case and Deaton (2005) include unhealthy consumption as well as physical effort on the job, but do not model longevity. In work independently developed around the same time as our theory, Dalgaard and Strulik (2014), Strulik (2015) and Dalgaard and Strulik (2017) present the so-called health-deficit models that contain health and health investment and interactions with earnings, labour-force participation and wealth, but lack other health behaviours. Strulik (2016) calibrates a related theoretical model in which health behaviour and SES influence health and longevity, but he does not model job conditions.

Because of the inclusion of a rich set of endogenous health behaviours, endogenous health and endogenous longevity, the theory can be employed to analyse the value of health as well as the value of life. Previous papers employing a life-cycle model for the value of life include Rosen (1988), Ehrlich (2000), Becker (2007) and Kuhn *et al.* (2015). In contrast to these papers, but in line with Murphy and Topel (2006), we distinguish between the value of health and the value of longevity. We go beyond Murphy and Topel (2006) by treating health and longevity endogenously: individuals seek to optimise both health and longevity. We find that the value of health and the value of life are distinct concepts that vary with age and by SES in distinct ways. For example, calibrated simulations suggest the value of life decreases near the end of life whereas the value of health increases. Assessments of the value of a statistical life generally involve investigating how much individuals need to be paid for taking a certain risk of death (usually in a setting of hazardous work). In practice, such estimates may capture the monetary value (compensating variation) of both the effect of changes in health and changes in longevity. In our theory, individuals are willing to engage in a certain amount of unhealthy consumption or job hazards for the instantaneous benefits it provides, as long as these benefits outweigh the associated health cost: the reduction in life-time utility due to health loss. Our theory suggests that unhealthy job conditions as well as risky health behaviours may be used in empirical work to evaluate both the value of health and the value of life.

Our second main contribution consists of deriving detailed predictions from the theory by performing comparative dynamic analyses of the effects of wealth, earnings, education and health on health behaviour and longevity. We are the first to perform such analyses analytically for a comprehensive theory with multiple health behaviours (to better model health) and multiple dimensions of SES (to model disparities in health for several measures of SES).[2] The comparative dynamic analyses not only deliver novel predictions regarding the SES-health gradient (see below), but more generally provide an alternative and complementary method to calibration and/or estimation in exploring model characteristics. They provide a method for researchers to explore their own research questions of interest. The method can be applied to models of similar form, for example, human capital, health deficits, firm investment, rational addiction, habit formation and resource extraction, to name a few. This work is therefore potentially also relevant to these and other areas of economics.

The comparative dynamic analyses provide insight into the mechanisms through which SES and health interact, and generate novel testable predictions. We highlight a few here and discuss these and several others more extensively in Section 3.

First, wealth, earnings and education affect health behaviour by increasing the marginal value of health relative to the marginal value of wealth. Intuitively, wealth,

[2] Ehrlich and Chuma (1990) generate a set of directional predictions (their table 3), based on a pioneering comparative dynamic analysis of the Grossman model with endogenous longevity. However, while they are able to generate the sign of the effects (broadly whether an effect is more or less likely to be positive or negative), they do not present dynamic results. Further, they do not consider other health behaviours besides health investment. Ried (1998) presents a comparative dynamic analysis of the Grossman model, but his model is limited to the linear case, which has distinct properties (Ehrlich and Chuma, 1990; Galama, 2015). Eisenring (1999) presents a comparative dynamic analysis of a simple model without consumption, wealth accumulation or health behaviours.

earnings, and the higher earnings associated with education relax the budget constraint. At higher levels of wealth, and hence consumption, only limited marginal utility is gained from additional consumption and it is not beneficial to consume more. In contrast, health extends length of life, providing additional time during which consumption, leisure, and health can be enjoyed. This leads to higher SES individuals to place a higher value on their health and to invest more in it (Becker, 2007; Hall and Jones, 2007).

A higher marginal value of health, in turn, increases the marginal benefits of healthy consumption, and the marginal costs of unhealthy working (and living) environments, and unhealthy consumption. This leads to healthier behaviour and gradually to greater health advantage with age. The more rapidly worsening health of low SES individuals may lead to early withdrawal from the labour force and associated lost earnings, further widening the gradient in early and mid-age. The model allows for a subsequent narrowing of the SES-health gradient, due to mortality selection and potentially because low SES individuals increase their health investment and improve their health behaviour faster as a result of their more rapidly worsening health. Our model is thus able to replicate the life-cycle patterns of the SES-health gradient.

Second, we predict a central role for our concept of a 'health cost' of unhealthy behaviours. The health cost is the marginal value (in terms of life-time utility) of health lost due to detrimental health behaviours. It takes into account all future consequences of current health behaviour. As a result of differences in the health cost, our theory predicts that high SES individuals are more likely to drink moderately but less likely to drink heavily (Van Kippersluis and Galama, 2014); and that individuals are willing to accept unhealthy working conditions in mid-life, given the high monetary benefits during those years, but that their willingness declines later in life due to an increasing health cost. Thus, the concept of a health cost has potential for explaining variation in health behaviours over the life cycle and across SES groups.[3]

Third, we predict that the ability to postpone death (endogenous longevity) is crucial in explaining observed associations between SES and health. Absent the ability to extend life (fixed horizon), associations between SES and health are small. If, however, life can be extended, SES and health are positively associated and the greater the degree of life extension, the greater is their association. The intuition behind this result is that the horizon (longevity) is a crucial determinant of the return to investments in

[3] While the concept of a health cost (benefit) of unhealthy (healthy) behaviour is not new, explicit theoretical modelling is, and so is our formal definition of the concept. The literature on the value of a statistical life (Viscusi and Aldy, 2003) focuses on the cost of reductions in life (mortality) rather than in health (morbidity) as in our theory. Even the seminal theory of rational addiction (Becker and Murphy, 1988), while arguing conceptually for an effect of unhealthy addictive consumption on health, does not explicitly model this effect (see Jones *et al.*, 2014, for an exception). To the best of our knowledge, only Forster (2001), Case and Deaton (2005), Strulik (2016) and Schuenemann *et al.* (2017) have previously explicitly modelled behaviour as a choice variable affecting health. Case and Deaton's (2005) model, however, focuses on the linear case, which has distinct properties (Ehrlich and Chuma, 1990; Galama, 2015). And while Forster (2001), Strulik (2016) and Schuenemann *et al.* (2017) model a health cost/health benefit by allowing consumption to affect health, our comparative dynamic analyses allow us to formally define, discuss and analyse the concept, investigate how it influences various health behaviours and the health gradient, and make predictions.

health. This suggests that in settings where it is difficult for wealthier, higher income and higher educated individuals to increase life expectancy (e.g. due to a high disease burden, competing risks, low efficiency of health investment, etc.), health disparities across socio-economic groups would be smaller.

These are just a few examples of how the theory can be used as a conceptual framework in conjunction with the comparative dynamic analyses to generate testable predictions for the complex relationships between SES and health. The theory is rich, and it is impossible to produce an exhaustive list of its possible uses. Researchers can use the theory and detailed comparative dynamic analyses presented here as a template to study their own questions of interest.

The article is organised as follows. Section 1 reviews the literature on health disparities by SES to determine the essential components required in a theoretical framework. Developing a theory requires simplification and a focus on the essential mechanisms relating SES and health. To keep the model relatively simple, we focus on explaining health disparities in adulthood.[4] We highlight potential explanations for the SES-health gradient that: (*i*) explain a large part of the gradient; and (*ii*) are relatively straightforward to include in our theoretical framework. Based on these principles, we develop our theoretical formulation in Section 2. Section 3 presents dynamics and calibrated simulations, Section 4 presents comparative dynamic analyses and makes predictions and Section 5 summarises and concludes.

1. Components of a Theory of the Gradient

In this Section, we review the empirical literature to determine the essential components of a theory of health disparities by SES in adulthood. Based on these findings, we present our theoretical formulation.

A significant body of research across multiple disciplines (including epidemiology, sociology, demography, psychology, evolutionary biology and economics) has been devoted to documenting and explaining the substantial disparity in health between low and high SES groups. The pathways linking the various dimensions of SES to health are diverse: some cause health, some are caused by health and some are jointly determined with health (Cutler *et al.*, 2011). Several key findings can be identified.

1.1. *Medical Care*

Utilisation of medical services and access to care explain only a relatively small part of the association between SES and health (Adler *et al.*, 1993). Therefore, additional mechanisms, besides medical care, have to be included in the model.

[4] James Heckman and colleagues have emphasised the role of childhood cognitive and non-cognitive abilities in determining both education and health outcomes in later life (Cunha and Heckman, 2007; Heckman, 2007; Conti *et al.*, 2010; Campbell *et al.*, 2014; Almond and Currie, 2011), and there is strong evidence that parental, especially maternal, SES influences the evolution of child health (Currie, 2009), suggesting that part of the SES-health gradient may be determined very early in life.

1.2. *Work Environment and Lifestyle*

Epidemiological research highlights the importance of lifestyles (e.g. smoking, drinking, caloric intake and exercise), psychosocial and environmental risk factors, neighbourhood social environment, acute and chronic psychosocial stress, social relationships and supports, sense of control, fetal and early childhood conditions, and physical, chemical, and psychosocial hazards and stressors at work (House *et al.*, 1994; Lynch *et al.*, 1997).

During adulthood, two of those mechanisms appear to be of particular importance: (*i*) working conditions; and (*ii*) lifestyles. Using three different data sets from the UK and the US, House *et al.* (1994) find that features of the psychosocial working environment, social circumstances outside work and health behaviour jointly account for much of the social gradient in health. Some epidemiological studies suggest that around two-thirds of the social gradient in health deterioration could be explained by working environment and lifestyle factors alone (Borg and Kristensen, 2000). Low SES individuals more often perform risky, manual labour than high SES individuals and their health deteriorates faster as a consequence (Marmot *et al.*, 1997; Ravesteijn *et al.*, 2013). Case and Deaton (2005) find that those who are employed in manual occupations have worse health than those who work in professional occupations and that the health effect of occupation operates at least in part independently of the personal characteristics of the workers. Extensive research further suggests an important role of lifestyle factors, particularly smoking, in explaining SES disparities in health (Mackenbach *et al.*, 2004). Fuchs (1986) argues that in developed countries, it is personal lifestyles that cause the greatest variation in health.

1.3. *Education*

Education appears to be a key dimension of SES and studies suggest education has a causal protective effect on health and mortality (Lleras-Muney, 2005; Conti *et al.*, 2010; Van Kippersluis *et al.*, 2010).[5] Education increases wages (Mincer, 1974), thereby enabling purchases of health investment goods and services (though higher wages also increase the opportunity cost of time). Education potentially increases the efficiency of medical and preventive care usage and time inputs into the production of health investment (Grossman, 1972*a*,*b*). The higher educated are also better able at managing their diseases (Goldman and Smith, 2002), and benefit more from new knowledge and new technology (Lleras-Muney and Lichtenberg, 2005; Glied and Lleras-Muney, 2008).

1.4. *Financial Measures of SES*

Financial measures of SES may have a more limited impact on health than education. Smith (2007) finds no effect of financial measures of SES (income, wealth and change in wealth) on changes in health. Cutler *et al.* (2011) provide an overview of empirical findings and conclude that the evidence points to no, or a very limited, impact of income

[5] Yet, see, e.g. Albouy and Lequien (2009) and Clark and Royer (2013) who could not establish a causal effect of education on mortality.

or wealth on health (Michaud and van Soest, 2008). Yet, this view is not unequivocally accepted. For example, Lynch *et al.* (1997) suggest that accumulated exposure to economic hardship causes bad health, and Herd *et al.* (2008) argue that there might be causal effects of financial resources on health at the bottom of the income or wealth distribution. Income and wealth enable purchases of medical care and thereby potentially allow for better health maintenance. Further, more affluent workers may choose safer working and living environments since safety is a normal good (Viscusi, 1978). But, higher wages are also associated with higher opportunity costs, which would reduce the amount of time devoted to health maintenance.

1.5 *Health and Labour-force Withdrawal*

In the other direction of causality, studies have shown that perhaps the most dominant causal relation between health and dimensions of SES in adulthood is the causal impact that poor health has on one's ability to work, and hence produce income and wealth (Case and Deaton, 2005; Smith, 2007). Healthy individuals are also more productive and earn higher wages (Currie and Madrian, 1999).

1.6 *Joint Determination*

Fuchs (1986) has argued that the strong correlation between SES and health may be due to differences in the time preferences of individuals, which affects investments in both education and health. Cutler and Lleras-Muney (2008) argue that differences in individual preferences (risk aversion and discount rates) appear to explain only a small portion of the SES-health gradient, but they also note that preferences are difficult to measure and that preferences with respect to health may differ from preferences with respect to finance. Other third factors known to contribute to the correlation between SES and health are cognitive and non-cognitive skills, in particular conscientiousness and self-esteem (Auld and Sidhu, 2005; Chiteji, 2010; Conti *et al.*, 2010; Savelyev, 2014).

1.7. *Gradient over the Life Cycle*

Health inequalities are largest in mid-life and narrow in later life. The literature provides competing explanations for this pattern. The cumulative advantage hypothesis states that health inequalities emerge by early adulthood and subsequently widen as economic and health advantages of higher SES individuals accumulate (House *et al.*, 1994; Ross and Wu, 1996; Lynch, 2003). Any apparent narrowing of SES inequalities in late life is largely attributed to mortality selection, i.e. lower SES people are more likely to die which results in an apparently healthier surviving disadvantaged population.[6] The competing age-as-leveler hypothesis maintains that later in life deterioration in health becomes more closely associated with age than with any other predictor, i.e. through

[6] Beckett (2000) and Baeten *et al.* (2013), however, have demonstrated that the convergence in health inequalities in later life cannot be explained entirely, or even mostly, by mortality selection.

a greater equalisation of health risks (House *et al.*, 1994), with the result that the SES-health gradient narrows.

2. Theory

2.1. *Theoretical Formulation*

In this Section, we formalise the above discussion on the features of a theoretical framework for the SES-health gradient over the life cycle. The aim is to understand the SES-health gradient as the outcome of rational constrained individual behaviour.

A natural starting point for a theory of the relation between health and SES is a model of life-cycle utility maximisation. Our model is based on the Grossman model of the demand for health (Grossman, 1972*a*,*b*, 2000) in continuous time (Wagstaff, 1986*a*; Ehrlich and Chuma, 1990; Zweifel and Breyer, 1997; Galama, 2015). The Grossman model provides a framework for the interrelationship between health, financial measures of SES (wealth, wages and earnings), the demand for consumption, the demand for medical goods and services and the demand for time investments in health (e.g. visits to the doctor, exercise). Health increases earnings (through reduced sick time) and provides utility. We add six additional features to the model.

First, we assume decreasing-returns-to-scale (DRTS) in investment of the health-production process. This addresses the degeneracy of the solutions for investment and health that characterises commonly employed linear investment models (Ehrlich and Chuma, 1990; Galama, 2015). It is further attractive in that the health-production process is generally thought of as being subject to diminishing returns (Wagstaff, 1986*b*).

Second, individuals choose their level of 'job-related health stress', which can be interpreted broadly, ranging from physical working conditions (e.g. hard or risky labour) to psychosocial aspects of work (e.g. low social status, lack of control, repetitive work, etc.) that are detrimental to health. Individuals may accept risky and/or unhealthy work environments, in exchange for higher pay (Muurinen, 1982; Case and Deaton, 2005), i.e. a compensating wage differential (Smith, 1776; Viscusi, 1978).

Third, we allow consumption patterns to affect the health deterioration rate (Grossman, 1972*b*; Forster, 2001; Case and Deaton, 2005; see Cawley and Ruhm, 2012, for a review of theoretical models of health behaviours). We distinguish healthy consumption (such as the consumption of healthy foods, sports and exercise) from unhealthy consumption (such as smoking, excessive alcohol consumption). Healthy consumption provides utility, and is associated with health benefits in that it lowers the health deterioration rate. We interpret healthy consumption broadly to include decisions regarding housing and neighbourhood.[7] Unhealthy consumption provides consumption benefits (utility) but increases the health deterioration rate.

[7] Living in an affluent neighbourhood is an expensive, yet health-promoting and utility-generating choice. It is a constrained choice: low SES individuals cannot afford to live in more affluent areas.

Fourth, the effect of education on income is included in a straightforward manner by assuming a Mincer-type wage relation, in which earnings are increasing in the level of education and in the level of experience of workers (Mincer, 1974).

Fifth, individuals endogenously optimise length of life (Ehrlich and Chuma, 1990). Longevity is an important health outcome and differential mortality by SES may explain part of the narrowing of the gradient in late life. Moreover, length of life is an essential horizon that determines the duration over which the benefits of health investments and healthy behaviours can be reaped.

Last, we include leisure, which jointly with sick time and time inputs into health investment and into health behaviour allows for the modelling of an implicit retirement decision. As health declines, increased sick time and increased demand for time inputs into health investment and healthy behaviour reduce the amount of time that can be devoted to work, capturing possible reverse causality from health to labour force participation, and thereby financial measures of SES.

Individuals maximise the life-time utility function:

$$\int_0^T U(t)e^{-\beta t}\mathrm{d}t, \tag{1}$$

where time t is measured from the time an individual has completed her education and joined the labour force (e.g. around age 25 or so), T denotes total lifetime (endogenous), β is a subjective discount factor and individuals derive utility $U(t) \equiv U[C_h(t), C_u(t), L(t), H(t)]$ from healthy consumption $C_h(t)$, unhealthy consumption $C_u(t)$, leisure $L(t)$ and health $H(t)$. Utility is assumed to be strictly concave and increasing in its arguments.

Healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption are produced following a Becker (1965) home-production model by combining goods, $X_h(t)$ and $X_u(t)$, purchased in the market and own time inputs, $\tau_{C_h}(t)$ and $\tau_{C_u}(t)$:

$$C_h(t) \equiv C_h[X_h(t), \tau_{C_h}(t); \mu_{C_h}(t)], \tag{2}$$

$$C_u(t) \equiv C_u[X_u(t), \tau_{C_u}(t); \mu_{C_u}(t)], \tag{3}$$

where $\mu_{C_h}(t)$ and $\mu_{C_u}(t)$ are (exogenous) efficiency factors.

The objective function (1) is maximised subject to three constraints. The first relates to the production of health capital:

$$\frac{\partial H(t)}{\partial t} = I[m(t), \tau_m(t); E] - d(t). \tag{4}$$

Health $H(t)$ can be improved through health production $I[m(t),\tau_m(t); E]$. Goods and services $m(t)$ (e.g. medical care) as well as own time inputs $\tau_m(t)$ (e.g. exercise, time spent visiting a doctor, etc.) are used in the production of health $I[m(t), \tau_m(t); E]$. We assume the following functional form:

$$I[m(t), \tau_m(t); E] \equiv \mu_I(t; E)\, m(t)^{\alpha_I} \tau_m(t)^{\beta_I}, \tag{5}$$

where the efficiency of the health-production process $\mu_I(t; E)$ is assumed to be a function of the consumer's stock of knowledge E as the more educated are assumed to be more

efficient consumers and producers of health (Grossman, 1972*a*,*b*). The health-production function $I[m(t), \tau_m(t); E]$ is assumed to exhibit DRTS ($0 < \alpha_I + \beta_I < 1$).

Health (see (4)) deteriorates at the health deterioration rate $d(t) \equiv d[t, C_h(t), C_u(t), z(t), H(t); \xi(t)]$. The health deterioration rate depends endogenously on healthy consumption $C_h(t)$, unhealthy consumption $C_u(t)$, job-related health stress $z(t)$, and health $H(t)$.[8] Consumption can be healthy ($\partial d/\partial C_h \leq 0$; e.g. healthy foods, healthy neighbourhood) or unhealthy ($\partial d/\partial C_u > 0$; e.g. smoking). Greater job-related health stress $z(t)$ accelerates the 'ageing' process ($\partial d/\partial z > 0$). The deterioration rate depends in a flexible way on health, instead of the usual assumption of a linear relationship $d(t) = \delta(t)H(t)$ as in Grossman (1972*a*,*b*) and the related literature (but see Dalgaard and Strulik, 2014, for an exception), and $\xi(t)$ denotes a vector of exogenous environmental conditions. Last, we have a fixed initial $H(0) = H_0$ and a fixed end condition $H(T) = H_{\min}$ for health.[9]

The second constraint is the total time budget Ω:

$$\Omega = \tau_w(t) + L(t) + \tau_m(t) + \tau_{C_h}(t) + \tau_{C_u}(t) + s[H(t)], \tag{6}$$

where the total time available in any period Ω is the sum of all possible uses $\tau_w(t)$ (work), $L(t)$ (leisure), $\tau_m(t)$ (health investment), $\tau_{C_h}(t)$ (healthy consumption), $\tau_{C_u}(t)$ (unhealthy consumption) and $s[H(t)]$ (sick time).

The third constraint concerns the dynamic relation for financial assets:

$$\frac{\partial A(t)}{\partial t} = rA(t) + Y(t) - p_{X_h}(t)X_h(t) - p_{X_u}(t)X_u(t) - p_m(t)m(t). \tag{7}$$

Assets $A(t)$ provide a return r (the return on capital), increase with income $Y(t)$ and decrease with purchases in the market of healthy consumption goods $X_h(t)$, unhealthy consumption goods $X_u(t)$ and medical care $m(t)$, at prices $p_{X_h}(t)$, $p_{X_u}(t)$ and $p_m(t)$, respectively. We have a fixed initial $A(0) = A_0$ and a fixed end condition $A(T) = A_T$ for wealth, and assume that individuals face no borrowing constraints.[10]

Income $Y(t) \equiv Y[H(t), z(t); E, x(t)]$ is assumed to be an increasing function of health $H(t)$ ($\partial Y/\partial H > 0$) and of job-related health stress $z(t)$ ($\partial Y/\partial z > 0$; Case and Deaton, 2005). We follow Grossman (1972*a*,*b*, 2000) and model income $Y(t)$ as the product of the wage rate $w(t)$ and time spent working $\tau_w(t)$:

$$Y[H(t)] \equiv w(t)\{\Omega - L(t) - \tau_m(t) - \tau_{C_h}(t) - \tau_{C_u}(t) - s[H(t)]\}. \tag{8}$$

[8] We follow Grossman (1972*a*,*b*) in distinguishing between the production of health $I[m(t), \tau_m(t); E]$ and the deterioration of health $d(t)$ and in allowing the deterioration rate $d(t)$ to be a function of health $H(t)$. We follow Case and Deaton (2005) in modelling health behaviours as operating through the deterioration rate $d(t)$. These choices are somewhat arbitrary. Mathematically, however, they are equivalent, with the exception that with our current choice, investment is not a direct function of health or health behaviour since the production process does not explicitly depend on them.

[9] In Grossman's original formulation (Grossman, 1972*a*,*b*) length of life T is determined by a minimum health level $H_{\min}$ below which life cannot be sustained. If health reaches this level $H(t) = H_{\min}$ an individual dies, hence $H(T) \equiv H_{\min}$.

[10] Imperfect capital markets itself could be a cause of socio-economic disparities in health if low income individuals face greater borrowing constraints and therefore cannot optimally invest in health.

Individuals receive wages $w(t) \equiv w[t, z(t); E, x(t)]$, which are a function of job-related health stress $z(t)$:

$$w(t) = w_*(t)[1 + z(t)]^{\gamma_w}, \tag{9}$$

where $\gamma_w \geq 0$ and $w_*(t) \equiv w_*[E, x(t)]$ represents the 'stressless' wage rate, i.e. the wage rate associated with the least job-related health stress $z(t) = 0$.[11] The stressless wage rate $w_*(t)$ is a function of the consumer's education E and experience $x(t)$ (Mincer, 1974):

$$w_*(t) = w_E e^{\rho_E E + \beta_x x(t) - \beta_{x^2} x(t)^2}, \tag{10}$$

where education E is expressed in years of schooling, $x(t)$ is years of working experience and ρ_E, β_x and β_{x^2} are coefficients, assumed to be positive.

Thus, we have the following optimal control problem: the objective function (1) is maximised with respect to the control functions $L(t)$, $X_h(t)$, $\tau_{C_h}(t)$, $X_u(t)$, $\tau_{C_u}(t)$, $m(t)$, $\tau_m(t)$, $z(t)$, the parameter T and subject to the constraints (4), (6) and (7).

The Hamiltonian (Seierstad and Sydsaeter, 1987) of this problem is as follows:

$$\Im = U(t)e^{-\beta t} + q_H(t)\frac{\partial H(t)}{\partial t} + q_A(t)\frac{\partial A(t)}{\partial t}, \tag{11}$$

where $q_H(t)$ is the marginal value of remaining life-time utility (from t onward) derived from additional health capital:

$$q_H(t) = \frac{\partial}{\partial H(t)}\int_t^{T^*} U(*)e^{-\beta s}\mathrm{d}s, \tag{12}$$

and $q_A(t)$ is the marginal value of remaining life-time utility derived from additional financial capital:

$$q_A(t) = \frac{\partial}{\partial A(t)}\int_t^{T^*} U(*)e^{-\beta s}\mathrm{d}s, \tag{13}$$

T^* denotes optimal length of life, and $U(*)$ denotes the maximised utility function (Caputo, 2005).

The condition for optimal longevity T follows from the dynamic envelope theorem (Caputo, 2005):

$$\frac{\partial}{\partial T}\int_t^{T^*} U(*)e^{-\beta s}\mathrm{d}s = \frac{\partial}{\partial T}\int_0^{T} \Im(t)\mathrm{d}t = \Im(T) = 0. \tag{14}$$

The marginal value of life extension is given by $\Im(T) = U(T)e^{-\beta T} + q_H(T)\partial H(t)/\partial t|_{t=T} + q_A(T)\partial A(t)/\partial t|_{t=T}$. When dividing by the marginal value of wealth, one obtains a measure for the monetary value of life, $\Im(T)/q_A(T) = U(T)/U_C(T) + q_{h/a}(T)\partial H(t)/\partial t|_{t=T} + \partial A(t)/\partial t|_{t=T}$. The monetary value of life is similar to the expressions obtained in Rosen (1988) (his equation (16)) and Murphy and Topel (2006) (their equations (7) and (8)), for health-neutral consumption. Our measure is richer since it additionally takes into account asset accumulation and health depreciation.

[11] We associate $z(t) = 0$ with the least amount of job-related health stress possible in employment, and since there is no obvious scale to job stress we employ the simple relationship shown in (9).

2.2. *First-order Conditions*

In this subsection, we discuss the first-order conditions for optimisation. We assume that an interior solution to the optimisation problem exists. Detailed derivations are provided in online Appendix A. The first-order condition for health investment is given by

$$q_{h/a}(t) = \pi_I(t), \tag{15}$$

where $q_{h/a}(t)$ represents the marginal benefit of health investment, defined as the ratio of the marginal value of health, $q_H(t)$, to the marginal value of wealth, $q_A(t)$ (throughout the article, we refer to $q_{h/a}(t)$ as the relative marginal value of health):

$$q_{h/a}(t) \equiv \frac{q_H(t)}{q_A(t)}, \tag{16}$$

and $\pi_I(t)$ represents the marginal cost of health investment:

$$\pi_I(t) \equiv \frac{p_m(t)}{\partial I/\partial m} = \frac{w(t)}{\partial I/\partial \tau_m}. \tag{17}$$

The marginal benefit of health investment increases in the marginal value of health $q_H(t)$ and decreases in the marginal value of wealth $q_A(t)$. If the marginal value of health is high individuals invest more in health, and if the marginal value of wealth is high individuals invest less, consume less and save more.

For the functional form (5), we can express the marginal cost of health investment $\pi_I(t)$ as follows (see online Appendix B for detail):

$$\pi_I(t) = \frac{p_m(t)^{1-\beta_I} w(t)^{\beta_I}}{\mu_I(t;E)\alpha_I^{1-\beta_I}\beta_I^{\beta_I}} m(t)^{1-\alpha_I-\beta_I}, \tag{18}$$

$$\pi_I(t) = \frac{p_m(t)^{\alpha_I} w(t)^{1-\alpha_I}}{\mu_I(t;E)\alpha_I^{\alpha_I}\beta_I^{1-\alpha_I}} \tau_m(t)^{1-\alpha_I-\beta_I}. \tag{19}$$

The marginal cost of health investment $\pi_I(t)$ increases with the level of health investment goods and services $m(t)$ and time inputs $\tau_m(t)$ due to decreasing returns to scale $0 < \alpha_I + \beta_I < 1$,[12] with the price of medical goods and services purchased in the market $p_m(t)$ and with the opportunity cost of time $w(t)$, where the latter is a function of job-related health stress, $z(t)$.

The first-order condition for leisure is

$$\frac{\partial U}{\partial L} = q_A(0) w(t) e^{(\beta - r)t}, \tag{20}$$

a standard result equating the marginal utility of leisure $\partial U/\partial L$ to the marginal cost of leisure, which is a function of the marginal value of initial wealth $q_A(0)$, the individual's

[12] Intuitively, at higher levels of investment, due to concavity of the health-production function $I[m(t), \tau_m(t); E]$, an additional increment of investment $m(t), \tau_m(t)$ produces a smaller improvement in health. Thus, the effective 'price' $\pi_I(t)$ increases with the level of investment $m(t), \tau_m(t)$.

wage rate $w(t)$, and the difference between the time preference rate β and the return on capital r.

The first-order condition for healthy consumption is

$$\frac{\partial U}{\partial C_h} = q_A(0)\left[\pi_{C_h}(t) - \varphi_{dC_h}(t)\right]e^{(\beta-r)t}, \tag{21}$$

where $\pi_{C_h}(t) \equiv \pi_{C_h}[t, C_h(t), z(t); E, x(t)]$ is the marginal monetary cost of healthy consumption $C_h(t)$:

$$\pi_{C_h}(t) \equiv \frac{p_{X_h}(t)}{\partial C_h/\partial X_h} = \frac{w(t)}{\partial C_h/\partial \tau_{C_h}}, \tag{22}$$

and $\varphi_{dC_h}(t) \equiv \varphi_{dC_h}[t, H(t), C_h(t), C_u(t), z(t); E, x(t), \xi(t)]$ is the marginal health benefit of healthy consumption:

$$\varphi_{dC_h}(t) \equiv -q_{h/a}(t)\frac{\partial d}{\partial C_h}. \tag{23}$$

The marginal monetary cost of healthy consumption $\pi_{C_h}(t)$ (see (22)) is a function of the price of healthy consumption goods and services $p_{X_h}(t)$ and the opportunity cost of time $w(t)$, and represents the *direct* monetary cost of consumption. The marginal health benefit of healthy consumption $\varphi_{dC_h}(t)$ (see (23)), is equal to the product of the relative marginal value of health $q_{h/a}(t)$ and the 'amount' of health saved $\partial d/\partial C_h$, and represents the marginal value of health saved.

Similarly, the first-order condition for unhealthy consumption is

$$\frac{\partial U}{\partial C_u} = q_A(0)\left[\pi_{C_u}(t) + \pi_{dC_u}(t)\right]e^{(\beta-r)t}, \tag{24}$$

where $\pi_{C_u}(t) \equiv \pi_{C_u}[t, C_u(t), z(t); E, x(t)]$ is the marginal monetary cost of unhealthy consumption $C_u(t)$ (direct monetary cost):

$$\pi_{C_u}(t) \equiv \frac{p_{X_u}(t)}{\partial C_u/\partial X_u} = \frac{w(t)}{\partial C_u/\partial \tau_{C_u}}, \tag{25}$$

and $\pi_{dC_u}(t) \equiv \pi_{dC_u}[t, H(t), C_h(t), C_u(t), z(t); E, x(t), \xi(t)]$ is the marginal health cost of unhealthy consumption (marginal value of health lost):

$$\pi_{dC_u}(t) \equiv q_{h/a}(t)\frac{\partial d}{\partial C_u}. \tag{26}$$

The first-order condition for unhealthy consumption (24) is similar to the condition for healthy consumption (21). The difference lies in the marginal health cost (rather than health benefit) of unhealthy consumption, which has to be added rather than subtracted from the marginal monetary cost of unhealthy consumption $\pi_{C_u}(t)$.

Last, the first-order condition for job-related health stress is

$$\varphi_z(t) = \pi_{dz}(t), \tag{27}$$

where $\varphi_z(t) \equiv \varphi_z[t, H(t), z(t); E, x(t)]$ is the marginal production benefit of job-related health stress:

$$\varphi_z(t) \equiv \frac{\partial Y}{\partial z}, \tag{28}$$

reflecting the notion that job-related health stress is associated with a compensating wage differential (greater earnings), and $\pi_{dz}(t) \equiv \pi_{dz}[t, H(t), C_h(t), C_u(t), z(t); E, x(t), \xi(t)]$ is the marginal health cost of job-related health stress (marginal value of health lost):

$$\pi_{dz}(t) \equiv q_{h/a}(t)\frac{\partial d}{\partial z}. \tag{29}$$

2.3. *The Health Cost and Health Benefit of Health Behaviour*

From the first-order conditions follow that lifestyle decisions regarding consumption and occupation provide utility (directly or indirectly), and are associated with a monetary cost and with an opportunity cost. However, in contrast to conventional economic models, in our theory, these lifestyle decisions are additionally associated with a 'health benefit' or a 'health cost'. The health cost (benefit) is given by

$$q_{h/a}(t)\frac{\partial d}{\partial x}, \tag{30}$$

where the variable x represents the relevant health behaviour (e.g. unhealthy consumption or hard physical labour), and the definition of $q_{h/a}(t)$ is given by (12), (13), and (16). From (16), (43) and (44) follows:

$$\begin{aligned} q_{h/a}(t) = {} & q_{h/a}(T)e^{-\int_t^T\left[r+\frac{\partial d}{\partial H}\right]\mathrm{d}x} \\ & + \int_t^T e^{-\int_t^s\left[r+\frac{\partial d}{\partial H}\right]\mathrm{d}x}\left(\frac{1}{q_A(0)}\frac{\partial U}{\partial H}e^{-(\beta-r)s} + \frac{\partial Y}{\partial H}\right)\mathrm{d}s. \end{aligned} \tag{31}$$

The health cost (benefit) is the amount of health 'lost' ('saved') due to the health behaviour $\partial d/\partial x$ times the relative marginal value of health $q_{h/a}(t)$ (the 'price' or 'value' of health, measured in life-time monetary units (e.g. US dollars) per unit of health, see (31)). In simple terms, unhealthy consumption worsens health, and the health cost is the value one attaches to the lifetime consequences of reduced health. These life-time consequences consist of the discounted additional (marginal) utility ($q_A(0)^{-1}\partial U/\partial H e^{-(\beta-r)s}$ in (31)) plus the additional earnings (production benefit) derived from better health ($\partial Y/\partial H$ in (31)), over the remainder of life (from t till T). They are discounted at the rate $r + \partial d/\partial H$, where $\partial d/\partial H$ represents the cost of holding the stock of health and r represents the opportunity cost (i.e. one could alternatively invest in assets).

There are assymetries between health costs and health benefits. For healthy consumption, the time investments and monetary costs are incurred today (and so is utility, if any, from its consumption), while the health benefits are reaped in the future. For unhealthy consumption, the opposite holds: it provides pleasure (and requires time investments and monetary costs) today, while the health costs occur in the future. This implies that discount rates matter, as well as information on the healthiness or unhealthiness of the good $\partial d/\partial x$. Individuals who discount the future

heavily (large β in (31)) or who underestimate the health consequences, will engage more in unhealthy consumption. Likewise, individuals who discount the future or underestimate the health gains from healthy consumption will engage less in healthy consumption.

Assessments of the value of a statistical life (VSL) (see Viscusi and Aldy, 2003, for a review) generally involve investigating the risk of death that people are willing to take (usually in a setting of hazardous work) and how much they should be paid for taking these risks. An analogous concept is captured in, e.g. the first-order condition for job-related health stress (27), which weighs the wage premium of engaging in unhealthy/risky jobs today with the costs in terms of the lifetime consequences of reduced health (and associated longevity). In our theory, individuals are also willing to engage in a certain amount of unhealthy consumption for the utility it provides, as long as this benefit outweighs the associated health cost: the reduction in life-time utility due to health loss associated with unhealthy consumption. Thus, our theory provides a framework for determining the value of life and the value of health in settings outside of hazardous work, e.g. by exploiting unhealthy behaviours.

As we discuss in the remainder of the article, the health cost and health benefit of behaviour are promising concepts for understanding a wide range of health behaviours, as well as the socio-economic disparities in these behaviours.

2.4. *Assumptions*

Apart from the earlier mentioned assumptions of diminishing returns to scale (DRTS) of investment $m(t)$, $\tau_m(t)$, in the health-production function $I[m(t), \tau_m(t); E]$ $(0 < \alpha_I + \beta_I < 1)$, and diminishing marginal utilities of healthy $C_h(t)$ and unhealthy consumption $C_u(t)$, of leisure $L(t)$ and of health $H(t)$, we further assume:

1. Diminishing marginal production benefit of health $\partial^2 Y/\partial H^2 < 0$, diminishing marginal production benefit of job-related health stress $\partial^2 Y/\partial z^2 < 0$, diminishing marginal health benefit of healthy consumption $\partial^2 d/\partial C_h^2 > 0$, and diminishing returns to longevity (see (14)):

$$\frac{\partial \Im(T)}{\partial T} = \frac{\partial^2}{\partial T^2}\int_t^{T^*} U(*)e^{-\beta s}\mathrm{d}s < 0, \tag{32}$$

   where T^* denotes optimal length of life and $U(*)$ denotes the maximised utility function (Caputo, 2005).
2. Increasing or constant returns to scale in the marginal health cost of unhealthy consumption $\partial^2 d/\partial C_u^2 \geq 0$ (Forster, 2001)[13] and in the marginal health cost of job-related health stress $\partial^2 d/\partial z^2 \geq 0$.
3. Cobb–Douglas CRTS relations between the inputs (goods/services purchased in the market and own-time) and the outputs healthy consumption $C_h(t)$,

[13] While it seems plausible that the health benefits of health investment and healthy consumption exhibit diminishing returns to scale, the health costs of unhealthy consumption and job-related health stress plausibly exhibit increasing returns to scale. In simple terms: whereas after a certain point more health investment, exercise or consumption of healthy foods, does not prevent eventual ageing, escalating risky behaviour (e.g. illicit drug use) or dangerous work can lead to rapid health deterioration.

and unhealthy consumption $C_u(t)$: $C_h(t) = \mu_{C_h}(t) X_h(t)^{\kappa_{C_h}} \tau_{C_h}(t)^{1-\kappa_{C_h}}$ and $C_u(t) = \mu_{C_u}(t) X_u(t)^{\kappa_{C_u}} \tau_{C_u}(t)^{1-\kappa_{C_u}}$, where κ_{C_h}, κ_{C_u} are the elasticities of the outputs with respect to goods and services and $1-\kappa_{C_h}$, $1-\kappa_{C_u}$ are the elasticities of the outputs with respect to time inputs. As a result, we have (see (22) and (25)):

$$\pi_{C_h}(t) = \frac{p_{C_h}(t)^{\kappa_{C_h}} w(t)^{1-\kappa_{C_h}}}{\kappa_{C_h}^{\kappa_{C_h}} (1-\kappa_{C_h})^{1-\kappa_{C_h}} \mu_{C_h}(t)}, \tag{33}$$

$$\pi_{C_u}(t) = \frac{p_{C_u}(t)^{\kappa_{C_u}} w(t)^{1-\kappa_{C_u}}}{\kappa_{C_u}^{\kappa_{C_u}} (1-\kappa_{C_u})^{1-\kappa_{C_u}} \mu_{C_u}(t)}. \tag{34}$$

4. First-order direct effects dominate higher-order indirect effects for control variables. For example, wealth affects healthy consumption directly, but also indirectly through its effect on unhealthy consumption, since unhealthy consumption affects healthy consumption through its effect on utility and on health deterioration. The assumption, in this particular example, is that the direct effect of wealth on healthy consumption dominates any indirect wealth effect that operates through unhealthy consumption or through any other control variable.

   Second and higher-order terms are often ignored since they tend to be quantitatively less important (e.g. first-order Taylor series approximations). Moreover, additively separable utility functions are fairly conventional in the literature (Zeckhauser, 1970; Gjerde *et al.*, 2005; Hall and Jones, 2007) and assuming that utility is additively separable and that the health deterioration rate is additively separable, would lead to similar results. The difference is that we require second-order terms to be small, not that they are identical to zero and we do not make the assumption for state variables since small differences can grow large over time. Thus, e.g. the marginal utility of healthy consumption is still a function of health (Finkelstein *et al.*, 2013). Therefore, our assumption is less restrictive than the fairly conventional assumption of functions being additively separable in their arguments. We provide more detail in online Appendix D.

In addition to these four assumptions, we need further information to be able to conduct the comparative dynamic analyses (Section 4). In particular, we need the sign of $\partial d/\partial H$, $\partial q_A(t)/\partial A(t)$ and $\partial q_A(t)/\partial H(t)$. For this reason, in the next Section, we calibrate a simple version of the model to establish which signs are plausible. We obtain these as our 'benchmark' relationships. From the calibrated model, we have:

(*1*) $\partial d/\partial H > 0$;
(*2*) $\partial q_A(t)/\partial A(t) < 0$; $\partial q_A(t)/\partial w_E < 0$; $\partial q_A(t)/\partial E < 0$; and
(*3*) $\partial q_A(t)/\partial H(t) < 0$.

The calibrated model suggests that health increases aging at a diminishing rate, i.e. $\partial d/\partial H \geq 0$ and $\partial^2 d/\partial H^2 \leq 0$. This suggests that the health of healthy individuals deteriorates faster in absolute terms (since they have more of it) but not in relative terms (as a percentage of total health). Further, the calibrated model suggests diminishing returns to wealth $\partial q_A(t)/\partial A(t) = \partial^2[\int_t^{T^*} U(*) e^{-\beta s} \mathrm{d}s]/\partial A(t)^2 < 0$,

i.e. poorer individuals derive greater benefits from an additional increment of wealth than wealthier individuals. Finally, the calibrated model suggests unhealthier individuals derive greater benefits from an additional increment of wealth $(\partial q_A(t)/\partial H(t) = \partial^2[\int_t^{T^*} U(*)e^{-\beta s}\mathrm{d}s]/\partial A(t)\partial H(t) < 0)$.[14] This suggests health and wealth are to some extent substitutable in financing consumption and leisure (Muurinen, 1982; Case and Deaton, 2005).[15]

In online Appendix C, we also explicitly investigate the sensitivity of our predictions to the opposite signs of the above relations to assess the robustness of our results.

3. Dynamics and Numerical Calibration

In this Section, we begin with a study of the life-cycle profiles of health and health investment. Theoretically, the model's predictions for these life-cycle patterns are ambiguous. We therefore calibrate a simple version of the model to investigate whether for a realistic set of parameter values it is able to reproduce empirical stylised facts. We then use findings from the calibrated simulations to make predictions for the life-cycle trajectories of healthy and unhealthy consumption and job-related health stress.

3.1. *Health Investment and Health over the Life Cycle*

A quick glance at the first-order conditions in subsection 2.2 shows that the relative marginal value of health $q_{h/a}(t)$ is an important driver of health investment (and more generally of health behaviour). Specifically, note that from (15), (18) and (19) we obtain

$$m(t) \propto q_{h/a}(t)^{1/(1-\alpha_I-\beta_I)}, \tag{35}$$

$$\tau_m(t) \propto q_{h/a}(t)^{1/(1-\alpha_I-\beta_I)}. \tag{36}$$

Thus, health investment increases with the marginal value of health. Hence, to understand health investment we first investigate the dynamics of the relative marginal value of health $q_{h/a}(t)$ and of health $H(t)$. We then discuss the implications of these dynamic patterns for health behaviours.[16]

Individuals start life generally in good health at $H(0) = H_0$, while the terminal health stock $H(T)$ is constrained to the minimum health level $H_{\min}$, below which life cannot be sustained. This implies that health decreases over the life cycle. Indeed, empirical evidence suggests that health falls over the life cycle, first slowly and then more rapidly in old age (Deaton and Paxson, 1998; Van Kippersluis *et al.*, 2009; Dalgaard and Strulik, 2014).

[14] Where T^* denotes optimal length of life and $U(*)$ denotes the maximised utility function (see (13)).

[15] Intuitively, health is a resource and having more of it relaxes the dynamic constraint for wealth: being in better health reduces the need for health investment and health provides earnings. Both health and wealth are resources that enable consumption and leisure.

[16] After solving the optimal control problem, the model's solutions can be fully expressed in the state and co-state functions. For this reason, it is useful to start with an analysis of the dynamics of the co-state function $q_{h/a}(t)$ and the state function $H(t)$.

The evolution of the relative marginal value of health $q_{h/a}(t)$ is given by

$$\frac{\partial q_{h/a}(t)}{\partial t} = q_{h/a}(t)\left[r + \frac{\partial d}{\partial H}\right] - \frac{1}{q_A(0)}\frac{\partial U}{\partial H}e^{-(\beta - r)t} - \frac{\partial Y}{\partial H} \tag{37}$$

(combine equations (43) and (44) of online Appendix A). Recall that the relative marginal value of health $q_{h/a}(t)$ is the ratio of the marginal value of health $q_H(t)$ and the marginal value of wealth $q_A(t)$. The marginal value of wealth depreciates with age at the rate of return on capital r (see (43)). The marginal value of health $q_H(t)$, however, may increase or decrease with age. As long as $q_H(t)$ appreciates, or depreciates more slowly than $q_A(t)$, the relative marginal value of health $q_{h/a}(t)$ will increase with age.

Empirical evidence suggests that health investments increase with age: medical expenditures peak in the final phase of life (Zweifel *et al.*, 1999),[17] and other components of health investment either increase or stay relatively flat with age (Podor and Halliday, 2012). This suggests that the relative marginal value of health increases with age (see (15)).

Theoretically, our model allows for these patterns but does not unambiguously predict them. In the next subsection, therefore, we resort to calibrated simulations to investigate whether our model, for a realistic set of parameter values, is able to reproduce these empirical stylised facts.

3.2. *Calibrated Simulations*

Our calibrated simulations proceed in several steps. First, we compute five-year averages for a health index and for health investment and treat these as the data 'moments' for our calibration. Following Poterba *et al.* (2011, 2013), we construct a health index using 1999–2013 US Panel Study of Income Dynamics (PSID) data of a rich set of health measures, including:

(*i*) a binary indicator of poor self-reported health;
(*ii*) diseases such as acute myocardial infarctions (AMI), arthritis, asthma, cancer, diabetes, heart conditions, high blood pressure, learning disorders, lung disease and stroke;
(*iii*) mental health problems;
(*iv*) activities of daily living (ADLs); and
(*v*) body-mass index (BMI).

We arbitrarily scale this health index such that initial health is equal to 100 and health at age 80 is equal to 15. We obtained age profiles for health investment from Halliday *et al.* (2017) who computed health investment expenditures from the US Medical Expenditure Survey (MEPS). Figure 1 shows the life-cycle patterns of health and health investment, where the dots indicate the average of the health index and of health investment in five year age groups. As the figure shows, health

[17] DeNardi *et al.* (2016) decompose medical expenditures into different components, and show that the rapidly increasing expenditures in old age can largely be attributed to long-term and nursing home care.

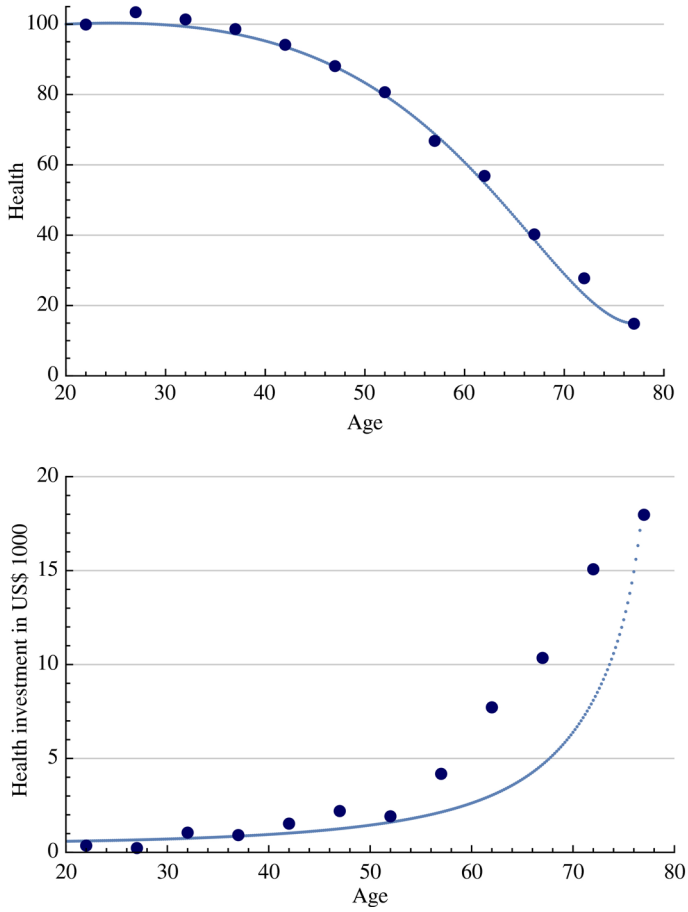

Fig. 1. *Model Simulation of Health (Top Part) and Health Investment (Bottom Part), Using Calibrated Parameter Values in Table 1*

Notes. The solid lines indicate the model simulations and the dots indicate the empirical data moments constructed from PSID (health) and MEPS (health investment). Colour figure can be viewed at https://academic.oup.com/ej.

declines in a concave manner over the life cycle, while health investment increases rapidly in old age. These are stylised facts our model of health production ought to reproduce.

Second, we estimate the hourly wage profile using the PSID data and a Mincer equation of log hourly wages on dummies for each educational level and a quadratic polynomial in age. The estimated parameters are as follows:

$$w_t = e^{2.3366+0.3731\times D+0.0314\times t-0.0006\times t^2}, \tag{38}$$

where D is a dummy for a college degree, t measures age and we multiply by two to obtain the yearly wages in thousands of dollars (assuming individuals work for 40 hours,

50 weeks per year). This Mincer-type wage equation for a college graduate starts with annual wages of \$30,000 per year at age $t = 20$ and peaks at \$45,000 per year at age $t = 50$ after which wages gradually decline.

Third, we numerically simulate a simplified version of our model (see online Appendix C for details). Since the aim is to investigate whether our model is capable of reproducing the stylised facts regarding the life-cycle profiles of health and health investment, we omit the time inputs into consumption and health investment, omit the distinction between healthy and unhealthy consumption, and omit job-related health stress.

We assume a constant relative risk aversion (CRRA) utility function with scale parameter μ_U of the form:

$$U[H_t, C_t] = \frac{1}{1-\rho}\mu_U \left[C_t^{\zeta} H_t^{1-\zeta}\right]^{1-\rho} + B, \tag{39}$$

where $0 \leq \zeta \leq 1$ is the relative 'share' of consumption *versus* health, ρ is the coefficient of relative risk aversion and B is a constant to ensure that utility is positive (Hall and Jones, 2007). We follow Scholz and Seshadri (2016) in modelling sick time as $s[H_t] = H_t^{-\gamma}$, and use a flexible functional form for health production:

$$H_{t+1} - H_t = \mu_I m_t^{\alpha} - d_t^* H_t^{\nu}, \tag{40}$$

where $0 < \alpha < 1$, and $d_t^* = e^{a+bt}$ following Cropper (1981) and Wagstaff (1986*a*). This functional form is flexible enough to capture various possible relationships between the deterioration rate ($d_t = d_t^* H_t^{\nu}$) and health. For example, $\nu = 1$ represents the Grossman case, while $\nu < 0$ is akin to the health-deficit model by Dalgaard and Strulik (2014), in which the health of unhealthy individuals deteriorates faster. The functional form also depends in a flexible way on calendar age t.

No consensus exists over what causes ageing with many theories attempting to explain what drives it. Two important classes of theories are prominent. The first class consists of 'deficit-based theories', where ageing occurs predominantly as a result of the accumulation of deficits (or 'damage'). For example, Arking (2006) advocates a theory of ageing in which health decline is independent of calendar age, depending solely on the level of health, i.e. as in a model of ageing due to wear and tear (Dalgaard and Strulik, 2014). The second class consists of 'programmed theories of ageing', where ageing is predominantly the result of genetically regulated or predetermined processes (De Magalhaes, 2003, 2011, 2014; Longo *et al.*, 2005). Such processes exhibit 'pure' calendar age effects. De Magalhaes (2011) writes for example: '… it is now largely recognised that ageing (…) is not primarily the result of damage to irreplaceable body parts (p. 34)' and that there is a role for '… genetically regulated processes or predetermined mechanisms (p. 38)' (De Magalhaes, 2011). Further, Longo *et al.* (2005) provide evidence for genetic programming 'that regulates the level of protection against stochastic damage, and therefore the length of time an organism remains healthy' (p. 866). Finally, the so-called ageing-clock theory, which also falls within the programmed theories, stipulates that 'Aging is programmed into our bodies, like a clock ticking away from the moment of conception (p. 61)' (Moody and Sasser, 2014). The two classes of theories are not mutually exclusive; and there is in fact wide recognition that ageing exhibits genetic programming or predetermined, and accumulated deficits aspects, with disagreement

about their relative importance. We therefore choose to remain agnostic about the ageing process, using a specification $d_t = d_t^* H_t^{\nu}$ that is flexible enough to encompass both deficits as well as calendar age aspects of ageing. In essence, we ask the data which aspects of ageing are important in our model for a good fit. We find that both calendar age and deficits features are needed.

Fourth, we fix a large number of model parameters by taking values from the literature (see Table 1 in online Appendix C). In particular, we set $\alpha = 0.75$ in line with Hugonnier *et al.* (2013, 2016) who use values of 0.69 and 0.77, respectively. We follow Scholz and Seshadri (2016), setting $A_0 = A_T = 0$, $\rho = 3$, $\zeta = 0.7$ and $\gamma = 1.7$.[18] We follow Blau (2008) and DeNardi *et al.* (2016), setting the time preference rate $\beta = 0.03$, and the interest rate r equal to β. We further normalise prices of consumption and health investment and the efficiency of consumption to one. The remaining parameters to be calibrated are ν, μ_I, μ_U, and the parameters a and b of the deterioration rate d_t^*.

Given the data moments and our model solutions, we calibrate the parameters ν, μ_I, μ_U, a and b using a method of simulated moments (MSM) approach (see online Appendix C for details). Admittedly, this parameter space leaves many degrees of freedom. We do not claim our parametrisation is unique. See, for example, Dalgaard and Strulik (2014) for an alternative framework using health deficits that is also able to reproduce the life-cycle profiles for health and health investment. Our aim is more modest: to illustrate that for a realistic set of parameter values the model can reproduce the empirical stylised facts. Table 1 in online Appendix C provides the fixed and calibrated parameter values, and Figure 1 shows the resulting simulated profiles (solid lines) alongside the empirical moments (dots). For this realistic set of parameter values, the fit of our simulated profiles is fairly accurate. The model can generate empirically plausible trajectories.[19]

The calibrated simulations provide three additional pieces of information. First, the calibrated simulations suggest an optimal length of life of $T = 76.8$ years.

Second, our calibration establishes a benchmark for the sign of relationships we require to obtain comparative dynamic results. The parameter value $\nu = 0.3$ in (40) is consistent with $\partial d/\partial H \geq 0$ and $\partial^2 d/\partial H^2 \leq 0$. Further, in counterfactual simulations where we increase A_0 and H_0 by a small amount, we obtain $\partial q_A(0)/\partial A_0 < 0$ and $\partial q_A(0)/\partial H_0 < 0$ (see Table 2 in online Appendix C).

Third, the calibrated model demonstrates that while health declines, the relative marginal value of health increases, with age. We will use these life-cycle patterns to derive from theory the implied life-cycle profiles for healthy and unhealthy consumption and job-related health stress in the general model.

[18] Scholz and Seshadri (2016) use $\gamma = 0.17$, we multiply with 10 to account for the difference in scaling between their and our health variable.

[19] The calibration suggests that a model where health depreciation depends only implicitly on age through the health stock $H(t)$, i.e. $d_t^* H_t^{\nu} = d_0 H_t^{\nu}$ and $\nu < 0$ is able to fit the health profile very accurately. However, this model cannot reproduce the quasi-exponentially increasing profile for health investment. Our understanding from the calibration exercise is that there are two different ways in which one can reproduce the empirical life-cycle profiles of health and health investment:

(*i*) a Grossman-style health accumulation model with health depreciation that depends explicitly on age as well as implicitly on the health stock, e.g., $d_t^* H_t^{\nu}$, with $0 < \nu < 1$; and
(*ii*) a health-deficit model in the style of Dalgaard and Strulik (2014).

3.3. *Health Behaviour over the Life-cycle*

As illustrated by the calibrated simulations, for plausible parameter values, the relative marginal value of health $q_{h/a}(t)$ increases with age. This would suggest that the health cost of unhealthy consumption and of job-related health stress and the health benefit of healthy consumption increase with age.[20] Smoking rates are 8.9% among the 65+ compared to 21.6% among the 25–44 (US DHHS, 2014), and intake of fruit and vegetables increases with age (Serdula *et al.*, 2004; Pearson *et al.*, 2005). These patterns are consistent with the notion that the health benefit of healthy behaviour and the health cost of unhealthy behaviour increase with age: individuals start caring more about their health when they get older.

With declining health $H(t)$ and an increasing marginal value of health $q_{h/a}(t)$ with age, we obtain the following life-cycle patterns for health behaviours. Early in life, individuals are generally healthy and, therefore, plausibly value health less (see Table 2 in online Appendix C and Propositions 5 and 7, subsection 4.3). As a result, they invest less in their health (see (15)), engage more in unhealthy consumption (see (24)), and less in healthy consumption (see (21)). As individuals age, declining health becomes a burden as poor health reduces utility and increased sick time reduces earnings. As a result, the benefits of health increase and individuals invest more in health, shift toward healthier consumption and reduce the level of job-related health stress. This general trend of improved health behaviour may be partially reversed in mid-life, as wages peak, leading to a higher opportunity cost of time. This may result in a reduction in health investment and healthy consumption in mid-life, relative to a general trend of improved health behaviours with age.

The pattern for job stress is distinct. Early in life the health cost of job stress $q_{h/a}(t)\partial d/\partial z]$ is low, but so is the marginal benefit of job stress $\partial Y/\partial z$ (see (27)). As wages increase, the benefit of job-stress increases. After mid-age, declining health reduces the marginal benefit of job-stress, as sick time reduces the time available for work and wages plateau or even decrease. This suggests a pattern in which job stress initially increases as the marginal benefit of job-stress increases due to wage growth, followed by a decline in job stress due to plateauing or declining wages, increasing sick time and an increasing health cost of job stress with age.

Prediction 1. *Individuals in mid-life plausibly accept unhealthy working conditions as they value the associated wage premium, but as they age they seek to engage in healthier work.*

4. Variation in Health Behaviour by SES and Health

Comparative dynamic analyses allow exploration of the effect of SES and health on the life-cycle trajectories of the control and state variables. We investigate the change in the optimal trajectory in response to variation in initial conditions or

[20] However, it is unclear whether $\partial d/\partial x$, for $x = C_h, C_u, z$ increases or decreases with age. Since health decreases with age, and since $\partial d/\partial H > 0$ (since $\nu > 0$ in our calibrated simulations), a declining health stock decreases the ageing rate $d(t)$ with age. Depending on the sign of $\partial^2 d/\partial H\partial C_h$, $\partial^2 d/\partial H\partial C_u$, and $\partial^2 d/\partial H\partial z$, this effect may be either reinforced or mitigated. As outlined in Assumption 4, we assume these second-order effects are quantitatively less important.

other model parameters, by comparing the 'perturbed' optimal trajectory with respect to the 'unperturbed' (or original) trajectory. Comparative dynamic analysis is an alternative to numerical simulation that has the benefit of being less restrictive by allowing for quite general functional forms (i.e. that utility be concave rather than, e.g. CRRA). Moreover, as Dalgaard and Strulik (2015, p. 2) put it: 'A purely numerical analysis entails a danger of neglecting general properties and implications of the model'.

Our emphasis is on exploring differences in constraints related to SES and health.[21] Common measures of SES employed in empirical research are wealth, earnings (income) and education. Here, we provide an intuitive discussion of the comparative dynamic results, supported by a number of propositions, whose formal proof we relegate to online Appendix D.

The effect of variation in an initial condition or other model parameter δZ, where we are particulary interested in $Z = \{A_0, w_E, E, H_0\}$, on any control or state variable $g(t)$ can be separated into two components:[22]

$$\frac{\partial g(t)}{\partial Z} = \frac{\partial g(t)}{\partial Z}\bigg|_T + \frac{\partial g(t)}{\partial T}\bigg|_Z \frac{\partial T}{\partial Z}, \tag{41}$$

where the first term on the right-hand side (RHS) represents the response to variation in Z for constant T and the second term on the RHS represents the additional response due to the associated variation in T.

4.1. *Variation in Initial Wealth, δA_0*

Let us focus first on the hypothetical case where length of life T is fixed. Contrasting the fixed T case with the general case where T is free provides us with useful insights regarding the properties of the model. This scenario may represent a developing nation with a high disease burden (where there may be lack of access to medical or public health technology, and competing risks from many diseases), the developed world, if it were faced with diminishing ability to further extend life or individuals with a disease that severely limits longevity, such as Huntington's disease (Oster *et al.*, 2013).

Proposition 1. *Absent ability to extend life, wealthy individuals, ceteris paribus, value health only marginally more than less wealthy individuals.*

For proof, see online Appendix D.1.

For fixed length of life T, additional wealth δA_0 increases the relative marginal value of health initially, $\partial q_{h/a}(t)/\partial A_0|_T > 0$ but eventually the relative marginal value of wealth decreases with respect to the unperturbed path, $\partial q_{h/a}(t)/\partial A_0|_T < 0$. This result is illustrated in Figure 2, where the thick solid line labelled 'Unperturbed' represents

[21] Part of the SES-health gradient may be explained by differences in individual's preferences. A lower rate of time preference β operates in a similar manner to wealth, earnings and education. A lower rate of time preference may also lead to greater investment in education (not part of our theory), and hence lead to joint determination of health and education (Fuchs, 1986).

[22] Note that we can restart the problem at any time t, taking $A(t)$ and $H(t)$ as the new initial conditions. Thus the comparative dynamic results derived for, e.g. variation in initial wealth δA_0 and initial health δH_0 have greater validity, applying to variation in wealth $\delta A(t)$ and in health $\delta H(t)$ at any time $t \in [0, T)$.

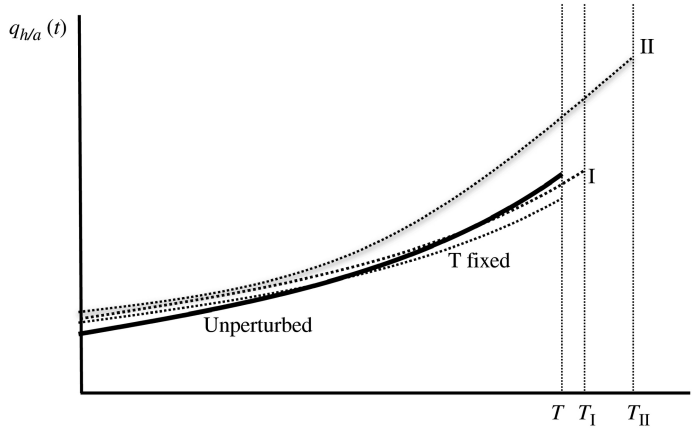

Fig. 2. *Evolution of the Relative Marginal Value of Health $q_{h/a}(t)$ with Age Due to Variation in A_0*
Notes. The solid thick line, labelled 'Unperturbed', represents the unperturbed path. The perturbed paths are shown for the T fixed case (dotted line, labelled 'T fixed'), scenario *I*, associated with small life extension T_{I} (dotted line, labelled 'I') and scenario *II*, associated with large life extension T_{II} (dotted line labelled 'II').

the unperturbed trajectory of the relative marginal value of health $q_{h/a}(t)$ *versus* age t and the dotted line labelled 'T fixed' represents the perturbed path for fixed T. Note that both curves end at $t = T$. The intuition is that the relative marginal value of health cannot be higher at all times, as this would be associated with higher health investment, improved health behaviour (see the first-order conditions in subsection 2.2), and a longer life, violating the condition that end of life occurs at $t = T$ (fixed) at the minimum health level $H(T) = H_{\min}$. Indeed, simulations confirm this pattern: in the calibrated model, for individuals with higher initial wealth, the marginal value of health is higher up to age 70 and lower thereafter.[23] The response to additional wealth is thus muted as the individual is forced to invest less later in life in order not to extend life. Hence, the first term on the RHS of (41) is small for variation in wealth A_0.

Let us now turn to the more interesting case where individuals can optimally choose T; they not only invest in the quality of life but also in the quantity of life.

Individuals optimally choose longevity T such that the marginal value of life extension is zero at this age, $\Im(T) = 0$ (see (14)):

$$\frac{\Im(T)}{q_A(T)} = \frac{1}{q_A(0)} U(T) e^{-(\beta - r)T} + q_{h/a}(T) \left.\frac{\partial H}{\partial t}\right|_{t=T} + \left.\frac{\partial A}{\partial t}\right|_{t=T} = 0, \tag{42}$$

where for reasons of exposition we have divided by $q_A(T)$. As (42) shows, the marginal benefit of extending life consists of the additional utility from consumption and health.

[23] Note that it can be understood that the opposite pattern, one of disinvestment early in life and increased investment later in life, is inferior as it would be associated with lower health at all ages and therefore a reduced consumption benefit.

The marginal costs consist of the increasingly binding wealth and health constraints, due to declining wealth and declining health near the end of life.[24] With declining health, the utility $U(t)$ gained over the extra years decreases, and thereby the marginal benefit of life extension. The calibrated simulation shows that $q_{h/a}(t)$ increases rapidly with age as health declines, suggesting that the marginal cost of life extension does, too. Eventually, the depreciation of health and decumulation of assets render additional life extension non-optimal.

As we will see, ability to extend life changes the picture dramatically as life extension increases the return to health investment by increasing the period over which a multitude of benefits of health can be accrued. The results can be summarised by Propositions 2 and 3.

PROPOSITION 2. *Wealthy individuals live longer:* $\partial T/\partial A_0 > 0$.

For proof, see online Appendix D.2.

PROPOSITION 3. *Wealthy individuals value health more and are healthier at all ages. The more life can be extended, the stronger is the increase in the value of health in response to additional wealth.*

For proof, see online Appendix D.3.

Intuitively, at high values of wealth (and hence consumption), individuals prefer investing in health over consuming, since health extends life, the period over which they can enjoy the benefits of health, leisure and consumption, whereas additional consumption per period would yield only limited marginal utility due to diminishing utility of consumption (Becker, 2007; Hall and Jones, 2007). With sufficient wealth, one starts caring more about other goods, in particular health. Wealthier individuals value health more relative to wealth, invest more, are healthier and live longer (Propositions 2 and 3).

The second part of Proposition 3 is best understood by its visual representation in Figure 2. The extent to which individuals are able to extend life, $\partial T/\partial A_0$, depends on the model's parameters r, α, μ_I, etc. These parameters are in turn determined by biology, medical technology and environmental factors. If these factors are unfavourable to life extension (scenario *I*; small life extension), then individuals value health more early in life, but value health less later in life (the perturbed path starts higher, but eventually crosses the unperturbed path). This is the case we observe in our calibrated simulations. The pattern of initially higher investment, and subsequently lower investment, closely resembles that of the fixed T case (see Proposition 1).

In contrast, if additional wealth affords considerable life extension (scenario *II*), the relative marginal value of health is higher at all times. Life extension raises the return to health investment and healthy behaviours. Further, utility from leisure and consumption can be enjoyed with additional years of life. Health also generates additional wealth from work, reinforcing the effect of the initial endowment of wealth δA_0. Together, these various benefits substantially raise the value of health, leading to

[24] Both $\partial H(t)/\partial t|_{t=T}$ and $\partial A(t)/\partial t|_{t=T}$ are negative since health approaches $H_{\min}$ from above, and assets decline near the end of life as poor health reduces earnings.

improved health behaviours, better health throughout life and greater longevity. This leads to the following prediction.

PREDICTION 2. *Health disparities are larger in environments where the wealthy can effectively use their resources to extend life.*

Propositions 2 and 3 also allow gauging the predicted response of an increase in wealth on health behaviours. A higher relative marginal value of health directly increases the health benefit of healthy behaviour and the health cost of unhealthy behaviour. Plausibly, this represents the dominant effect,[25] consistent with wealthy individuals behaving healthier (Cutler and Lleras-Muney, 2010; Cutler *et al.*, 2011; Cawley and Ruhm, 2012) and consistent with less affluent individuals responding more strongly to an unanticipated wealth shock (Van Kippersluis and Galama, 2014).

PREDICTION 3. *Wealthy individuals shift consumption towards healthy consumption: they consume more healthy and moderately unhealthy consumption goods and services but fewer severely unhealthy consumption goods and services.*

The comparative dynamic effect of wealth on healthy consumption can be decomposed into a 'direct' and an 'indirect' wealth effect. The direct wealth effect is positive: an increase in wealth affords more healthy consumption (see Table 2 in online Appendix C). Yet, wealth also has an indirect effect: an increase in wealth leads to a higher relative marginal value of health $q_{h/a}(t)$ (Proposition 3), which increases the health benefit of healthy consumption $[-q_{h/a}(t)\partial d/\partial C_h]$. Both the direct and indirect effects operate in the same direction, and wealthy individuals engage more in healthy consumption: $\partial C_h(t)/\partial A_0 > 0$, at least initially.[26]

Similar to healthy consumption, additional wealth enables purchases of more unhealthy consumption goods – the direct wealth effect is positive. Yet, additional wealth also increases the marginal health cost of unhealthy consumption $q_{h/a}(t)\partial d/\partial C_u$ (the indirect wealth effect), through a higher relative marginal value of health $q_{h/a}(t)$. The indirect wealth effect competes with the direct wealth effect.

While we cannot *a priori* sign the relation between unhealthy consumption and wealth, the two competing effects predict an interesting pattern of behaviour. The health cost increases in the severity of its impact on health, $\pi_{dC_u(t)} \propto \partial d/\partial C_u$ (the degree of 'unhealthiness' of the consumption good). This suggests that for moderately unhealthy goods, the direct wealth effect would dominate, while for severely unhealthy goods the indirect wealth effect would dominate.

4.2. *Variation in Wages, δw_E and Education δE*

PROPOSITION 4. *Permanently higher wages and education operate in a similar manner to an increase in wealth δA_0 (Propositions 1–3), with some differences:*

[25] An indirect effect operates through the effect that wealth has on health, and health in turn has on the deterioration rate. It is not clear what the signs of these effects are, since signing these terms requires assumptions on the signs of $\partial d^2/\partial H \partial C_u$, $\partial d^2/\partial H \partial C_h$, etc. As outlined in Assumption 4, we assume these second-order effects are quantitatively less important. The effect of wealth on health is gradual and is therefore at least initially unlikely to drive the effect of wealth on health behaviours.

[26] If wealth enables limited life extension (scenario *I* in Proposition 3), it is possible that the health benefit decreases late in life, since wealth reduces $q_{h/a}(t)$ late in life compared to the unperturbed path.

(*i*) *the wealth effect is muted by the increased opportunity cost of time;*
(*ii*) *permanent wages w_E and education E also raise the production benefit of health; and*
(*iii*) *education raises the efficiency of health investment.*

For proof, see online Appendix D.4.

It is important to distinguish between an evolutionary wage change and permanent differences in the wage rate $w(t)$, i.e. permanent income. In our model of perfect certainty and perfect capital markets, an evolutionary increase in the wage rate $w(t)$ raises the opportunity cost of time but does not affect the marginal value of wealth $q_A(t)$ (i.e. the life-cycle trajectory is unchanged). In contrast, if wages are permanently higher, i.e. larger w_E in (10), earnings are higher over the entire life cycle,[27] and in addition to the opportunity cost of time effect, there is also a wealth effect (operating by decreasing the marginal value of wealth $q_A(t)$; see Table 2 in online Appendix C). Further, the production benefit of health is higher as higher wages increase the value of health in reducing sick time.

There are reasons to believe that the wealth effect and the effect of a higher production benefit of health dominate the opportunity cost of time effect. First, this is consistent with the result by Dustmann and Windmeijer (2000) and Contoyannis *et al.* (2004) that a permanent wage change affects health positively, while a transitory wage increase affects health negatively. Second, it is consistent with the rich literature on SES and health that consistently finds that high-income individuals are generally in better health than low-income individuals.

Permanently higher wages due to education E (see (10)) are also associated with an increased opportunity cost of time effect, a wealth effect and higher production benefits. But, education also increases the efficiency $\mu_I(t; E)$ of health investment, as the educated are assumed to be more efficient consumers and producers of health. The efficiency effect of education has two implications. First, it increases total health investment $I[m(t), \tau_m(t); E]$. However, improved efficiency implies that fewer inputs are required to obtain a certain level of investment, potentially reducing the market $m(t)$ and time inputs $\tau_m(t)$ devoted to health investment (Grossman, 1972*a*,*b*). Second, it could potentially explain the stronger evidence for an effect of education on health and the weaker evidence for effects of income and wealth on health, since the efficiency effect does not operate for income and wealth (see Section 1). Thus, among the socioeconomic indicators, education improves health behaviours and health potentially the most.

[27] Earnings $Y(t)$ are a function of the wage rate $w(t)$ times the amount of time spent working $\tau_w(t)$ (see (8)). A higher wage rate w_E implies that the individual has higher earnings $Y(t)$ because the direct effect of higher wages is to increase earnings. There are also two secondary effects. First, individuals may work more because of the higher opportunity cost of not working (substitution effect). Second, individuals may work fewer hours to spend their increased income on leisure or consumption (income effect). Empirical studies suggest that the substitution and income effects are of the same magnitude (Blundell and MaCurdy, 1999), and hence that the direct effect of a wage increase is to increase earnings, while the secondary effect is small, consisting of two competing effects that roughly cancel out. Thus, a higher wage rate translates into higher earnings.

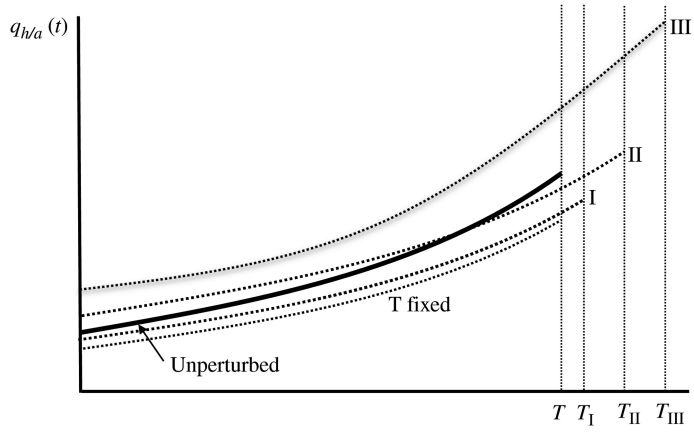

Fig. 3. *Evolution of the Relative Marginal Value of Health $q_{h/a}(t)$ with Age Due to Variation in H_0*
Notes. The solid thick line, labelled 'Unperturbed', represents the unperturbed path. The perturbed paths are shown for the T fixed case (dotted line, labelled 'T fixed'), scenario *I*, associated with small life extension T_{I} (dotted line, labelled 'I'), scenario *II*, associated with intermediate life extension T_{II} (dotted line labelled 'II') and scenario *III*, associated with large life extension (dotted line labelled 'III').

4.3. *Variation in Initial Health, δH_0*

Proposition 5. *Absent ability to extend life, healthy individuals, ceteris paribus, value health cumulatively less, $\int_0^T [\partial q_{h/a}(t)/\partial H_0|_T]\mathrm{d}t < 0$.*

For proof, see online Appendix D.5.

For fixed length of life T, when starting off with a higher level of health, cumulatively the relative marginal value of health has to be lower over the life cycle, leading to cumulatively unhealthier behaviour and lower health investment, in order to arrive at $H_{\min}$ over the same duration of life T. The calibrated model simulations confirm this pattern and Figure 3 illustrates it: the perturbed fixed T path (dotted line) lies below the unperturbed curve and both end at $t = T$.[28]

Proposition 6. *Healthy individuals live longer $\partial T/\partial H_0 \geq 0$.*

For proof, see online Appendix D.6.

Proposition 7. *Individuals with greater endowed health are healthier at all ages, $\partial H(t)/\partial H_0 > 0$, $\forall t$. For small life extension, they cumulatively value health less $\int_0^T [\partial q_{h/a}(t)/\partial H_0]\mathrm{d}t < 0$, for intermediate life extension they value health cumulatively more $\int_0^T [\partial q_{h/a}(t)/\partial H_0]\mathrm{d}t > 0$ and for large life extension they value health more at all ages, $\partial q_{h/a}(t)/\partial H_0 > 0$, $\forall t$.*

For proof, see online Appendix D.7.

[28] Cases are also possible where the relative marginal value of health is initially higher but eventually lower. See online Appendix Section D.5 and Figure 9 for more details.

When T can be optimally chosen, individuals with greater endowed health are healthier throughout life, and live longer (Propositions 6 and 7).[29] We distinguish between three scenarios: 'small', 'intermediate' and 'large' life extension, as illustrated in Figure 3 (see online Appendix Section D.7 and Figure 10 for more detail). For small life extension, healthier individuals value health more than in the fixed T case, but cumulatively still less than for the unperturbed path, and life is extended to T_I. For intermediate life extension, the relative marginal value of health is cumulatively higher compared with the unperturbed path, but health is still valued less in old age. Life is extended to T_{II}. For large life extension, the relative marginal value of health is higher at all ages, and life is extended to T_{III}. In such a scenario, healthy individuals care more about their health as for them investment pays off in terms of a longer lifespan over which the benefits of health, consumption, and leisure may be enjoyed. Since health investment increases in the value of health (see (15) and (17)) it follows similar patterns in these three scenarios.

Which scenario is more plausible depends on the extent to which medical technology, institutional, and environmental factors, allow for endowed health to extend life ($\partial T/\partial H_0$): in developed countries with few competing risks from diseases, universal access to health care, and cutting-edge medical technology, or in developing nations where large gains in longevity can potentially be achieved with relatively low cost interventions such as provision of clean water and improving sanitation. In our calibrated simulations for an average US college graduate, we find $\partial q_{h/a}(t)/\partial H(t) < 0$ for all ages (see Table 2 in Section C), which would be consistent with the 'small' life extension scenario T_I (see Figure 3 and for more detail Figure 10 in online Appendix Section D.7) and is consistent with the empirical regularity that healthy individuals consume less medical care (Van de Ven and van der Gaag, 1982; Wagstaff, 1986*a*; Erbsland *et al.*, 2002).[30]

4.4. *Variation in Work, Leisure and Retirement by SES and Health*

Important variation exists across individuals both in the type of work and in the amount of time spent working, during a day and over the life cycle. We first discuss variation in the type of work and then turn to time spent working.

4.4.1. *Work and job-related health stress*

A higher relative marginal value of health $q_{h/a}(t)$ induced by wealth (Proposition 3) increases the health cost of job-related health stress (see (29)). Eventually, however, wealth leads to better health (Proposition 3) and better health increases the marginal benefit $\partial Y/\partial z$ of job-related health stress through reduced sick time. Permanently higher wages, e.g. through better education, are also associated with the above competing wealth and

[29] If, in contrast to our benchmark scenario, $\partial q_A(0)/\partial H_0 > 0$, then theoretically we cannot rule out that $\partial T/\partial H_0 < 0$ and that $\partial H(t)/\partial H_0 < 0$ after some age t (see online Appendix D.5). This case would represent an extreme scenario in which the complementarity between consumption and health in utility is so strong that healthier individuals choose to spend much more on consumption and much less on health investment, living shorter lives. This scenario is inconsistent with empirical stylised facts, e.g. that healthier people live longer lives and with our calibrated simulations.

[30] However, after accounting for the endogeneity of health, Galama *et al.* (2012) find that the effect of health on medical care use becomes statistically insignificant, suggesting the negative association obtained in empirical studies may not be sufficiently robust.

health effects. In addition, the marginal benefit of job-related health stress $\partial Y/\partial z$ increases directly with the wage rate. Empirical evidence suggests that high SES individuals on average work in less demanding occupations (Ravesteijn *et al.*, 2013). This suggests that higher SES increases the marginal costs of job-related health stress more than it increases its marginal benefits.

The effect of health on job-related health stress is plausibly positive. Better health reduces sick time, which increases the marginal benefit of job-related health stress. Further, if the relative marginal value of health decreases in health (Propositions 5 and 7), then healthier individuals will have lower marginal costs of engaging in job-related health stress. With higher benefits and lower costs, we expect the healthy to engage in unhealthy jobs, consistent with empirical evidence (Kemna, 1987).

4.4.2. *Leisure and retirement*

To analyse retirement, we informally treat a small amount of time devoted to work $\tau_w(t)$, i.e. below a certain threshold, say τ_R, as a retirement phase. With declining health, time spent working $\tau_w(t)$ (see (6)) gradually decreases, as a result of increasing sick time and the increasing demand for time devoted to health investment. Hence, the model produces a phase of life in old age that naturally qualifies as retirement. During working life, individuals divide their time between work, leisure and time inputs into consumption and health investment (see (6)). Therefore, we can infer the effect on the time spent working by investigating effects on leisure and time inputs.

Wealth increases the demand for leisure, for (time inputs into) healthy and unhealthy consumption, and for (time devoted to) health investment, through a 'direct' wealth effect, reducing the marginal value of wealth $q_A(t)$ and an 'indirect' value of health effect, increasing the relative marginal value of health (Proposition 3). This leads wealthier individuals, *ceteris paribus*, to work less (see (6)), and hence retire earlier, in line with empirical evidence (Imbens *et al.*, 2001; Brown *et al.*, 2010).[31] Permanently higher wages, e.g. through education, are also associated with the above wealth and value of health effects, but higher wages also increase the cost of time inputs, i.e. of leisure, of healthy and unhealthy consumption and of health investment (see subsections D.8.2 and D.8.3 in online Appendix D). The higher opportunity cost of not working encourages higher educated individuals to retire later. Thus, the wealthy retire earlier, but the higher educated and those with higher permanent wages may retire later.

Healthier individuals spend more time working, as good health reduces sick time and reduces the demand for (time inputs into) health investment (Propositions 5 and 7). This encourages healthier individuals to work more and retire later (see (6)). However, health is also associated with a wealth effect, reducing the marginal value of wealth $q_A(t)$ (see Table 2 in online Appendix C), which increases the demand for leisure and for (time inputs into) healthy and unhealthy consumption, and thereby encourages early retirement.[32] While the net effect is therefore ambiguous, it seems plausible that the direct effect of health on reducing sick time and reducing time inputs into health investment,

[31] However, wealthy individuals are healthier (Proposition 3), which reduces sick time and increases the benefit of work (earnings). The disparity in health grows over time, so that eventually the effect of better health among the wealthy may become important in the retirement decision.

[32] The amount of time devoted to work further depends on whether leisure, consumption and health are complements or substitutes in utility.

outweighs the indirect effect of health on leisure by relaxing the wealth constraint. This is consistent with an extensive literature showing quantitatively large effects of health on labour force participaton, with unhealthier individuals retiring earlier (Currie and Madrian, 1999; Smith, 2007).

PREDICTION 4. *Under plausible assumptions, healthier individuals retire later. This, combined with an effect of health on earnings, leads to reverse causality as healthier individuals accumulate more wealth by earning more and retiring later.*

5. Discussion and Conclusions

We have developed a theory of the relation between health and SES over the life cycle. Our life-cycle model incorporates health, longevity, wealth, earnings, education, work, job-related physical and psychosocial health stressors, leisure, health investment (e.g. exercise, medical care) and healthy and unhealthy consumption (including housing, neighbourhood social environment). Our review of the literature identifies these as essential mechanisms in the formation and evolution of disparities in health.

The theory is capable of reproducing stylised facts regarding the life-cycle profiles of health and health investment, as illustrated by calibrated simulations of the model. The theory is further able to reproduce stylised facts characteristic of the SES-health gradient. We find that greater SES, as measured by wealth, earnings and education, induces a healthy lifestyle: it encourages investment in health, encourages healthy consumption, discourages unhealthy consumption and protects individuals from the health risks of physically and psychosocially demanding working conditions. The healthier lifestyle of high SES individuals causes the health trajectories of high and low SES individuals to diverge. As a result, they are healthier and live longer (Propositions 2–4). In addition, health generates earnings and the worsening health of low SES individuals potentially leads to early withdrawal from the labour force (Prediction 4). This reverse causality from health to financial measures of SES potentially reinforces the widening of the SES-health gradient, as documented in empirical studies (Smith, 2007).

In middle to late life, the divergence of health trajectories potentially slows as lower levels of health encourages low SES individuals to invest more in health and engage in healthier behaviour in order to slow down their health deterioration (Propositions 3 and 4). Also, mortality selection, i.e. the least healthy among lower SES individuals die sooner, results in an apparently healthier surviving disadvantaged population, potentially narrowing the gradient in late age.[33] Thus, the theory is capable of reproducing the characteristic life-cycle patterns of the SES-health gradient.

Apart from providing a framework to interpret stylised facts, the theory also makes novel testable predictions and provides new intuition. In particular, we emphasise the importance of our concept of a health cost (benefit) of unhealthy (healthy) behaviours, in explaining health behaviour. Individuals make decisions regarding health by taking into account not just monetary prices and preferences, but additionally the life-time

[33] The narrowing of the gradient due to unhealthier individuals engaging in healthier behaviour would represent an economic variant of the age-as-leveller hypothesis, while the narrowing of the gradient due to mortality selection would be consistent with a process of cumulative advantage (House *et al.*, 1994; Ross and Wu, 1996; Lynch, 2003).

health consequences of their choices, as embodied by the health cost (benefit). Variation in the health cost over the life cycle and across SES potentially explains several empirical phenomena.

For example, we predict that individuals in mid-life, particularly the healthy and poor, engage in work associated with unhealthy working conditions as they value the associated wage premium. However, as individuals age they engage in healthier work, as the health cost of unhealthy working conditions increase with declining health (Prediction 1). Another implication of our concept of a health cost is a pattern in which high SES individuals consume more of moderately unhealthy consumption goods (e.g. moderate alcohol consumption) and less of severely unhealthy consumption goods (e.g. cigarettes, high alcohol consumption, illicit drugs) than do lower SES individuals (Prediction 3). Greater wealth permits more consumption but also increases the health cost. This could provide an explanation for the observation that high SES individuals are less likely to smoke cigarettes (bad for health) but are more likely to be moderate drinkers (moderately bad) than low SES individuals (Cutler and Lleras-Muney, 2008).

Another insight is that (endogenous) longevity is crucial in explaining observed associations between SES and health (cf. Propositions 1, 3 and 4, and Prediction 2). Absent ability to extend life (fixed horizon), the association between SES and health is small (Proposition 1). If, however, life can be extended, SES and health are positively associated and the greater the degree of life extension afforded by SES, the greater is their association (Propositions 3 and 4). Thus, health disparities are larger in environments where higher SES individuals can effectively use their resources to extend life (Prediction 2). For example, if the latest medical technology is more easily accessible to higher SES individuals, health disparities across SES groups may be larger.

In deriving predictions from the comparative dynamic analyses, we have made a number of assumptions, most of which are conventional, such as diminishing marginal utility and Cobb–Douglas production functions. Our calibrated model corroborated other necessary relations. However, one assumption we had to make is that first-order effects dominate second-order effects (Assumption 4 in subsection 2.4). Future work could investigate the sensitivity of results to this assumption, although it is less restrictive than the fairly conventional assumption of functions being additively separable in their arguments, and other (potentially stronger) assumptions are likely needed to derive unambiguous predictions from the comparative dynamic analyses.

Future work may also extend the model to incorporate the joint determination of SES and health (Chiteji, 2010; Conti *et al.*, 2010), the evolution of child health (Case *et al.*, 2002; Currie and Stabile, 2003; Heckman, 2007) and the impact of fetal and early-childhood conditions on health in adulthood (Barker *et al.*, 1993; Case *et al.*, 2005).[34] Early childhood could be included modelling the production of health by the family, similar to, e.g. Jacobson (2000) and Bolin *et al.* (2001). We do not explicitly take into account the influence of the wider social context and social relationships of the family or neighbourhood on health (Kawachi and Berkman, 2003) or of social capital on health (Bolin *et al.*, 2003). Insights from the behavioural-economic and psychological literature regarding myopia and lack of self-control (Blanchflower *et al.*, 2009) might be incorporated following Laibson (1998). Uncertainty (e.g. health shocks) could be

[34] See Galama and van Kippersluis (2015) and Strulik (2016) for theories of endogenous education and health.

included similar to, e.g. Cropper (1977), Dardanoni and Wagstaff (1990), Liljas (1998) and Ehrlich (2000).

University of Southern California
Erasmus University Rotterdam, Tinbergen Institute and Chinese University of Hong Kong

Accepted: 13 November 2017

Additional Supporting Information may be found in the online version of this article:

Appendix A. First-order Conditions.
Appendix B. Functional Relations.
Appendix C. Calibrated Simulations.
Appendix D. Proofs of Propositions.
Data S1.

References

Adler, N.E., Boyce, T., Chesney, M., Folkman, S. and Syme, L. (1993). 'Socioeconomic inequalities in health: no easy solution', *Journal of the American Medical Association*, vol. 269(24), pp. 3140–5.
Adler, N.E., Boyce, T., Chesney, M.A., Cohen, S., Folkman, S., Kahn, R.L. and Syme, S.L. (1994). 'Socioeconomic status and health: the challenge of the gradient', *American Psychologist*, vol. 49(1), pp. 15–24.
Albouy, V. and Lequien, L. (2009). 'Does compulsory education lower mortality?', *Journal of Health Economics*, vol. 28(1), pp. 155–68.
Almond, D. and Currie, J. (2011). 'Killing me softly: the fetal origins hypothesis', *Journal of Economic Perspectives*, vol. 25(3), pp. 153–72.
Arking, R. (2006). *Biology of Aging: Observations and Principles*, Oxford: Oxford University Press.
Auld, M.C. and Sidhu, N. (2005). 'Schooling, cognitive ability, and health', *Health Economics*, vol. 14(10), pp. 1019–34.
Baeten, S., Van Ourti, T. and van Doorslaer, E. (2013). 'The socioeconomic health gradient across the life cycle: what role for selective mortality and institutionalization?', *Social Science and Medicine*, vol. 97, pp. 66–74.
Barker, D., Gluckman, P.D., Godgrey, K.M., Harding, J.E., Owens, J.A. and Robinson, J.S. (1993). 'Fetal nutrition and cardiovascular disease in adult life', *Lancet*, vol. 341(8850), pp. 938–41.
Becker, G. (1965). 'A theory of the allocation of time', Economic Journal, vol. 75(299), pp. 493–517.
Becker, G. (2007). 'Health as human capital: synthesis and extensions', *Oxford Economic Papers*, vol. 59(3), pp. 379–410.
Becker, G. and Murphy, K. (1988). 'A theory of rational addiction', *Journal of Political Economy*, vol. 96(4), pp. 675–700.
Beckett, M. (2000). 'Converging health inequalities in later life: an artifact of mortality selection?', *Journal of Health and Social Behavior*, vol. 41(1), pp. 106–19.
Blanchflower, D.G., Oswald, A.J. and van Landeghem, B. (2009). 'Imitative obesity and relative utility', *Journal of the European Economic Association*, vol. 7(2–3), pp. 528–38.
Blau, D.M. (2008). 'Retirement and consumption in a life cycle model', *Journal of Labor Economics*, vol. 26(1), pp. 35–71.
Blundell, R. and MaCurdy, T. (1999). 'Labor supply', in (O. Ashenfelter and D. Card, eds.), *Handbook of Labor Economics*, pp. 1559–695, Amsterdam: North-Holland.
Bolin, K., Jacobson, L. and Lindgren, B. (2001), 'The family as the producer of health – when spouses are Nash bargainers', *Journal of Health Economics*, vol. 20(3), pp. 349–62.
Bolin, K., Lindgren, B., Lindstrom, M. and Nystedt, P. (2003). 'Investments in social capital – implications of social interactions for the production of health', *Social Science & Medicine*, vol. 56(12), pp. 2379–90.
Borg, V. and Kristensen, T.S. (2000). 'Social class and self-rated health: can the gradient be explained by differences in life style or work environment?', *Social Science & Medicine*, vol. 51(7), pp. 1019–30.
Brown, J.R., Coile, C.C. and Weisbenner, S.J. (2010). 'The effect of inheritance receipt on retirement', *Review of Economics and Statistics*, vol. 92(2), pp. 425–34.
Campbell, F., Conti, G., Heckman, J.J., Moon, S.H., Pinto, R., Pungello, E. and Pan, Y. (2014). 'Early childhood investments substantially boost adult health', *Science*, vol. 343(6178), pp. 1478–85.
Caputo, M.R. (2005). *Foundations of Dynamic Economic Analysis*, Cambridge: Cambridge University Press.

Case, A. and Deaton, A. (2005). 'Broken down by work and sex: how our health declines', in (D.A. Wise, ed.), *Analyses in the Economics of Aging*, pp. 185–212, Chicago, IL: The University of Chicago Press.
Case, A., Fertig, A. and Paxson, C. (2005). 'The lasting impact of childhood health and circumstance', *Journal of Health Economics*, vol. 24(2), pp. 365–89.
Case, A., Lubotsky, D. and Paxson, C. (2002). 'Economic status and health in childhood: the origins of the gradient', *American Economic Review*, vol. 92(5), pp. 1308–34.
Cawley, J. and Ruhm, C. (2012). 'The economics of risky health behaviors', in (M.V. Pauly, T.G. McGuire and P.P. Barros, eds.), *Handbook of Health Economics*, vol. 2, pp. 95–200, Waltham, MA: Elsevier.
Chiteji, N. (2010). 'Time preference, noncognitive skills and well being across the life course: do noncognitive skills encourage healthy behavior?', *American Economic Review: Papers & Proceedings*, vol. 100(2), pp. 200–4.
Clark, D. and Royer, H. (2013). 'The effect of education on adult mortality and health: evidence from Britain', *American Economic Review*, vol. 103(6), pp. 2087–120.
Conti, G., Heckman, J.J. and Urzua, S. (2010). 'The education-health gradient', *American Economic Review Papers and Proceedings*, vol. 100(2), pp. 234–8.
Contoyannis, P., Jones, A.M. and Rice, N. (2004). 'The dynamics of health in the British Household Panel Survey', *Journal of Applied Econometrics*, vol. 19(4), pp. 473–503.
Cropper, M.L. (1977). 'Health, investment in health, and occupational choice', *Journal of Political Economy*, vol. 85(6), pp. 1273–94.
Cropper, M.L. (1981). 'Measuring the benefits from reduced morbidity', *American Economic Review*, vol. 71(2), pp. 235–40.
Cunha, F. and Heckman, J.J. (2007). 'The technology of skill formation', *American Economic Review*, vol. 97(2), pp. 31–47.
Currie, J. (2009). 'Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development', *Journal of Economic Literature*, vol. 47(1), pp. 87–122.
Currie, J. and Madrian, B.C. (1999). 'Health, health insurance and the labour market', in (O. Ashenfelter and D. Card, eds.), *Handbook of Labour Economics*, vol. 3, pp. 3309–415, Amsterdam: Elsevier Science.
Currie, J. and Stabile, M. (2003). 'Socioeconomic status and child health: why is the relationship stronger for older children?', *American Economic Review*, vol. 93(5), pp. 1813–23.
Cutler, D.M. and Lleras-Muney, A. (2008). 'Education and health: evaluating theories and evidence', in (R.F. Schoeni, J.S. House, G.A. Kaplan and H. Pollack, eds.), *Making Americans Healthier: Social and Economic Policy As Health Policy* (National Poverty Center Series on Poverty and Public Policy), pp. 29–60, New York: Russell Sage Foundation.
Cutler, D.M. and Lleras-Muney, A. (2010). 'Understanding differences in health behaviors by education', *Journal of Health Economics*, vol. 29(1), pp. 1–28.
Cutler, D.M., Lleras-Muney, A. and Vogl, T. (2011). 'Socioeconomic status and health: dimensions and mechanisms', in (P. Smith and S. Glied, eds.), *Oxford Handbook of Health Economics*, pp. 124–63, Oxford: Oxford University Press.
Dalgaard, C-J. and Strulik, H. (2014). 'Optimal aging and death: understanding the Preston curve', *Journal of the European Economic Association*, vol. 12(3), pp. 672–701.
Dalgaard, C-J. and Strulik, H. (2015). 'The economics of health demand and human aging: health capital vs. health deficits', Discussion Paper, University of Goettingen.
Dalgaard, C-J. and Strulik, H. (2017). 'The genesis of the golden age – accounting for the rise in health and leisure', *Review of Economic Dynamics*, vol. 24, pp. 132–51.
Dardanoni, V. and Wagstaff, A. (1990). 'Uncertainty and the demand for medical care', *Journal of Health Economics*, vol. 9(1), pp. 23–38.
De Magalhaes, J.P. (2003). 'Is mammalian aging genetically controlled?', *Biogerontology*, vol. 4(2), 119–20.
De Magalhaes, J.P. (2011). 'The biology of ageing: a primer', in (I. Stuart-Hamilton, ed.), *An Introduction to Gerontology*, pp. 21–47, Cambridge: Cambridge University Press.
De Magalhaes, J.P. (2014). 'senescence.info', available at: http://www.senescence.info (last accessed: 21 July 2017)
Deaton, A. (2002). 'Policy implications of the gradient of health and wealth', *Health Affairs*, vol. 21(2), pp. 13–30.
Deaton, A. and Paxson, C. (1998). 'Aging and inequality in income and health', *American Economic Review*, vol. 88(2), pp. 248–53.
DeNardi, M., French, E., Bailey Jones, J. and McCauley, J. (2016). 'Medical spending of the U.S. elderly', *Fiscal Studies*, vol. 37(3–4), pp. 717–47.
DHHS (2014). 'Current cigarette smoking among adults – United States, 2005–2012', US Department of Health and Human Services, *MMWR*, vol. 63(2), pp. 29–46.
Dustmann, C. and Windmeijer, F. (2000). 'Wages and the demand for health – a life cycle analysis', IZA Discussion Paper No. 171.
Ehrlich, I. (2000). 'Uncertain lifetime, life protection, and the value of life saving', *Journal of Health Economics*, vol. 19(3), pp. 341–67.

Ehrlich, I. and Chuma, H. (1990). 'A model of the demand for longevity and the value of life extension', *Journal of Political Economy*, vol. 98(4), pp. 761–82.
Eisenring, C. (1999). 'Comparative dynamics in a health investment model', *Journal of Health Economics*, vol. 18(5), pp. 655–60.
Erbsland, M., Ried, W. and Ulrich, V. (2002). 'Health, health care, and the environment: econometric evidence from German micro data', in (A. M. Jones and O. O'Donnell, eds), *Econometric Analysis of Health Data*, pp. 25–36, New York: John Wiley & Sons.
Finkelstein, A., Luttmer, E.F. and Notowidigdo, M.J. (2013). 'What good is wealth without health? The effect of health on the marginal utility of consumption', *Journal of the European Economic Association*, vol. 11(s1), pp. 221–58.
Forster, M. (2001). 'The meaning of death: some simulations of a model of healthy and unhealthy consumption', *Journal of Health Economics*, vol. 20(4), pp. 613–38.
Fuchs, V.R. (1986). *The Health Economy*, Cambridge, MA: Harvard University Press.
Galama, T.J. (2015). 'A contribution to health capital theory', CESR Schaeffer Working Paper No. 2015-004.
Galama, T.J. and Kapteyn, A. (2011). 'Grossman's missing health threshold', *Journal of Health Economics*, vol. 30(5), pp. 1044–56.
Galama, T.J. and van Kippersluis, H. (2013). 'Health inequalities through the lens of health capital theory: issues, solutions, and future directions', in (J. Bishop, P.R. Dias and O. O'Donnell, eds.), *Research on Economic Inequality. Vol. 21: Health and Inequality*, pp. (263–284), Bingley: Emerald Group.
Galama, T.J. and van Kippersluis, H. (2015). 'A theory of education and health', CESR–Schaeffer Working Paper Series No. 2015-001.
Galama, T.J., Hullegie, P., Meijer, E. and Outcault, S. (2012). 'Is there empirical evidence for decreasing returns to scale in a health capital model?', *Health Economics*, vol. 21(9), pp. 1080–100.
Gjerde, J., Grepperud, S. and Kverndokk, S. (2005). 'On adaptation and the demand for health', *Applied Economics*, vol. 37(11), pp. 1283–301.
Glied, S. and Lleras-Muney, A. (2008). 'Technological innovation and inequality in health', *Demography*, vol. 45(3), pp. 741–61.
Goldman, D.P. and Smith, J.P. (2002). 'Can patient self-management help explain the SES health gradient?', *Proceedings of the National Academy of Science*, vol. 99(16), pp. 10929–34.
Grossman, M. (1972*a*). *The Demand for Health – a Theoretical and Empirical Investigation*, New York: National Bureau of Economic Research.
Grossman, M. (1972*b*). 'On the concept of health capital and the demand for health', *Journal of Political Economy*, vol. 80(2), pp. 223–55.
Grossman, M. (2000). 'The human capital model', in (A.J. Culyer and J.P. Newhouse, eds.), *Handbook of Health Economics*, vol. 1, pp. 347–408, Amsterdam: Elsevier Science.
Hall, R.E. and Jones, C.I. (2007). 'The value of life and the rise in health spending', *Quarterly Journal of Economics*, vol. 122(1), pp. 39–72.
Halliday, T.J., He, H., Ning, L. and Zhang, H. (2017). 'Health investment over the life-cycle', *Macroeconomic Dynamics*, forthcoming.
Heckman, J.J. (2007). 'The economics, technology, and neuroscience of human capability formation', *Proceedings of the National Academy of Sciences*, vol. 104(33), pp. 13250–5.
Herd, P., Schoeni, R.F. and House, J.S. (2008). 'Upstream solutions: does the supplemental security income program reduce disability in the elderly?', *Milbank Quarterly*, vol. 86(1), pp. 5–45.
House, J.S., Lepkowski, J.M., Kinney, A.M., Mero, R.P., Kessler, R.C. and Regula Herzog, A. (1994). 'The social stratification of aging and health', *Journal of Health and Social Behavior*, vol. 35(3), pp. 213–34.
Hugonnier, J., Pelgrin, F. and St-Amour, P. (2013). 'Health and (other) asset holdings', *Review of Economic Studies*, vol. 80(2), pp. 663–710.
Hugonnier, J., Pelgrin, F. and St-Amour, P.(2016). 'Closing down the shop: optimal health and wealth dynamics near the end of life', manuscript, Faculty of Business and Economics (HEC), University of Lausanne.
Imbens, G., Rubin, D. and Sacerdote, B. (2001). 'Estimating the effect of unearned income on labor supply, earnings, savings and consumption: evidence from a survey of lottery players', *American Economic Review*, vol. 91(4), pp. 778–94.
Jacobson, L. (2000). 'The family as producer of health – an extension of the Grossman model', *Journal of Health Economics*, vol. 19(5), pp. 611–37.
Jones, A., Laporte, A., Rice, N. and Zucchelli, E. (2014). 'A synthesis of the Grossman and Becker-Murphy models of health and addiction: theoretical and empirical implications', Centre for Health Economics, University of York Working Paper.
Kawachi, I. and Berkman, L. (Eds.) (2003). *Neighborhoods and Health*, Oxford: Oxford University Press.
Kemna, H. (1987). 'Working conditions and the relationship between schooling and health', *Journal of Health Economics*, vol. 6(3), pp. 189–210.
Kuhn, M., Wrzaczek, S., Prskawetz, A. and Feichtinger, G. (2015). 'Optimal choice of health and retirement in a life-cycle model', *Journal of Economic Theory*, vol. 158(Part A), pp. 186–212.

Laibson, D. (1998). 'Life-cycle consumption and hyperbolic discount functions', *European Economic Review*, vol. 42(3–5), pp. 861–71.
Liljas, B. (1998). 'The demand for health with uncertainty and insurance', *Journal of Health Economics*, vol. 17(2), pp. 153–70.
Lleras-Muney, A. (2005). 'The relationship between education and adult mortality in the United States', *Review of Economic Studies*, vol. 72(1), pp. 189–221.
Lleras-Muney, A. and Lichtenberg, F. (2005). 'The effect of education on medical technology adoption: are the more educated more likely to use new drugs', Annales d'Economie et Statistique, special issue in memory of Zvi Griliches, No. 79/80, pp. 671–96.
Longo, V.D., Mitteldorf, J. and Skulachev, V.P. (2005). 'Programmed and altruistic ageing', *Nature Reviews Genetics*, vol. 6(11), pp. 866–72.
Lynch, S.M. (2003). 'Cohort and life-course patterns in the relationship between education and health: a hierarchical approach', *Demography*, vol. 40(2), pp. 309–31.
Lynch, J.W., Kaplan, G.A. and Shema, S.J. (1997). 'Cumulative impact of sustained economic hardship on physical, cognitive, psychological, and social functioning', *New England Journal of Medicine*, vol. 337(26), pp. 1889–95.
Mackenbach, J.P., Huisman, M., Andersen, O., Bopp, M., Borgan, J.K., Borrell, C., Costa, G., Deboosere, P., Donkin, A., Gadeyne, S., Minder, C., Regidor, E., Spadea, T., Valkonen, T. and Kunst, A.E. (2004). 'Inequalities in lung cancer mortality by the educational level in 10 European populations', *European Journal of Cancer*, vol. 40(1), pp. 126–35.
Marmot, M. (1999). 'Multi-level approaches to understanding social determinants', in (L. Berkman and I. Icawachi, eds.), *Social Epidemiology*, pp. 349–66, Oxford: Oxford University Press.
Marmot, M., Ryff, C.D., Bumpass, L.L., Shipley, M. and Marks, N.F. (1997). 'Social inequalities in health: next questions and converging evidence', *Social Science & Medicine*, vol. 44(6), pp. 901–10.
Michaud, P.-C. and van Soest, A. (2008). 'Health and wealth of elderly couples: causality tests using dynamic panel data models', *Journal of Health Economics*, vol. 27(5), pp. 1312–25.
Mincer, J.A. (1974). *Schooling, Experience, and Earnings*, New York: Columbia University Press.
Moody, H.R. and Sasser, J.R. (2014). *Aging: Concepts and Controversies*, 8th edn. Los Angeles, CA: Sage Publications.
Murphy, K.M. and Topel, R.H. (2006). 'The value of health and longevity', *Journal of Political Economy*, vol. 114(5), pp. 871–904.
Muurinen, J-M. (1982). 'Demand for health: a generalized Grossman model', *Journal of Health Economics*, vol. 1(1), pp. 5–28.
Oster, E., Shoulson, I. and Dorsey, E.R. (2013). 'Limited life expectancy, human capital and health investments', *American Economic Review*, vol. 103(5), 1977–2002.
Pearson, T., Russell, J., Campbell, M.J. and Barker, M.E. (2005). 'Do "food deserts" influence fruit and vegetable consumption? A cross-sectional study', *Appetite*, vol. 45(2), pp. 195–7.
Podor, M. and Halliday, T.J. (2012). 'Health status and the allocation of time', *Health Economics*, vol. 21(5), pp. 514–27.
Poterba, J.M., Venti, S.F. and Wise, D.A. (2011). 'Were they prepared for retirement? Financial status at advanced ages in the HRS and AHEAD cohorts', in (D.A. Wise, ed.), *Investigations in the Economics of Aging*, pp. 21–69. Chicago, IL: University of Chicago Press.
Poterba, J., Venti, S. and Wise, D.A. (2013). 'Health, education, and the postretirement evolution of household assets', *Journal of Human Capital*, 7(4), pp. 297–339.
Ravesteijn, B., Van Kippersluis, H. and Van Doorslaer, E. (2013). 'The contribution of occupation to health inequality', in (P.R. Dias and O. O'Donnell, eds.), *Research on Economic Inequality. Vol. 21: Health and Inequality*, pp. 263–284, Bingley: Emerald Group Publishing.
Ried, W. (1998). 'Comparative dynamic analysis of the full Grossman model', *Journal of Health Economics*, vol. 17(4), pp. 383–425.
Rosen, S. (1988). 'The value of changes in life expectancy', *Journal of Risk and Uncertainty*, vol. 1(3), pp. 285–304.
Ross, C.E. and Wu, C.L. (1996). 'Education, age, and the cumulative advantage in health', *Journal of Health and Social Behavior*, 37(1), pp. 104–20.
Savelyev, P. (2014). 'Socioemotional skills, education, and longevity of high-ability individuals', Working Paper, Vanderbilt University.
Scholz, J.K. and Seshadri, A. (2016). 'Health and wealth in a lifecycle model', University of Wisconsin-Madison Working Paper.
Schuenemann, J., Strulik, H. and Trimborn, T. (2017). 'The gender gap in mortality: how much is explained by behavior?', *Journal of Health Economics*, vol. 54, pp. 79–90.
Seierstad, A. and Sydsaeter, K. (1987). 'Optimal control theory with economic applications', in (C.J. Bliss and M.D. Intriligator, eds.), *Advanced Texbooks in Economics*, vol. 24, Amsterdam, North Holland: Elsevier.

Serdula, M.K., Gillespie, C., Kettel-Khan, L., Farris, R., Seymour, J. and Denny, C. (2004). 'Trends in fruit and vegetable consumption among adults in the United States: behavioral risk factor surveillance system, 1994–2000', *American Journal of Public Health*, vol. 94(6), pp. 1014–8.
Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*, Reprinted R.H. Campbell and A.S. Skinner, eds., Oxford: Clarendon Press, 1976.
Smith, J.P. (2007). 'The impact of socioeconomic status on health over the life course', *Journal of Human Resources*, vol. 42(4), pp. 739–64.
Strulik, H. (2015). 'Frailty, mortality, and the demand for medical care', *Journal of the Economics of Ageing*, vol. 6, pp. 5–12.
Strulik, H. (2016). 'The return to education in terms of wealth and health', Cege Discussion Paper No. 293.
Van de Ven, W. and van der Gaag, J. (1982). 'Health as an unobservable: a MIMIC-model of the demand for health care', *Journal of Health Economics*, vol. 1, pp. 157–83.
Van Kippersluis, H. and Galama, T.J. (2014). 'Wealth and health behavior: testing the concept of a health cost', *European Economic Review*, vol. 72, pp. 197–220.
Van Kippersluis, H., Van Ourti, T., O'Donnell, O.A. and van Doorslaer, E. (2009). 'Health and income across the life cycle and generations in Europe', *Journal of Health Economics*, vol. 28(4), pp. 818–30.
Van Kippersluis, H., O'Donnell, O., van Doorslaer, E. and Van Ourti, T. (2010). 'Socioeconomic differences in health over the life cycle in an Egalitarian Country', *Social Science and Medicine*, vol. 70(3), pp. 428–38.
Van Kippersluis, H., O'Donnell, O. and Van Doorslaer, E. (2011). 'Long-run returns to education does schooling lead to an extended old age?', *Journal of Human Resources*, vol. 46(4), pp. 695–721.
Viscusi, K.W. (1978). 'Wealth effects and earnings premiums for job hazards', *Review of Economics and Statistics*, vol. 60(3), pp. 408–16.
Viscusi, K.W. and Aldy, J.E. (2003). 'The value of a statistical life: a critical review of market estimates throughout the world', *Journal of Risk and Uncertainty*, vol. 27(1), pp. 5–76.
Wagstaff, A. (1986*a*). 'The demand for health: some new empirical evidence', *Journal of Health Economics*, vol. 5(3), pp. 195–233.
Wagstaff, A. (1986*b*). 'The demand for health: theory and applications', *Journal of Epidemiology and Community Health*, vol. 40(1), pp. 1–11.
Zeckhauser, R. (1970). 'Medical insurance: a case study of the tradeoff between risk spreading and appropriate incentives', *Journal of Economic Theory*, vol. 2(1), pp. 10–26.
Zweifel, P. and Breyer, F. (1997). *Health Economics*, New York: Oxford University Press.
Zweifel, P., Felder, S. and Meiers, M. (1999). 'Ageing of population and health care expenditures: a red herring?', *Health Economics*, vol. 8(6), pp. 485–96.