WORKING PAPER

A Theory of Socioeconomic Disparities in Health over the Life Cycle

TITUS GALAMA, HANS VAN KIPPERSLUIS

WR-773

July 2010

This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the NICHD funded RAND Population Research Center (R24HD050906).

RAND LABOR AND POPULATION

A theory of socioeconomic disparities in health over the life cycle*

Titus J. Galama† Hans van Kippersluis‡

July 22, 2010

Abstract

Understanding of the substantial disparity in health between low and high socioeconomic status (SES) groups is hampered by the lack of a sufficiently comprehensive theoretical framework to interpret empirical facts and to predict yet untested relations. We present a life-cycle model that incorporates multiple mechanisms explaining (jointly) a large part of the observed disparities in health by SES. In our model, lifestyle factors, working conditions, retirement, living conditions and curative care are mechanisms through which SES, health and mortality are related. Our model predicts a widening and possibly a subsequent narrowing with age of the gradient in health by SES.

Keywords: socioeconomic status, education, health, demand for health, health capital, medical care, life cycle, age, labor, retirement, mortality

JEL Codes : D91, I10, I12, J00, J24

*This research was supported by the National Institute on Aging, under grants R01AG030824, R01AG037398, P30AG012815 and P01AG022481. We are grateful to Arie Kapteyn, Michael Grossman, Owen O'Donnell, Arthur van Soest, Erik Meijer, Eddy van Doorslaer, Tom Van Ourti, Rosalie Pacula, Jan Brinkhuis, Bas Jacobs, Mauricio Avendano, Megan Beckett, and seminar participants at the University of Southern California and Erasmus University for useful discussions and comments.

†RAND Corporation

‡Erasmus School of Economics, Erasmus University Rotterdam (EUR)

1 Introduction

Disparities in health across socioeconomic status (SES) groups – often called the SES health gradient – are substantial. For example, Case and Deaton (2005) show how in the United States, 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 Glasgow, U.K., life expectancy of men in the most deprived areas is 54 years, compared with 82 years in the most affluent (Hanlon et al. 2006). 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. Similar patterns hold for other measures of SES, such as education and wealth and other indicators of health, such as onset of chronic diseases, disability and mortality (e.g., Adler et al. 1994; Marmot, 1999; Smith, 1999). This pattern is remarkably similar between countries with relatively low levels of protection from loss of work and health risks, such as the U.S., and those with stronger welfare systems, such as the Netherlands (House et al. 1994; Kunst and Mackenbach, 1994; Preston and Elo, 1995; Smith 1999; 2004; 2007; Case and Deaton, 2005; van Kippersluis et al. 2010).

Recent significant contributions to the understanding of socioeconomic 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 health (Lleras-Muney, 2005; Oreopoulos, 2006; Silles, 2009) but it is not known exactly how the more educated achieve their health advantage.

Understanding of the relative importance of underlying mechanisms responsible for the observed relationships is hampered by the lack of a sufficiently comprehensive theory. Case and Deaton (2005) argue that it is extremely difficult to understand the relationships between health, education, income, wealth and labor-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 (e.g., Case and Deaton, 2005; Cutler et al. 2010) have pointed to the absence of a theory of SES and health over the life cycle and have emphasized the importance of developing one.

A suitable framework in which multiple mechanisms and their cumulative long-term effects can be studied is a structural model of SES and health over the life cycle. Case and Deaton (2005) have attempted to develop a model for the role of work and consumption behavior in explaining the SES-health gradient. Their starting point is the canonical life cycle model of the demand for health and medical care, due to Grossman (1972a; 1972b). Case and Deaton (2005) present a simplified Grossman model and extend the model to include the detrimental effect of hard/risky labor and of unhealthy consumption behavior on health. However, the authors conclude that the model is not able to explain a number of the most salient features of the SES health gradient. For example, Case and Deaton (2005) argue that while the model can explain differences in the *level* of health between low and high SES groups, it cannot explain differences in the *rate* of health decline. In other words, it cannot account for the widening of the SES health gradient with age through late middle age or early late life, as is observed in empirical studies. Other problems with some of the predictions and properties of health production models have been pointed out in the literature (see Grossman, 2000, for a review and rebuttal of these).

The aim of this paper is to develop a conceptual framework for health and socioeconomic status over the life cycle. The framework includes simplified representations of major mechanisms, which allows us to improve our understanding of their operational roles in explaining the SES health gradient and make predictions. Our starting point is the health production literature spawned by Grossman (Grossman, 1972a; 1972b) and the extensions presented by Ehrlich and Chuma (1990) and Case and Deaton (2005). Our contribution is as follows. First, we address a number of issues identified with this strand of the literature by noting that what is generally interpreted as the equilibrium condition for health can alternatively be interpreted as the first-order condition for health investment (as in Galama, 2010). This interpretation necessitates the assumption of decreasing-returns-to-scale (DRTS) in the health production function (as in Ehrlich and Chuma, 1990), and addresses (i) the indeterminacy problem ("bang-bang" solution) for investment in medical care (Ehrlich and Chuma, 1990), (ii) the inability to reproduce the observed negative relation between health and the demand for medical care (e.g., Zweifel and Breyer, 1997),[1] (iii) the lack of history in the model solutions (e.g., Usher, 1975) and (iv) the lack of capacity to explain differences in the rate of health decline between different socioeconomic groups (Case and Deaton, 2005). With these essential issues addressed our formulation can account for a greater number of observed empirical patterns and suggests that the Grossman model provides a suitable foundation for the development of a life-cycle model of the SES-health gradient.

Yet, utilization of medical services and access to care explain only part of the association between SES and health (e.g., Adler et al. 1993). Our second contribution is therefore to incorporate many potential mechanisms in the model that could explain disparities in health by SES and to include a multitude of potential bi-directional pathways between health and dimensions of SES. One important concept in our work is "job-related health stress", which can be interpreted broadly and can range from physical working conditions (e.g., hard labor) to the psychosocial aspects of work (e.g., low status, limited control, repetitive work, etc). The notion here is that job-related health stress can include any aspect of work that is detrimental to health and as such is associated with a wage premium (a compensating wage differential). Other important features of the model are lifestyle factors (preventive care, healthy and unhealthy consumption), curative (medical) care, labor force withdrawal (retirement) and mortality. The model integrates a life cycle approach, and the concepts of financial, education and health capital (Muurinen and Le

[1] It is not entirely correct to assert that health production models always produce the incorrect sign for the relationship between health and investment in curative care. For the pure investment model and assuming that the biological aging rate increases with age, investment in curative care increases with age while the health stock falls if the elasticity of the marginal production benefit of health with respect to health is less than one (Grossman refers to this as the MEC schedule; Grossman, 2000 [p. 369]). This produces a negative correlation between health and medical care.

Grand, 1985). The focus is on understanding the SES-health gradient as the outcome of rational (constrained) individual behavior, and the framework applies to individuals who have completed their education and participate (or have participated) in the labor-force.

We explore the characteristics of the first-order conditions for a fixed retirement age and a fixed age of death. We find that greater initial wealth, permanently higher earnings (over the life cycle) and a higher level of education induce individuals to invest more in curative and in preventive care, shift consumption toward healthy consumption, and enable individuals to afford healthier working environments (associated with lower levels of physical and psychosocial health stresses) and living environments. The mechanism through which initial wealth, permanent income and education operates is by increasing the demand for curative care and raising the marginal cost of curative care. A higher marginal cost of curative care, in turn, increases the health benefit of (and hence demand for) preventive care and healthy consumption, and the health cost of (and hence reduced demand for) unhealthy working and living environments, and unhealthy consumption. Jointly these behavioral choices gradually lead to growing health advantage with age. Further, the model predicts an initial widening and potentially a subsequent narrowing of the SES-health gradient, as low SES individuals increase their health investment and improve their health-related behavior faster as a result of their worse health. Results from earlier studies (Ehrlich and Chuma, 1990; Ehrlich, 2000; Galama et al. 2009) suggest that the more rapidly worsening health of low SES individuals could lead to early withdrawal from the labor force, potentially widening the gradient in early and mid age, and shorter life spans, potentially narrowing the gradient in late age. Our model thus holds promise in explaining empirical health patterns. Such a model has not been available before and economists have highlighted the significance of its development (e.g., Cutler et al. 2010; Case and Deaton, 2005).

The paper is organized as follows. Section 2 briefly reviews the literature on health disparities by SES to determine the essential components required in a theoretical framework. The relation between SES and health is complex and developing a theory requires simplification and a focus on the essential mechanisms relating SES and health. We highlight potential explanations for the SES health gradient that a) explain a large part of the gradient and b) are relatively straightforward to include in our theoretical framework. Based on these findings we develop our theoretical formulation. Section 3 presents and discusses first-order conditions and the characteristics of the model solutions for a fixed age of retirement and a fixed age of death. The section also highlights potential mechanisms through which SES and health influence each other. In section 4 we summarize and conclude.

2 Components of a model capturing the SES-health gradient

In this section we review the literature to determine the essential components of a theory of health disparities by SES. Based on these findings we extend and refine prior work and present our theoretical formulation.

2.1 Background

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 socioeconomic status (SES) groups. Progress has been made in recent years in characterizing the relationships between the various dimensions of SES and health over the life cycle and in understanding the relative importance and directions of causal pathways. Epidemiological research has used longitudinal studies to examine the role of behavioral, material, psychosocial and healthcare related pathways in explaining SES-health associations (House et al. 1990; 1994; Lynch et al. 1997; Marmot et al. 1997a; Lantz et al. 1998; Yen and Kaplan, 1999; van Oort et al. 2005; Skalicka et al. 2009). Economists have recently re-emphasized the importance of the reverse impact of health on SES through ability to work (Smith, 1999; 2004; 2007; Case and Deaton, 2005).

These studies suggest that education is the key dimension of SES for which there appears to be robust evidence of a substantial causal protective effect on health. Secondly, an important part of the health differences by financial indicators of SES can be explained by the fact that bad health impinges on the ability to work, thereby reducing income. Further, these studies highlight the importance of health behaviors (such as smoking, drinking and exercise), curative and preventive care, psychosocial and environmental risk factors, neighborhood social environment, acute and chronic psychosocial stress, social relationships and supports, sense of control, fetal and early childhood conditions, and physical, chemical, biological and psychosocial hazards and stressors at work.

Below we provide more detail on the potential role of the working environment and lifestyle factors and the role of various potential pathways between health and SES and vice versa.

Working environment and lifestyle factors: Low SES individuals more often perform risky, manual labor than high SES individuals, and their health deteriorates faster as a consequence (Marmot et al. 1997b; Schrijvers et al. 1998; Borg and Kristensen, 2000). 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. Cutler et al. (2010) present similar results using mortality as an indicator of health. Schrijvers et al. (1998) use Dutch cross-sectional data to study the impact of working conditions on the association between occupational class and self-reported health. Hazardous physical working conditions are more prevalent in lower occupational classes, and this explains a substantial part (for males up to 83 percent) of the association between health and occupational (social) class. Extensive research further suggests an important role of lifestyle factors, particularly smoking, in explaining SES disparities in health (Mackenbach et al. 2004; Khang et al. 2009). Fuchs (1986) even argues that in developed countries, it is personal lifestyles that cause the greatest variation in health. Using three different datasets from the U.K. and the U.S., Marmot et al. (1997a) find that features of the psycho-social working environment, social circumstances outside work, and health behavior jointly account for much of the social gradient in health (see also House et al. 1994). Some epidemiological studies estimate that around two thirds of the social gradient in health deterioration could be explained by working environment and life style factors alone (Borg and Kristensen, 2000).

A multitude of potential pathways between health and SES and vice versa: As Cutler et al. (2010) note, the mechanisms 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.

- *Education on health:* Education is found to have a causal effect on health and mortality (Lleras-Muney, 2005; Oreopoulos, 2006; Smith, 2007; Silles, 2009). However, Cutler et al. (2010) note that the mechanisms by which education affects health are not well understood. While consumption behavior and curative and preventive care can partly explain the effect of education on health, it remains largely unclear why more educated individuals behave in a healthier manner (Cutler et al. 2010). Education increases earnings (e.g., Mincer, 1974) and thereby enables the purchase of health investments (though higher earnings may also increase the opportunity cost of time). Education potentially increases the efficiency of curative and preventive care usage and time inputs into the production of health investment (Grossman, 1972a; 1972b). It appears that the higher educated are better able at managing their diseases (Goldman and Smith, 2002), and high SES individuals appear to benefit more from new knowledge and new technology (Lleras-Muney and Lichtenberg, 2005; Glied and Lleras-Muney, 2008).

- *Health on education:* The existence of an effect of early childhood health on educational attainment has been established in studies from developed as well as developing countries. Studies from the U.S., U.K., and Norway show convincingly that low birth weight individuals have worse schooling outcomes (Behrman and Rosenzweig, 2004; Case et al. 2005; Black et al. 2007; Royer, 2009). Another piece of evidence is derived from the 1918 influenza epidemic in the U.S., and the hookworm eradication from the American South, where adverse conditions in childhood caused a lower educational attainment of the affected cohorts (Almond, 2006; Bleakley, 2007). From developing countries similar evidence is presented by, e.g., Miguel and Kremer (2004).

- *Income or wealth on health:* Income and wealth enable purchases of curative and preventive care and thereby potentially allow for better health maintenance. The impact of financial resources on health is likely to depend on the manner of health care provision in a country. In the case of market provision, income, wealth and employment may determine access to health care, whereas in the case of universal health care provision these factors may be less important. On the other hand, higher wages are associated with greater opportunity costs, which would reduce the amount of time devoted to health maintenance. Further, more affluent workers may choose safer working (associated with a lower level of job-related health stress) and living environments since safety is a normal good (Viscusi, 1978; 1993).

Smith (2007) finds no effect of financial measures of SES (income, wealth and change in wealth) on changes in health in the U.S. Financial indicators of SES do not seem to cause the onset of health problems at any age (Smith, 2007). Cutler et al. (2010) provide an overview of empirical findings and conclude that the evidence points to no or a very limited impact of income or wealth on health (see also Michaud and van Soest, 2008). Yet, this view is not unequivocally accepted. Replication is still needed and controversy remains on the extent to which these findings apply uniformly to different population segments. 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.

- *Health on income and wealth:* Healthy individuals are more productive, earn higher wages and are able to accrue greater wealth (Currie and Madrian, 1999; Contoyannis and Rice, 2001). Studies have shown that perhaps the most dominant causal relation between health and dimensions of SES is the causal impact that health has on one's ability to work and hence produce income and wealth (e.g., Smith, 2004; 2007; Case and Deaton, 2005).

- *Joint determination:* Fuchs (1982; 1986) (see also Barsky et al. 1997) has argued that the strong correlation between education and health may be due to differences in the time preferences of individuals, which affects investments in both education and health and helps to explain variations in cigarette smoking, diet, and exercise. Other third factors of interest that may produce a spurious correlation between SES and health are intelligence, cognitive ability, and non-cognitive skills (Auld and Sidhu, 2005; Deary, 2008; Chiteji, 2010). In a review of the literature on the relationship between 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 (see also Elo and Preston, 1996). But the authors also note that few studies have attempted to investigate the role of individual preferences, that preferences are difficult to measure, and that preferences with respect to health may differ from preferences with respect to finance, measures of which are usually employed in such studies.

2.2 Theoretical formulation

In this section we formalize 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 behavior.

A natural starting point for a theory of the relation between health and SES is a model of life cycle utility maximization. The model is based on the Grossman model of the demand for health (Grossman, 1972a; 1972b; 2000) in continuous time (see also Wolfe, 1985; Wagstaff, 1986; Ehrlich and Chuma, 1990; Zweifel and Breyer, 1997) with seven essential additional features.

First, we assume decreasing-returns-to-scale (DRTS) in the health production function (as in Ehrlich and Chuma, 1990).

Second, individuals choose their level of undesirable job characteristics which potentially have health consequences, denoted as "job-related health stress". The concept of job-related health stress can be interpreted broadly and can range from physical working conditions (e.g., hard or risky labor) to psychosocial aspects of work (e.g., low social status, lack of control, repetitive work, etc). The decision to engage in unhealthy labor is governed by the relative benefit of a possible wage premium – a compensating wage differential (Smith, 1776; Viscusi, 1978; 1979) – versus the cost in terms of a higher health deterioration rate. Evidence is strong that there is a wage premium for jobs with higher mortality risk (Smith, 1978), and also for less serious, non-fatal, health risks (e.g. Viscusi, 1978; Olson, 1981; Duncan and Holmlund, 1983). Thus we introduce the notion that individuals may accept risky and/or unhealthy work environments, in exchange for higher pay (Muurinen, 1982; Case and Deaton, 2005), and explore solutions in which the decision to rapidly "wear one's body down" (i.e., to perform "hard" labor or engage in work with psychosocial health risks) is endogenous.

Third, individuals engage in preventive care (such as check up doctor visits) to slow the biological aging rate. Hence, we explicitly model health investment as consisting of two components: (i) curative care (as in Grossman, 1972a; 1972b), and (ii) a new concept of preventive care. Fourth, consumption may affect the biological aging rate (Case and Deaton, 2005; see also Forster, 2001). We distinguish healthy consumption (such as the consumption of healthy foods, sports and exercise) from unhealthy consumption (such as smoking, excessive alcohol consumption).[2] Preventive care and healthy consumption are associated with health benefits in that they lower the biological aging rate. Healthy consumption also provides direct utility whereas preventive care is assumed to solely provide health benefits (similar to curative care, individuals demand preventive care solely for the health benefits it provides).[3] We interpret healthy consumption broadly to include decisions regarding housing and neighborhood.[4] Unhealthy consumption provides consumption benefits (utility) but increases the biological rate of aging.

Fifth, we include the decision to withdraw from the labor force (Galama et al. 2009).

Sixth, an essential component of the disparity in health by SES is the observed difference in mortality between SES groups. Further, length of life might be an important determinant of the age of retirement and the level of consumption and health investment over the life-course. Individuals

[2] It is useful to interpret the endogenous functions as bundles of goods and services (e.g., various consumption goods/services) or composite environmental factors (e.g., various physical and psychosocial health stresses).

[3] The distinction between healthy consumption and preventive care could in practice be difficult for some activities and could differ across individuals (e.g., some individuals exercise because they derive utility from it, whereas others solely exercise because it is healthy).

[4] Living in an affluent neighborhood is an expensive, yet health-promoting and utility-generating choice of individuals. However, the choice of neighborhood (housing) is a constrained choice: low SES individuals cannot afford to live in more affluent areas.

optimize length of life as in Ehrlich and Chuma (1990).[5]

Last, the causal 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 the level of experience of workers (e.g., Mincer, 1974).

With the exception of the above seven additional features, the discussion below follows the usual formulation (e.g., Grossman, 1972a; 1972b; 2000; Wagstaff, 1986; Zweifel and Breyer, 1997). Health is treated as a form of human capital (health capital) and individuals derive both consumption (health provides utility) and production benefits (health increases earnings) from it. Health is modeled as a stock that deteriorates over the life cycle and its deterioration can be counteracted by health investment. The demand for health investment (broadly interpreted as curative and/or preventive care) is a derived demand: individuals demand "good health", not the consumption of curative or preventive care.

Individuals maximize the life-time utility function

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

where T denotes the life span, β is a subjective discount factor and individuals derive utility $U(t) \equiv U[C_h(t), C_u(t), H(t)]$ from healthy consumption $C_h(t)$, unhealthy consumption $C_u(t)$ and from health $H(t)$. Time t is measured from the time an individual has completed her education and joined the labor force (e.g., around age 25 or so). Utility increases with healthy consumption $\partial U(t)/\partial C_h(t) \geq 0$, unhealthy consumption $\partial U(t)/\partial C_u(t) \geq 0$ and with health $\partial U(t)/\partial H(t) \geq 0$.

The objective function (1) is maximized subject to the following dynamic equations,

$$\dot{H}(t) = I_m(t)^{\alpha} - d(t)H(t), \tag{2}$$

$$\dot{A}(t) = \delta A(t) + Y(t) - p_{X_h}(t)X_h(t) - p_{X_u}(t)X_u(t) - p_m(t)m_m(t) - p_p(t)m_p(t), \tag{3}$$

the total time budget Ω,

$$\Omega = \tau_w(t) + \tau_{I_m}(t) + \tau_{I_p}(t) + \tau_{C_h}(t) + \tau_{C_u}(t) + s[H(t)], \tag{4}$$

and we have initial and end conditions: $H(0)$, $H(T)$, $A(0)$ and $A(T)$ are given.[6]

$\dot{H}(t)$ and $\dot{A}(t)$ in equations (2) and (3) denote time derivatives of health $H(t)$ and assets $A(t)$. Health (equation 2) can be improved through investment in curative (medical) care $I_m(t)$ and deteriorates at the biological aging rate $d(t) \equiv d[t, C_h(t), C_u(t), z(t), I_p(t); \xi(t)]$. The health

[5]However, to allow qualitative exploration of the characteristics of the solutions we treat mortality and retirement as exogenous (fixed) in this work.

[6]In Grossman's original formulation (Grossman, 1972a; 1972b) length of life T is determined by a minimum health level $H_{\min}$. If health falls below this level $H(t) \leq H_{min}$ an individual dies, hence $H(T) \equiv H_{\min}$.

production function $I_m(t)^\alpha$ is assumed to exhibit DRTS ($0 < \alpha < 1$).[7] The biological aging rate depends endogenously on healthy consumption $C_h(t)$, unhealthy consumption $C_u(t)$, job-related health stress $z(t)$, and investment in preventive care $I_p(t)$ and on a vector of exogenous functions $\xi(t)$. Consumption can be healthy ($\partial d(t)/\partial C_h(t) \leq 0$; e.g., healthy foods, healthy neighborhood) or unhealthy ($\partial d(t)/\partial C_u(t) > 0$; e.g., smoking). Preventive care is modeled analogous to curative care as an activity that provides no utility ($\partial U(t)/\partial I_p(t) = 0$) but is demanded for its health benefits ($\partial d(t)/\partial I_p(t) < 0$). Greater job-related health stress $z(t)$ accelerates the "aging" process ($\partial d(t)/\partial z(t) > 0$).

Assets $A(t)$ (equation 3) provide a return δ (the interest rate), increase with income $Y(t)$ and decrease with purchases in the market of healthy consumption goods $X_h(t)$, unhealthy consumption goods $X_u(t)$, curative care $m_m(t)$ and preventive care $m_p(t)$ at prices $p_{X_h}(t)$, $p_{X_u}(t)$, $p_m(t)$ and $p_p(t)$, respectively. Income $Y(t) \equiv Y[H(t), z(t); E, x(t)]$ is assumed to be an increasing function of health $H(t)$ ($\partial Y(t)/\partial H(t) > 0$) and an increasing function of job-related health stress $z(t)$ ($\partial Y(t)/\partial z(t) > 0$; Case and Deaton, 2005). Further, income depends exogenously on the consumer's stock of knowledge (an individual's human capital exclusive of health capital), usually assumed to be a function of years of schooling E and years of working experience $x(t)$ (e.g., Mincer, 1974). Last, we assume that individuals face no borrowing constraints.[8]

Goods and services $m_m(t)$ and $m_p(t)$ as well as own time inputs $\tau_{I_m}(t)$ and $\tau_{I_p}(t)$ are used in the production of curative care $I_m(t)$ and preventive care $I_p(t)$, respectively. Similarly, goods $X_h(t)$ and $X_u(t)$ purchased in the market and own time inputs $\tau_{C_h}(t)$ and $\tau_{C_u}(t)$ are used in the production of healthy and unhealthy consumption, $C_h(t)$ and $C_u(t)$, respectively.[9] The efficiencies of production $\mu_{I_m}(t;E)$, $\mu_{I_p}(t;E)$, $\mu_{C_h}(t;E)$ and $\mu_{C_u}(t;E)$ are 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 curative (medical) and preventive care (based on the interpretation of education as a productivity factor in own time inputs and in identifying and seeking effective care; Grossman, 1972a; 1972b),

$$I_m(t) \equiv I_m[m_m(t), \tau_{I_m}(t), \mu_{I_m}(t;E)], \tag{5}$$
$$I_p(t) \equiv I_p[m_p(t), \tau_{I_p}(t), \mu_{I_p}(t;E)], \tag{6}$$
$$C_h(t) \equiv C_h[X_h(t), \tau_{C_h}(t), \mu_{C_h}(t;E)], \tag{7}$$
$$C_u(t) \equiv C_u[X_u(t), \tau_{C_u}(t), \mu_{C_u}(t;E)]. \tag{8}$$

[7] Mathematically, this assumption is equivalent to assuming a linear process ($\alpha = 1$) and DRTS in the relation between the inputs of health investment goods / services $m_m(t)$ and own time $\tau_{I_m}(t)$ (as in Ehrlich and Chuma, 1990). Conceptually, however, there is an important distinction. In principle, one could imagine a scenario where the investment function $I_m(t)$ has constant or even increasing returns to scale in its inputs of health investment goods / services $m_m(t)$ and own time $\tau_{I_m}(t)$, but where the resulting health improvement (through the health production process) exhibits diminishing returns to scale in its inputs.

[8] Imperfect capital markets itself could be a cause of socioeconomic disparities in health if low income individuals face more borrowing constraints than higher income peers, and as such cannot optimally invest in their health.

[9] Because consumption consists of time inputs and purchases of goods/services in the market one can conceive leisure as a form of consumption consisting entirely or mostly of time inputs. Leisure, similar to consumption, provides utility and its cost consists of the price of goods/services utilized and the opportunity cost of time.

Further, we implicitly assume that curative care $I_m(t)$, preventive care $I_p(t)$ and job-related health stress $z(t)$ are non negative. We do so by assuming DRTS of the health production function in investment in curative care (see equation 2) and diminishing marginal benefits for job-related health stress and for investment in preventive care. The notion here is that one cannot "sell" ones health through negative curative care (see Galama and Kapteyn, 2009) or negative preventive care nor can one "buy" health through negative job-related health stress.

The total time available in any period Ω is the sum of all possible uses $\tau_w(t)$ (work), $\tau_{I_m}(t)$ (curative care), $\tau_{I_p}(t)$ (preventive care), $\tau_{C_h}(t)$ (healthy consumption), $\tau_{C_u}(t)$ (unhealthy consumption) and $s[H(t)]$ (sick time). The resulting time budget constraint is shown in equation (4).

We follow Grossman (1972a; 1972b; 2000) and assume that income $Y(t)$ is equal to the wage rate $w(t)$ times the amount of time spent working $\tau_w(t)$,

$$Y(t) = w(t)\left\{\Omega - \tau_{I_m}(t) - \tau_{I_p}(t) - \tau_{C_h}(t) - \tau_{C_u}(t) - s[H(t)]\right\}. \tag{9}$$

After the age of retirement R we have $\tau_w(t) = 0$ and $Y(t) = b(t)$, where $b(t)$ is a pension benefit function (potentially accrued over time as in Galama et al. 2009).

The wage rate $w(t) \equiv w[t, z(t); E, x(t)]$ is a function of job-related health stress $z(t)$

$$w(t) = w_*(t)[1 + z(t)]^{\gamma_w}, \tag{10}$$

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 $z(t) = 0$.[10] The stressless wage rate $w_*(t)$ is a function of the consumer's education E and experience $x(t)$ (e.g., Mincer, 1974),

$$w_*(t) = w_E e^{\rho_E E + \beta_x x(t) - \beta_{x^2} x(t)^2}, \tag{11}$$

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 constants, assumed to be positive.

Thus, we have the following optimal control problem: the objective function (1) is maximized with respect to the control functions $X_h(t)$, $\tau_{C_h}(t)$, $X_u(t)$, $\tau_{C_u}(t)$, $m_m(t)$, $\tau_{I_m}(t)$, $m_p(t)$, $\tau_{I_p}(t)$ and $z(t)$ and subject to the constraints (2, 3 and 4). The Hamiltonian (see, e.g., Seierstad and Sydsaeter, 1977; 1987) of this problem is:

$$\Im \quad = \quad U(t)e^{-\beta t} + q_H(t)\dot{H}(t) + q_A(t)\dot{A}(t), \tag{12}$$

where $q_H(t)$ is the adjoint variable associated with the differential equation (2) for health $H(t)$ and $q_A(t)$ is the adjoint variable associated with the differential equation (3) for assets $A(t)$.

[10] Our model concerns individuals who participate in the labor force. Given that our frame of reference is the labor force 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-related health stress we employ the simple relationship shown in equation (10).

The conditions for the optimal retirement age R and the optimal length of life T are for the Hamiltonian $\Im$ to equal zero at these ages

$$\Im(R) = 0, \quad (13)$$
$$\Im(T) = 0. \quad (14)$$

3 Solutions

In this section we discuss the first-order conditions for optimization (section 3.1), the characteristics of the solutions (section 3.2), the effect of SES on behavior (section 3.3), and the effect of health on behavior (section 3.4). Throughout the discussion we assume that an interior solution to the optimization problem exists.

3.1 First-order conditions

The first-order condition for maximization of (12) with respect to the control function health investment is

$$\frac{\partial U(t)}{\partial H(t)} = q_A(0)\left[\sigma_H(t) - \varphi_H(t)\right] e^{(\beta-\delta)t}, \quad (15)$$

where the Lagrange multiplier $q_A(0)$ is the shadow price of life-time wealth (see, e.g, Case and Deaton, 2005), $\sigma_H(t) \equiv \sigma_H[t, I_m(t), C_h(t), C_u(t), z(t), I_p(t); E, x(t), \xi(t)]$ is the user cost of health capital at the margin

$$\sigma_H(t) \equiv \pi_{I_m}(t)\left[d(t) + \delta - \widetilde{\pi_{I_m}}(t)\right], \quad (16)$$

$\pi_{I_m}(t) \equiv \pi_{I_m}[t, I_m(t), z(t); E, x(t)]$ is the marginal monetary cost of curative care $I_m(t)$

$$\pi_{I_m}(t) \equiv \frac{p_m(t) I_m(t)^{1-\alpha}}{\alpha[\partial I_m(t)/\partial m_m(t)]} = \frac{w(t) I_m(t)^{1-\alpha}}{\alpha[\partial I_m(t)/\partial \tau_{I_m}(t)]}, \quad (17)$$

and $\widetilde{\pi_{I_m}}(t) = \pi_{I_m}(t)^{-1}\,(\partial \pi_{I_m}(t)/\partial t)$.[11] The marginal monetary cost of curative care (equation 17) is a function of the price of medical goods and services purchased in the market $p_m(t)$ and the opportunity cost of time $w(t)$ (hence monetary). Note that the marginal monetary cost of investment in curative care $\pi_{I_m}(t)$ increases with the level of investment in curative care $I_m(t)$ due to decreasing-returns-to-scale of the health production function $I_m(t)^\alpha$ ($0 < \alpha < 1$; see equation 2). Further, $\varphi_H(t) \equiv \varphi_H[t, H(t), z(t); E, x(t)]$ is the marginal production benefit of health

$$\varphi_H(t) \equiv \frac{\partial Y(t)}{\partial H(t)}, \quad (18)$$

[11] In the remainder of this paper the symbol $\sim$ is used to denote the relative time derivative of a function: $\widetilde{f(t)} \equiv \frac{\partial f(t)}{\partial t} f(t)^{-1}$.

reflecting the notion that health increases earnings $Y(t)$.

The first-order condition for maximization of (12) with respect to the control function healthy consumption is

$$\frac{\partial U(t)}{\partial C_h(t)} = q_A(0)\left[\pi_{C_h}(t) - \varphi_{dC_h}(t)\right] e^{(\beta-\delta)t}, \tag{19}$$

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 $C_h(t)$

$$\pi_{C_h}(t) \equiv \frac{p_{X_h}(t)}{\partial C_h(t)/\partial X_h(t)} = \frac{w(t)}{\partial C_h(t)/\partial \tau_{C_h}(t)}, \tag{20}$$

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

$$\varphi_{dC_h}(t) \equiv -\pi_{I_m}(t)\frac{\partial d(t)}{\partial C_h(t)}H(t). \tag{21}$$

The marginal monetary cost of healthy consumption $\pi_{C_h}(t)$ (equation 20) 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 $\varphi_{dC_h}(t)$ (equation 21), is equal to the product of the marginal monetary cost of investment in curative care $\pi_{I_m}(t)$ and the "amount" of health saved $[\partial d(t)/\partial C_h(t)]H(t)$, and represents the marginal monetary value of health saved.[12] Similarly, the first-order condition for maximization of (12) with respect to the control function unhealthy consumption is

$$\frac{\partial U(t)}{\partial C_u(t)} = q_A(0)\left[\pi_{C_u}(t) + \pi_{dC_u}(t)\right] e^{(\beta-\delta)t}, \tag{22}$$

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)

$$\pi_{C_u}(t) \equiv \frac{p_{X_u}(t)}{\partial C_u(t)/\partial X_u(t)} = \frac{w(t)}{\partial C_u(t)/\partial \tau_{C_u}(t)}, \tag{23}$$

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

$$\pi_{dC_u}(t) \equiv \pi_{I_m}(t)\frac{\partial d(t)}{\partial C_u(t)}H(t). \tag{24}$$

[12]The marginal health benefit can be understood intuitively as the reduced need for health investment because of a lower health deterioration rate. While the health benefit is expressed in terms of the marginal cost of curative care, this is essentially arbitrary, as the monetary value of health saved could also be expressed in terms of the reduced need for other types of health investments such as preventive care or healthy consumption.

The first-order condition for maximization of (12) with respect to the control function job-related health stress is

$$\pi_{dz}(t) \quad = \quad \varphi_z(t), \tag{25}$$

where $\pi_{dz}(t) \equiv \pi_{dz}[t, H(t), I_m(t), C_h(t), C_u(t), z(t), I_p(t); E, x(t), \xi(t)]$ is the marginal health cost of job-related health stress (marginal monetary value of health lost)

$$\pi_{dz}(t) \equiv \pi_{I_m}(t)\frac{\partial d(t)}{\partial z(t)}H(t), \tag{26}$$

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

$$\varphi_z(t) \equiv \frac{\partial Y(t)}{\partial z(t)}, \tag{27}$$

reflecting the notion that job-related health stress is associated with a compensating wage differential (greater earnings).

Lastly, the first-order condition for maximization of (12) with respect to the control function preventive care is

$$\pi_{I_p}(t) \quad = \quad \varphi_{dI_p}(t), \tag{28}$$

where $\pi_{I_p}(t) \equiv \pi_{I_p}[t, z(t), I_p(t); E, x(t)]$ is the marginal monetary cost of preventive care $I_p(t)$

$$\pi_{I_p}(t) \equiv \frac{p_p(t)}{\partial I_p(t)/\partial m_p(t)} = \frac{w(t)}{\partial I_p(t)/\partial \tau_{I_p}(t)}, \tag{29}$$

and $\varphi_{dI_p}(t) \equiv \varphi_{dI_p}[t, H(t), I_m(t), C_h(t), C_u(t), z(t), I_p(t); E, x(t), \xi(t)]$ is the marginal health benefit of preventive care (marginal monetary value of health saved)

$$\varphi_{dI_p}(t) \equiv -\pi_{I_m}(t)\frac{\partial d(t)}{\partial I_p(t)}H(t). \tag{30}$$

The five first-order equations (15, 19, 22, 25 and 28) and the transversality conditions (13) and (14) define the dynamics of the problem we are interested in. Solving the first-order equations provides solutions for the time paths of the control functions $I_m(t)$, $C_h(t)$, $C_u(t)$, $z(t)$ and $I_p(t)$. The state functions health $H(t)$ and assets $A(t)$ can subsequently be obtained through the dynamic equations (2) and (3). Lastly, the optimal retirement age R and the optimal length of life T follow from the transversality conditions (13) and (14).

We have thus arrived at a life cycle model that incorporates labor force participation, healthy and unhealthy consumption (including housing, neighborhood social environment), health, curative (medical) and preventive care, job-related physical and psychosocial health stresses, wealth and mortality.

The Grossman model (Grossman, 1972a; 1972b) is a special case of our model and is defined by the first-order equations (15) and (19) for an exogenous biological aging rate $d(t)$. The first-order conditions (19), (22) and (25) are similar (but not identical) to those presented by Case and Deaton (2005). Ehrlich and Chuma (1990) have extended the Grossman model with the transversality condition (14) for optimal length of life T. The inclusion of endogenous retirement follows Galama et al. (2009).

3.2 Characteristics of the solutions

In the remainder of this paper we qualitatively explore the properties of the solutions for health $H(t)$, investment in curative care $I_m(t)$, investment in preventive care $I_p(t)$, healthy consumption $C_h(t)$ and unhealthy consumption $C_u(t)$ and job-related health stress $z(t)$. We do so by assessing the effects of parameter changes on the endogenous functions of interest, utilizing stylized representations (graphs) of the first-order conditions. However, stylized representations are less useful in assessing the nature of the transversality conditions for retirement R (13) and length of life T (14); this requires numerical approaches to solving the model. Thus, in practice, we explore the characteristics of the model conditional on retirement age R and length of life T (i.e., for fixed R and T).[13]

3.2.1 Assumptions

In the remainder, we assume:

1. Diminishing returns to scale (DRTS) in the health production function $I_m(t)^\alpha$ $(0 < \alpha < 1)$,

2. Diminishing marginal utilities of healthy $C_h(t)$ and unhealthy consumption $C_u(t)$ and of health $H(t)$,

$$\frac{\partial^2 U(t)}{\partial C_h(t)^2} < 0, \quad \frac{\partial^2 U(t)}{\partial C_u(t)^2} < 0, \quad \frac{\partial^2 U(t)}{\partial H(t)^2} < 0;$$

3. Diminishing marginal production benefit of health $\varphi_H(t)$, diminishing marginal production benefit of job-related health stress $\varphi_z(t)$, diminishing marginal health benefit of healthy consumption $\varphi_{dC_h}(t)$, and diminishing marginal health benefit of investment in preventive

[13] Treating retirement R and length of life T as exogenous (fixed) does not significantly affect our qualitative results regarding the formation of the SES health gradient (discussed in this work). Optimizing the age of retirement R and length of life T affects the overall level of health investment and consumption over the life cycle, as the transversality conditions (13) and (14) in combination with the initial and end conditions ($A(0)$, $A(T)$, $H(0)$ and $H(T)$), determine the parameters $q_A(0)$ and $q_H(0)$ in equations (31) and (32), but have limited effect on the direction of changes in the level of health investment and consumption.

care $\varphi_{dI_p}(t)$

$$\begin{aligned}
\frac{\partial \varphi_H(t)}{\partial H(t)} &= \frac{\partial^2 Y(t)}{\partial H(t)^2} < 0,\\
\frac{\partial \varphi_z(t)}{\partial z(t)} &= \frac{\partial^2 Y(t)}{\partial z(t)^2} < 0,\\
\frac{\partial \varphi_{dC_h}(t)}{\partial C_h(t)} &= -\pi_{I_m}(t)\frac{\partial^2 d(t)}{\partial C_h(t)^2} H(t) < 0,\\
\frac{\partial \varphi_{dI_p}(t)}{\partial I_p(t)} &= -\pi_{I_m}(t)\frac{\partial^2 d(t)}{\partial I_p(t)^2} H(t) < 0;
\end{aligned}$$

4. Constant returns to scale (CRTS) in the marginal health cost of unhealthy consumption $\pi_{dC_u}(t)$ and in the marginal health cost of job-related health stress $\pi_{dz}(t)$[14]

$$\begin{aligned}
\frac{\partial \pi_{dC_u}(t)}{\partial C_u(t)} &= \pi_{I_m}(t)\frac{\partial^2 d(t)}{\partial C_u(t)^2} H(t) = 0,\\
\frac{\partial \pi_{dz}(t)}{\partial z(t)} &= \frac{\partial}{\partial z(t)}\left[\pi_{I_m}(t)\frac{\partial d(t)}{\partial z(t)}\right] H(t) = 0;
\end{aligned}$$

5. CRTS in the inputs (goods/services purchased in the market and own-time) for investment in curative care $I_m(t)$, healthy consumption $C_h(t)$, unhealthy consumption $C_u(t)$ and preventive care $I_p(t)$.[15] As a result we have (see equations 17, 20, 23 and 29):

$$\pi_{I_m}(t) \propto I_m(t)^{1-\alpha}, \quad \frac{\partial \pi_{C_h}(t)}{\partial C_h(t)} = 0, \quad \frac{\partial \pi_{C_u}(t)}{\partial C_u(t)} = 0, \quad \frac{\partial \pi_{I_p}(t)}{\partial I_p(t)} = 0;$$

[14]While it seems plausible that the health benefits of investment in curative care, healthy consumption and investment in preventive care exhibit diminishing returns to scale, it is unclear whether the health costs of unhealthy consumption and job-related health stress exhibit decreasing or increasing returns to scale. For example, Forster (2001) assumes decreasing returns to scale for healthy consumption and increasing returns to scale for unhealthy consumption. In simple terms: escalating risky behavior (e.g., illicit drug use) or more hours of dangerous work can lead to rapid health deterioration, whereas after a certain point more investment in curative or preventive care, more exercise or more consumption of healthy foods does not prevent eventual aging. Since it is unclear a priori whether the effect of unhealthy consumption and the effect of job-related health stress on the biological aging rate $d(t)$ exhibits in- or decreasing returns to scale, we assume CRTS for simplicity.

[15]A priori it is not clear whether the relationships between the inputs (good/services and own time) and investment exhibit decreasing- or increasing-returns-to-scale. Hence we assume CRTS in these relations for simplicity.

6. Complementarity in utility of consumption $C_h(t)$, $C_u(t)$ and health $H(t)$[16]

$$\frac{\partial^2 U(t)}{\partial C_h(t)\partial H(t)} > 0, \quad \frac{\partial^2 U(t)}{\partial C_u(t)\partial H(t)} > 0;$$

7. Substitutability in utility of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption[17]

$$\frac{\partial^2 U(t)}{\partial C_h(t)\partial C_u(t)} < 0;$$

Assumptions 1 through 5 ensure that solutions to the optimal control problem exist.[18] The remaining assumptions are made to illustrate the potential of the model to describe a wide range of behaviors.

3.2.2 Stylized representations

Figures 1 and 2 provide stylized representations of the first-order conditions for health investment $I_m(t)$ (equation 15), healthy consumption $C_h(t)$ (equation 19), unhealthy consumption $C_u(t)$ (equation 22), job-related health stress $z(t)$ (equation 25) and investment in preventive care $I_p(t)$ (equation 28). In Section 3.3 we consider individuals a and b who differ in one particular SES indicator, but are otherwise identical, and in Section 3.4 we consider individuals a and c who differ in their health, but are otherwise identical. Figures 1 and 2 therefore show each of the five first-order conditions for individuals a and b or c. In this section we do not yet vary SES or health indicators and focus on the curves labeled with a.

Investment in curative care: The solution for investment in curative care $I_m(t)$ is determined by the first-order condition for health investment (15), conditional on the level of the health stock $H(t)$.[19] The evolution of the health stock $H(t)$ then follows from the initial condition $H(0)$ and the health investment $I_m(s)$ and biological aging $d(s)$ histories ($s < t$) through the dynamic equation (2).

[16]Indeed, Finkelstein et al. (2008) find evidence that the marginal utility of consumption declines as health deteriorates. This would rule out strongly separable functional forms for the utility function where the marginal utility of consumption is independent of health and forms where the marginal utility of consumption would decrease in health.

[17]The substitutability in utility of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption allows us to model substitution from unhealthy to healthy consumption (or vice versa).

[18]Optimal solutions for the state functions $A(t)$, $H(t)$ and the control functions $X_h(t)$, $\tau_{C_h}(t)$, $X_u(t)$, $\tau_{C_u}(t)$, $m_m(t)$, $\tau_{I_m}(t)$, $m_p(t)$, $\tau_{I_p}(t)$ and $z(t)$ exist if the Hamiltonian $\Im$ (see equations 2, 3 and 12) is concave in each of the state and control functions and differentiable w.r.t. the state and control functions (see, e.g., Seierstad and Sydsaeter, 1977; 1987).

[19]Note that the first-order condition (15) is interpreted in the health production literature spawned by Grossman (1972a; 1972b) as the condition for optimal health, and not as the condition for optimal health investment. However, this condition was derived by optimizing the optimal control problem with respect to health investment (it follows directly from relations 32, 37 and 38) and hence an alternative interpretation is that it determines the optimal level of

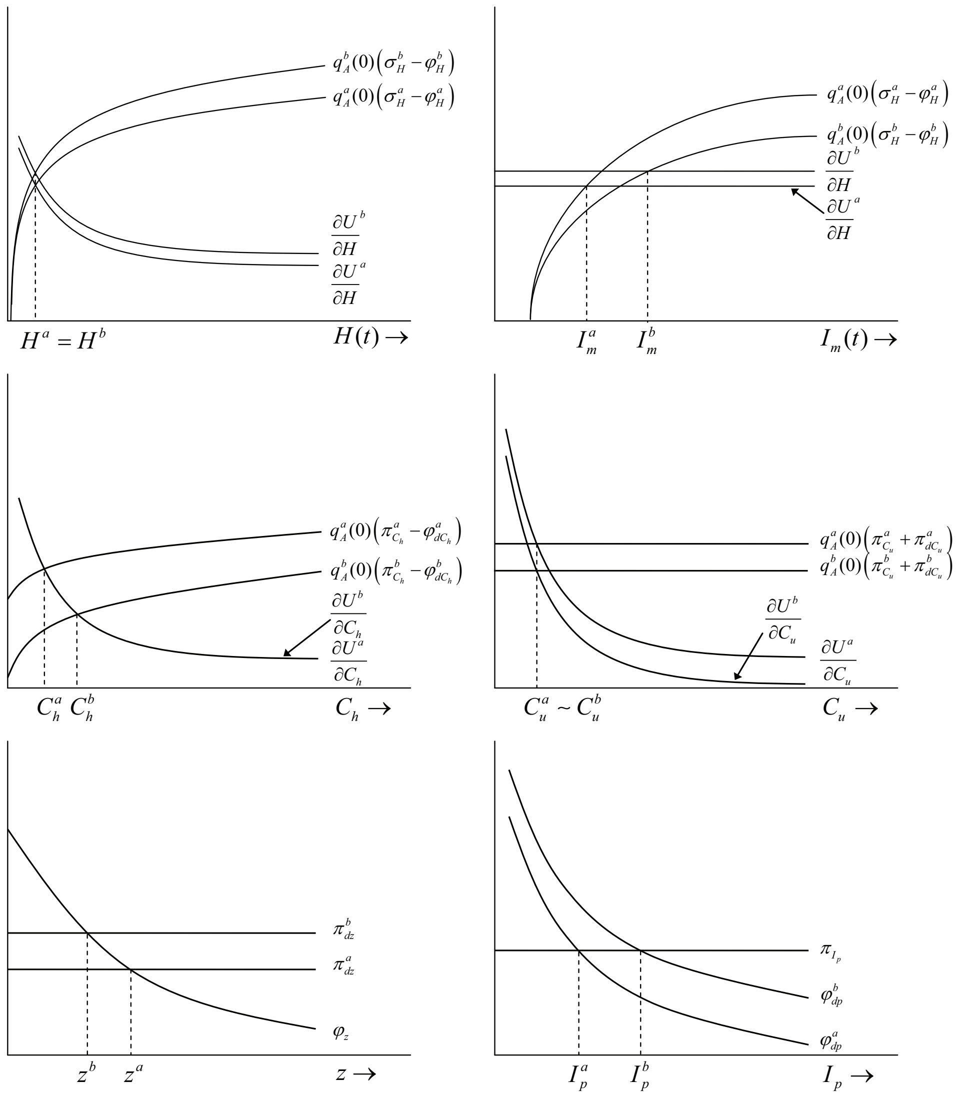

Figure 1: **Differences in SES:** *Marginal consumption $\partial U/\partial H$ and marginal production benefit φ_H of health versus the user cost of health capital at the margin σ_H as a function of health (top left) and as a function of health investment (top right). Marginal utility of healthy consumption $\partial U/\partial C_h$ versus the marginal monetary cost π_{C_h} and the health benefit φ_{dC_h} of healthy consumption $C_h(t)$ (center left); Marginal utility of unhealthy consumption $\partial U/\partial C_u$ versus the marginal monetary cost π_{C_u} and the marginal health cost φ_{dC_u} of unhealthy consumption $C_u(t)$ (center right); Marginal health cost π_{dz} versus the marginal production benefit φ_z of job-related health stress $z(t)$ (bottom-left); Marginal monetary cost π_{I_p} versus the marginal health benefit φ_{dI_p} of investment in preventive care $I_p(t)$ (bottom-right). In labeling the curves we have omitted the time varying term with exponent $(\beta - \delta)t$.*

The first-order condition for health investment (15) equates the consumption benefit of health $\partial U(t)/\partial H(t)$ with the cost of maintaining the health stock $q_A(0)[\sigma_H(t) - \varphi_H(t)]e^{(\beta-\delta)t}$. Figure 2 shows a simple stylized representation of this relation as a function of health $H(t)$ (top left-hand panel) and as a function of investment in curative care $I_m(t)$ (top right-hand panel).

Consider the top left-hand panel and individual a first. The marginal utility of health (labeled $\partial U^a/\partial H$) is diminishing in health (assumption 2). The user cost of health capital $\sigma_H(t) = \pi_{I_m}(t)[d(t) + \delta - \widetilde{\pi_{I_m}}(t)]$ is independent of health and the marginal production benefit of health $\varphi_H(t) = \partial Y(t)/\partial H(t)$ (increased earnings) exhibits DRTS in health (assumption 3). The resulting curve (labeled $q_A(0)(\sigma^a_H - \varphi^a_H)$) is upward sloping in health. Since health is a state function its level is given and provides a constraint: the two curves have to intersect at H^a.

Now consider the top right-hand panel of Figure 2. The marginal monetary cost of curative care $\pi_{I_m}(t)$ and hence the user cost of health capital $\sigma_H(t)$ is increasing in the level of curative care $I_m(t)$ ($\pi_{I_m}(t) \propto I_m(t)^{1-\alpha}$; see equations 16 and 17; assumptions 1 and 5). The marginal production benefit of health $\varphi_H(t)$ (see equations 9 and 18) is independent of the level of investment in curative care $I_m(t)$. The resulting curve is upward sloping (labeled $q_A(0)(\sigma^a_H - \varphi^a_H)$). Further, the marginal utility of health $\partial U(t)/\partial H(t)$ is independent of the level of investment in curative care $I_m(t)$ (horizontal line labeled $\partial U^a/\partial H$). Its level is determined by the level of the health stock H^a (draw a horizontal line from the top left-hand to the top right-hand panel of Figure 2). The intersection of the two curves determines the optimal level of investment in curative care I^a_m.

The top-right hand panel of Figure 2 further illustrates that the level of investment in curative care $I_m(t)$ increases with the consumption benefit of health $\partial U(t)/\partial H(t)$, the production benefit of health $\varphi_H(t)$, and with wealth (lower $q_A(0)$), and decreases with the user cost of health capital $\sigma_H(t)$. Further, the level of health investment $I_m(t)$ is a function of the health stock $H(t)$ (more details are provided in section 3.4).

Healthy and unhealthy consumption: The center-left panel of Figure 2 shows the first-order condition for healthy consumption $C_h(t)$ (equation 19) which equates the marginal utility of healthy consumption (solid line labeled $\partial U^a/\partial C_h$) to the net marginal cost of healthy consumption (solid line labeled $q_A(0)(\pi^a_{C_h} - \varphi^a_{dC_h})$). The marginal utility of healthy consumption is diminishing in the level of consumption (assumption 2). The net marginal cost of healthy consumption increases with the marginal monetary cost of healthy consumption $\pi_{C_h}(t)$ (equation 20; CRTS [assumption 5]) and decreases with the marginal health benefit $\varphi_{dC_h}(t) = -\pi_{I_m}(t)[\partial d(t)/\partial C_h(t)]H(t)$ (DRTS [assumption 3]). Hence, the net marginal cost of healthy consumption is upward sloping. The

the control function health investment $I_m(t)$. Health $H(t)$ is a state function and is determined by the dynamic relation (2). At a given time t an individual cannot decide about its level (hence conditional). This seemingly subtle difference in interpretation of the first-order condition (together with the assumption of DRTS in the health production function) addresses a number of issues with the health production literature and allows us to accommodate a wider range of health behaviors than existing health production models (see Galama, 2010). Importantly, the first-order condition (15) is of a simpler form than the condition used in the health production literature, allowing us to develop a better understanding of the characteristics of the optimal solution for health investment.

point of intersection defines the optimal solution for healthy consumption C_h^a (vertical dashed line).

The center-right panel of Figure 2 shows the first-order condition for unhealthy consumption $C_u(t)$ (equation 22). The first-order condition is similar to the condition for healthy consumption described in the preceding paragraph. The difference lies in the marginal health cost (rather than health benefit) of unhealthy consumption $\pi_{dC_u}(t) = \pi_{I_m}(t)[\partial d(t)/\partial C_u(t)]H(t)$ (CRTS [assumption 4]) which has to be added rather than subtracted from the marginal monetary cost of unhealthy consumption $\pi_{C_u}(t)$ (equation 23; CRTS [assumption 5]). The net marginal cost of unhealthy consumption is represented by the solid horizontal line labeled $q_A(0)(\pi_{C_u}^a + \pi_{dC_u}^a)$. The point of intersection defines the optimal solution for unhealthy consumption C_u^a (vertical dashed line).

The level of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption increases with the marginal utility of consumption ($\partial U(t)/\partial C_h(t)$ and $\partial U(t)/\partial C_u(t)$), increases with wealth (lower $q_A(0)$), decreases with the marginal monetary costs of consumption ($\pi_{C_h}(t)$ and $\pi_{C_u}(t)$), increases with the marginal health benefit of healthy consumption $\varphi_{dC_h}(t)$, and decreases with the marginal health cost of unhealthy consumption $\pi_{dC_u}(t)$.

Job-related health stress: The bottom-left panel of Figure 2 shows the first-order condition for job-related health stress $z(t)$ (equation 25) which equates the marginal production benefit $\varphi_z(t) = \partial Y(t)/\partial z(t)$ (increased earnings; DRTS [assumption 3]) to the marginal health cost of job-related health stress $\pi_{dz}(t) = \pi_{I_m}(t)[\partial d(t)/\partial z(t)]H(t)$ (CRTS [assumption 4]). The optimal solution for job-related health stress z^a is indicated by the vertical dashed line. The optimal level increases with the marginal production benefit $\varphi_z(t)$ and decreases with the marginal health cost $\pi_{dz}(t)$.

Investment in preventive care: The bottom-right panel of Figure 2 represents the first-order condition for investment in preventive care $I_p(t)$ (equation 28) which equates the marginal health benefit of investment in preventive care $\varphi_{dI_p}(t) = -\pi_{I_m}(t)[\partial d(t)/\partial I_p(t)]H(t)$ (DRTS [assumption 3]) to the marginal monetary cost $\pi_{I_p}(t)$ (equation 29; CRTS [assumption 5]). The optimal solution for investment in preventive care I_p^a is indicated by the vertical dashed line. The optimal level of investment in preventive care $I_p(t)$ increases with the marginal health benefit $\varphi_{dI_p}(t)$ and decreases with the marginal monetary cost $\pi_{I_p}(t)$ of investment in preventive care.

3.3 SES and its effect on behavior

In this section we explore the (cumulative) effect on health over the life cycle of choices made in curative care, in life style and in working environment. Our emphasis is on exploring differences in constraints (e.g., in wealth, skills, experience, education and prices).

Common measures of SES employed in empirical research are wealth, earnings (income) and education. In the following subsections we discuss the relations between wealth and health, earnings and health and education and health. We consider two individuals a and b who differ in one particular SES indicator, but are otherwise identical. Both individuals have the same initial level of health $H(t)$, are of the same age t, face the same environments (e.g., same interest rate δ), and have the same preferences (i.e., same utility function $U[C_h(t), C_u(t), H(t)]$ and same time preference β).[20] We are interested in the predictions of our model for the subsequent evolution of health for these two individuals, given a ceteris paribus change in one SES indicator.

3.3.1 Wealth and health: pure "asset" effect

Consider two individuals a and b who differ in life-time wealth $q_A(0)$. Individual b has greater life-time wealth, i.e., $q_A^b(0) < q_A^a(0)$, but is otherwise identical. Because of the similarities between the two individuals the difference in life-time wealth is to be interpreted as due to differences in endowed physical capital (e.g., assets $A(0)$).[21]

Investment in curative care: Figure 1 shows a stylized representation of the first-order condition for investment in curative care as a function of health $H(t)$ (top left-hand panel) and as a function of investment in curative care $I_m(t)$ (top right-hand panel). Consider the top right-hand panel first. As a result of greater endowed wealth ($q_A^b(0) < q_A^a(0)$) the net marginal cost of maintaining the health stock shifts downward (curve labeled $q_A^b(0)(\sigma_H^b - \varphi_H^b)$). As a result, the optimal level of investment in curative care is higher $I_m^b > I_m^a$.

An indirect effect operates through consumption. Greater endowed wealth allows individual b to consume more consumption goods and services (see discussion below for further detail). A higher level of consumption increases (or at a minimum leaves unchanged) the marginal utility of health $\partial U(t)/\partial H(t)$ (assumption 6). Thus the marginal utility of health shifts upward (or is unchanged) in both the top left-hand panel and in the top right-hand panel of Figure 1 (curves labeled $\partial U^b/\partial H$). This reinforces the effect on curative care, and wealthier individuals invest more in curative care than less affluent peers $I_m^b > I_m^a$.

Now turn to the top-left panel of Figure 1. Because the health stock of individuals a and b is the same $H^b = H^a$ the curves need to intersect at the same level of health. The net result is an upward shift in the marginal cost of maintaining the health stock (curve labeled $q_A^b(0)(\sigma_H^b - \varphi_H^b)$),

[20] Part of the SES health gradient may be explained by differences in individuals preferences. A lower rate of time preference β operates in a similar manner to wealth, earnings and education. However in contrast to SES, differences between low and high discounting individuals grow larger with time (the discount factor $e^{(\beta-\delta)t}$ grows slower with age t for an individual with a low discount rate). 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 (e.g., Fuchs, 1982; 1986).

[21] Endowments need not necessarily be available to the individual at age $t = 0$, but could also be received at later ages. Conceptually there is no distinction between early and late endowments: an endowment at later ages also lowers life-time wealth $q_A(0)$. Our model is deterministic and the individual knows with certainty about the timing and amount of physical assets she will receive.

as a result of the upward shift in the marginal utility of health $\partial U/\partial H$.[22,23]

Healthy and unhealthy consumption: The center-left panel of Figure 1 shows the shift in the level of healthy consumption $C_h(t)$. The product $q_A(0)\pi_{C_h}(t)$ shifts downward as a result of greater endowed physical capital ($q_A^b(0) < q_A^a(0)$) and because the marginal monetary cost of healthy consumption $\pi_{C_h}(t)$ (equation 20) is unchanged. Essentially, greater endowed assets enable more purchases of healthy consumption goods. Further, endowed assets increase the health benefit of healthy consumption, $\varphi_{dC_h}(t) = -\pi_{I_m}(t)[\partial d(t)/\partial C_h(t)]H(t)$ (for $H^b = H^a$ and $\pi_{I_m}^b > \pi_{I_m}^a$ [see earlier discussion in "Investment in curative care"]). The resulting net marginal cost of healthy consumption (solid line labeled $q_A^b(0)(\pi_{C_h}^b - \varphi_{dC_h}^b)$) is lower in level (the "wealth" effect) and steeper in slope (the "savings in care" effect; $\varphi_{dC_h}^b > \varphi_{dC_h}^a$). In the example the marginal utility of healthy consumption (curve labeled $\partial U^b/\partial C_h$; center-left panel of Figure 1) is shown as unchanged (i.e., we have assumed the level of unhealthy consumption $C_u(t)$ has not changed). The optimal solution for healthy consumption C_h^c (vertical dashed line) of an individual with greater endowed wealth is higher than that of a poorer individual $C_h^b > C_h^a$.

The center-right panel of Figure 1 shows the shift in the level of unhealthy consumption $C_u(t)$. As with healthy consumption, greater endowed wealth shifts the product of the shadow price of life-time wealth $q_A(0)$ (lowered) and the marginal monetary cost of unhealthy consumption $\pi_{C_u}(t)$ (unchanged; equation 23) downward. Unlike healthy consumption this shift is countered by an increase in the marginal health cost of unhealthy consumption $\pi_{dC_u}(t) = \pi_{I_m}(t)[\partial d(t)/\partial C_u(t)]H(t)$ (for $H^c = H^a$ and $\pi_{I_m}^b > \pi_{I_m}^a$; see earlier discussion in "Investment in curative care"). Greater endowed wealth allows purchasing more unhealthy consumption goods, but also increases the marginal health cost. Further, the marginal utility of unhealthy consumption $\partial U(t)/\partial C_u(t)$ shifts downward as a result of the higher level of healthy consumption $C_h(t)$ (assumption 7; see previous paragraph). The optimal level of unhealthy consumption C_u^b (vertical dashed line) is shown as unchanged $C_u^b \sim C_u^a$.[24]

Job-related health stress and investment in preventive care: The first-order conditions for job-related health stress $z(t)$ (equation 25; bottom-left corner of Figure 1) and for investment in preventive care (equation 28; bottom-right corner of Figure 1) do not depend on life-time wealth $q_A(0)$. However, there is an indirect effect of greater endowed wealth. Both the marginal health cost of job-related health stress $\pi_{dz}(t)$ (equation 26) and the marginal health benefit of preventive

[22]One can obtain the same result as follows. Since, $\partial U^b/\partial H \geq \partial U^a/\partial H$ we have $q_A^b(0)[\sigma_H^b(t) - \varphi_H^b(t)] \geq q_A^a(0)[\sigma_H^a(t) - \varphi_H^a(t)]$ (equation 15) and the net result is an upward shift of the curve labeled $q_A^b(0)(\sigma_H^b - \varphi_H^b)$.

[23]While the net marginal cost of maintaining the health stock (curve labeled $q_A^b(0)(\sigma_H^b - \varphi_H^b)$) shifts downward as a result of greater endowed wealth ($q_A^b(0) < q_A^a(0)$), it shifts upward due to a higher user cost of health capital $\sigma_H(t)$, since a higher optimal level of investment in curative care ($I_m^b > I_m^a$) increases the marginal monetary cost of curative care $\pi_{I_m}(t)$ ($\pi_{I_m}(t) \propto I_m(t)^{1-\alpha}$) (see equation 16 and assumption 5).

[24]Note that the marginal utility of healthy consumption $\partial U(t)/\partial C_h(t)$ is unchanged if the level of unhealthy consumption $C_u(t)$ is unchanged (as was assumed).

care $\varphi_{dI_p}(t)$ (equation 30) are proportional to the marginal monetary cost of investment in curative care $\pi_{I_m}(t)$. Higher endowed wealth (individual b) implies $\pi^b_{I_m}(t) > \pi^a_{I_m}(t)$ (see earlier discussion in "Investment in curative care"). Thus wealthier individuals have greater marginal health cost of job-related health stress $\pi_{dz}(t)$ and greater marginal health benefit of preventive care $\varphi_{dI_p}(t)$. Consequently the optimal level of job-related health stress is lower $z^b < z^a$ and the optimal level of investment in preventive care is higher $I^b_p > I^a_p$ for individuals with greater endowed wealth, compared to less-affluent peers.

3.3.2 Income and health: pure "wage" effect

Again, consider two individuals a and b but this time the difference is in their wage rate. Individual b has a higher "stressless" wage rate than individual a ($w^b_*(t) > w^a_*(t)$) and hence has a higher level of earnings over the life cycle $Y^b(t) > Y^a(t)$.[25] It is important to distinguish between an *evolutionary wage change* (differences along the wage path of an individual) and *differences in life-time wage profiles* (between individuals).

Evolutionary wage change: In our model of perfect certainty a change in wage does not affect the parameter $q_A(0)$ (life-time wealth) as the change is fully anticipated by the individual. Such a response is referred to as an evolutionary wage change (along an individual's wage profile). An evolutionary increase in the wage rate $w(t)$ increases the marginal production benefit of health $\varphi_H(t) = \partial Y(t)/\partial H(t)$ and of job-related health stress $\varphi_z(t) = \partial Y(t)/\partial z(t)$ (see equations 9 and 10).[26] It also increases the opportunity cost of time.[27] As a result, the various marginal costs and benefits of the functions of interest increase, and the net effect on the level of investment in curative care, healthy consumption, job-related health stress, and preventive care is ambiguous. An exception is the level of unhealthy consumption, which is lower since both the marginal monetary cost and the marginal health cost of unhealthy consumption increase with the wage rate. Thus, an evolutionary wage increase could be either good or bad for health.

However, consider the case where the marginal production benefit of health $\varphi_H(t)$ is small

[25]Earnings $Y(t)$ are a function of the wage rate $w(t)$ times the amount of time spent working $\tau_w(t)$ (see equation 9). A higher wage rate $w_*(t)$ implies that individual b has higher earnings $Y(t)$ than individual a because the direct effect of higher wages is to increase earnings (the wage rate multiplied by the time spent working). A secondary effect operates through time spent working, where individuals may work more because of the higher opportunity cost of not working (substitution effect). On the other hand individuals may work fewer hours to spend their increased income on care or consumption (income effect). Empirical studies suggest that the substitution and income effects are of the same magnitude (e.g., Blundell and MaCurdy, 1999) and hence that the direct effect of a wage increase is to increase earnings.

[26]In our formulation the marginal benefit of job-related health stress is increasing in the wage rate. Case and Deaton (2005) in their narrower definition of $z(t)$ as manual, risky labor (i.e., not including the psychosocial aspects of work), assume that the marginal benefit of additional manual labor is lower among those with higher wages.

[27]The wage rate might not be the most appropriate measure of the opportunity cost of time since time is not always mutually exclusive. Sick time is usually used in the production of curative care, and often institutional arrangements make it possible to continue earning wages while seeking curative care (De Serpa, 1971; Muurinen, 1982).

compared to the user cost of health capital $\sigma_H(t)$.[28] Since individuals a and b possess the same health stock, and $\sigma_H(t) - \varphi_H(t) \sim \sigma_H(t)$, it follows from the first-order condition for health investment (equation 15) that $\sigma_H(t)$ is unchanged and hence $\pi_{I_m}(t)$ is unchanged. Yet a higher wage rate increases the opportunity cost of time, and consequently the level of investment in curative care $I_m(t)$ is lower.[29] Further, the health benefit of healthy consumption $\varphi_{dC_h}(t)$ and of preventive care $\varphi_{dI_p}(t)$ (equations 21 and 30), and the health cost of unhealthy consumption $\pi_{dC_u}(t)$ and of job-related health stress $\pi_{dz}(t)$ (equations 24 and 26) are unchanged (because $\pi_{I_m}(t)$ is unchanged). The marginal monetary cost of healthy consumption $\pi_{C_h}(t)$ (equation 20) and of unhealthy consumption $\pi_{C_u}(t)$ (equation 23) however increase with the wage rate $w(t)$ (reflecting the higher opportunity cost of time) and the level of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption is lower. In addition, the marginal production benefit of job-related health stress $\varphi_z(t)$ increases with the wage rate (equation 27) as does the marginal monetary cost of investment in preventive care $\pi_{I_p}(t)$ (equation 29). As a result, the level of job-related health stress $z(t)$ is higher and the level of investment in preventive care $I_p(t)$ lower. Thus, on balance, if the production benefit of health is small, an evolutionary wage change is bad for health. An exception to this pattern is the level of unhealthy consumption, which is lower.

Differences in life-time wage profiles: Now consider again two individuals a and b. Individual b earns higher wages over the life cycle, i.e., person b has greater life-time wealth (and hence $q_A^b(0) < q_A^a(0)$). Thus the net result of higher earnings over the life cycle would be similar to the "pure" asset effect described in section 3.3.1, except that apart from the life-time wealth effect ($q_A^b(0) < q_A^a(0)$) there is also a competing effect of the greater opportunity cost of time (see discussion above).

3.3.3 Education and health: the additional "efficiency" effect

Consider two individuals a and b who differ in their level of education E. Individual b has obtained more education but is otherwise identical. As a result, individual b has a higher wage rate $w_*(t)$ (equation 11). Thus the effect of education is similar to the effect of higher earnings over the life cycle and the discussion presented in section 3.3.2 applies here as well.

But education potentially also improves the efficiencies $\mu(t;E)$ of investment in curative and preventive care, and to a lesser extent healthy and unhealthy consumption (equations 5 to 8).[30] The marginal cost of investment in curative care $\pi_{I_m}(t)$ is determined by the first-order condition for health investment (equation 15) and is unchanged. Since the marginal cost of investment in curative care $\pi_{I_m}(t)$ increases in the level $I_m(t)$ and decreases in the efficiency $\mu_{I_m}(t;E)$ of

[28]Note that it is always true that $\sigma_H(t) > \varphi_H(t)$, otherwise the investment in curative care would finance itself through negative net marginal costs of maintaining the health stock and individuals would achieve infinite health.

[29]This can be seen from equation (17): if $w(t)$ increases, $I_m(t)$ has to decrease to maintain $\pi_{I_m}(t)$ at the same level.

[30]Grossman (1972a; 1972b) assumes that the higher educated are more efficient producers and consumers of curative care. We extend his definition to preventive care. However, it is less clear whether the higher educated are more efficient producers and consumers of consumption goods and services.

investment in curative care (see equation 17), a higher efficiency due to education implies a higher level of investment in curative care compared to the pure "wage" effect described in section 3.3.2.

A higher efficiency of investment in preventive care $\mu_{I_p}(t;E)$ lowers the marginal monetary cost of preventive care $\pi_{I_p}(t)$ (equation 29) while the marginal benefit $\varphi_{dI_p}(t) \propto \pi_{I_m}(t)$ (equation 30) is unchanged. Thus the optimal level of investment in preventive care is higher compared with the pure "wage" effect.

If the efficiencies of healthy and unhealthy consumption do not (or only moderately) respond to education then the levels of healthy and unhealthy consumption are unchanged compared to the pure "wage" effect.

3.3.4 Summary and discussion – the effect of SES on behavior

The left column of Table 1 provides a brief overview of the effect of greater *endowed wealth* on behavior. The direct effect of endowed wealth (through $q_A(0)$) is to enable a higher level of investment in curative care $I_m(t)$, healthy consumption $C_h(t)$ and unhealthy consumption $C_u(t)$. In addition, associated with a higher level of investment is a higher marginal monetary cost of curative care $\pi_{I_m}(t)$ (assumption 1). As a result, individuals derive greater marginal health benefit from healthy consumption $\varphi_{dC_h}(t)$ and from preventive care $\varphi_{dI_p}(t)$ because of the greater monetary value represented by the amount of health saved. Similarly, the marginal health cost of unhealthy consumption $\pi_{dC_u}(t)$ and of job-related health stress $\pi_{dz}(t)$ are greater because of the greater monetary value represented by the amount of health lost.

Wealthier individuals invest more in curative $I_m(t)$ and preventive care $I_p(t)$ and their level of healthy consumption $C_h(t)$ is higher. Wealthy individuals also engage in work that is more conducive to health: jobs associated with lower levels of job-related health stress $z(t)$. Wealth protects health by encouraging healthy life styles and enabling individuals to work and live in healthy environments. The net effect is ambiguous only for the level of unhealthy consumption as the direct effect of endowed wealth is to enable a higher level of unhealthy consumption $C_u(t)$, whereas the indirect effect is an increase in the marginal health cost of unhealthy consumption $\pi_{dC_u}(t)$.

With regards to consumption, consider a situation where the severity of the health detriment $\partial d(t)/\partial C_u(t)$ resulting from unhealthy consumption is greater than in the example shown in the right-hand panel of Figure 1. In this case greater marginal health cost of unhealthy consumption $\pi_{dC_u}(t) \propto \partial d(t)/\partial C_u(t)$ shifts the net marginal cost of unhealthy consumption ($q^b_A(0)(\pi^b_{C_u} + \pi^b_{dC_u})$; center-right panel of Figure 1) upward. This lowers the level of unhealthy consumption C^b_u.[31] Because the marginal health cost of unhealthy consumption increases in the severity of the health detriment ($\pi_{dC_u}(t) \propto \partial d(t)/\partial C_u(t)$) we expect to observe a pattern in which wealthy individuals

[31] This shift is exacerbated due to substitutability in utility of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption (assumption 7), as the marginal utility of healthy consumption $\partial U(t)/\partial C_h(t)$ increases for a lower level of unhealthy consumption $C_u(t)$ and the marginal utility of unhealthy consumption $\partial U(t)/\partial C_u(t)$ decreases for a higher level of healthy consumption $C_h(t)$.

	Endowed wealth	Evolutionary wage change
$I_m(t)$	+	−
$C_h(t)$	+	−
$C_u(t)$	+/−	−
$z(t)$	−	+
$I_p(t)$	+	−

Table 1: *The effect of greater endowed wealth* $q_A^b(0) < q_A^a(0)$ *(left column) and of an evolutionary wage increase (right column) on behavior. The results for greater endowed wealth (left column) are also valid for the effect of greater earnings over the life cycle and for the effect of a higher level of education.*

consume more of moderately unhealthy consumption goods (e.g., moderate alcohol consumption) and less of more severe unhealthy consumption goods (e.g., cigarettes, high alcohol consumption, illicit drugs) when compared to less wealthy individuals.

Differences between individuals in *life-time earnings* (comparing different individuals with different life cycle wage profiles) operate similar to an increase in endowed wealth. Our model suggests that the health benefit of a "pure" asset endowment would be larger than the effect of a "comparable" change in life-time earnings (similar change in the shadow-price of life-time wealth $q_A(0)$) due to the competing effect of the increased opportunity cost of time. There are reasons to believe that the life-time wealth effect may dominate the effect of the opportunity cost of time (higher current wages). First, this is consistent with the result by Dustmann and Windmeijer (2000) and Contoyannis et al. (2004) that a transitory wage increase affects health negatively while a permanent wage change affects health positively. 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.

A higher level of *education* operates similar to greater earnings over the life cycle. But education has an independent effect on health, over and above generating greater life-time earnings and wealth, through enhancing the efficiency of curative and preventive care. This leads to a higher demand for both curative and preventive care.

The right-hand column of Table 1 summarizes the effect of an *evolutionary* increase in the wage rate if the marginal production benefit of health $\varphi_H(t)$ is small compared to the user cost of health capital $\sigma_H(t)$. On average, as a result of the increased opportunity cost of time and the greater marginal production benefit of job-related health stress, an evolutionary increase in the wage rate is bad for health.

If there is no complementarity in utility of consumption and health (assumption 6) the predictions would remain the same. If the relation were instead one of substitutability (worse health improves the utility of consumption), solutions are possible in which greater SES leads to less investment in curative care. This is generally not observed. If there were no substitutability in utility of healthy and unhealthy consumption (assumption 7) higher SES would generally not lead to a reduction in unhealthy consumption, except for extremely unhealthy consumption goods, but would still be associated with a shift toward healthy consumption (i.e., a smaller fraction of a larger budget is devoted to unhealthy consumption).

3.4 Health and its effect on behavior

In this section we consider two identical individuals a and c that differ only in their health. Individual c is in better health than individual a ($H^c > H^a$), but is otherwise identical to individual a, i.e., all exogenous variables and functions are assumed to be the same as for person a.

Investment in curative care: Consider the top right-hand panel of Figure 2 first. There is no direct effect of a higher level of health on the user cost of health capital at the margin $\sigma_H(t) = \pi_{I_m}(t)[d(t) + \delta - \widetilde{\pi_{I_m}}(t)]$.[32] However, the production benefit of health $\varphi_H(t) = \partial Y(t)/\partial H(t)$ is lower (DRTS [assumption 3]), and the resulting curve shifts upward (labeled $q_A(0)(\sigma^c_H - \varphi^c_H)$). Further, the marginal utility of health is lower (curve labeled $\partial U^c/\partial H$; diminishing marginal utility [assumption 2]). An indirect effect operates through consumption, is assumed to be smaller than the direct effects, and is discussed below. These shifts are associated with a lower optimal level of investment in curative care $I^c_m < I^a_m$.

Now turn to the top left-hand panel of Figure 2 which shows the associated shifts as a function of health. Assuming a lower level of investment in curative care $I_m(t)$, the user cost of health capital $\sigma_H(t)$ is smaller (assumption 5) and the net marginal user cost of health capital shifts downward (labeled $q_A(0)(\sigma^c_H - \varphi^c_H)$). Further, an indirect effect on the marginal utility of health $\partial U(t)/\partial H(t)$ operates through consumption. Higher health increases the marginal utility of consumption and hence increases the level of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption (assumption 6). This in turn increases the marginal utility of health and the curve shifts upward (labeled $\partial U^c/\partial H$). Thus higher health is associated with two competing shifts: (i) a shift *of* the curve through higher consumption, and (ii) a shift *along* the curve because of diminishing marginal utility (assumption 2). Again, health provides a constraint and both curves have to intersect at H^c. The marginal utility at the point of intersection $\partial U^c/\partial H$ determines the marginal utility of health in the top right-hand panel of Figure 2 (draw a horizontal line from the top left-hand to the top right-hand panel in Figure 2). There, the intersection of the two curves determines the optimal level of health

[32] There is however an indirect effect on the biological aging rate $d(t)$ because health affects choices made in working environment and in life style (operating through $C_h(t)$, $C_u(t)$, $z(t)$ and $I_p(t)$). This secondary effect is assumed to be small, which would be the case if $\partial d(t)/\partial C_h(t)$, $\partial d(t)/\partial C_u(t)$, $\partial d(t)/\partial z(t)$ and $\partial d(t)/\partial I_p(t)$ are small. However, as we will see, even under this assumption, as time passes lower levels of healthy consumption, curative and preventive care and higher levels of unhealthy consumption and job-related health stress lead to increasing health disadvantage over the life cycle.

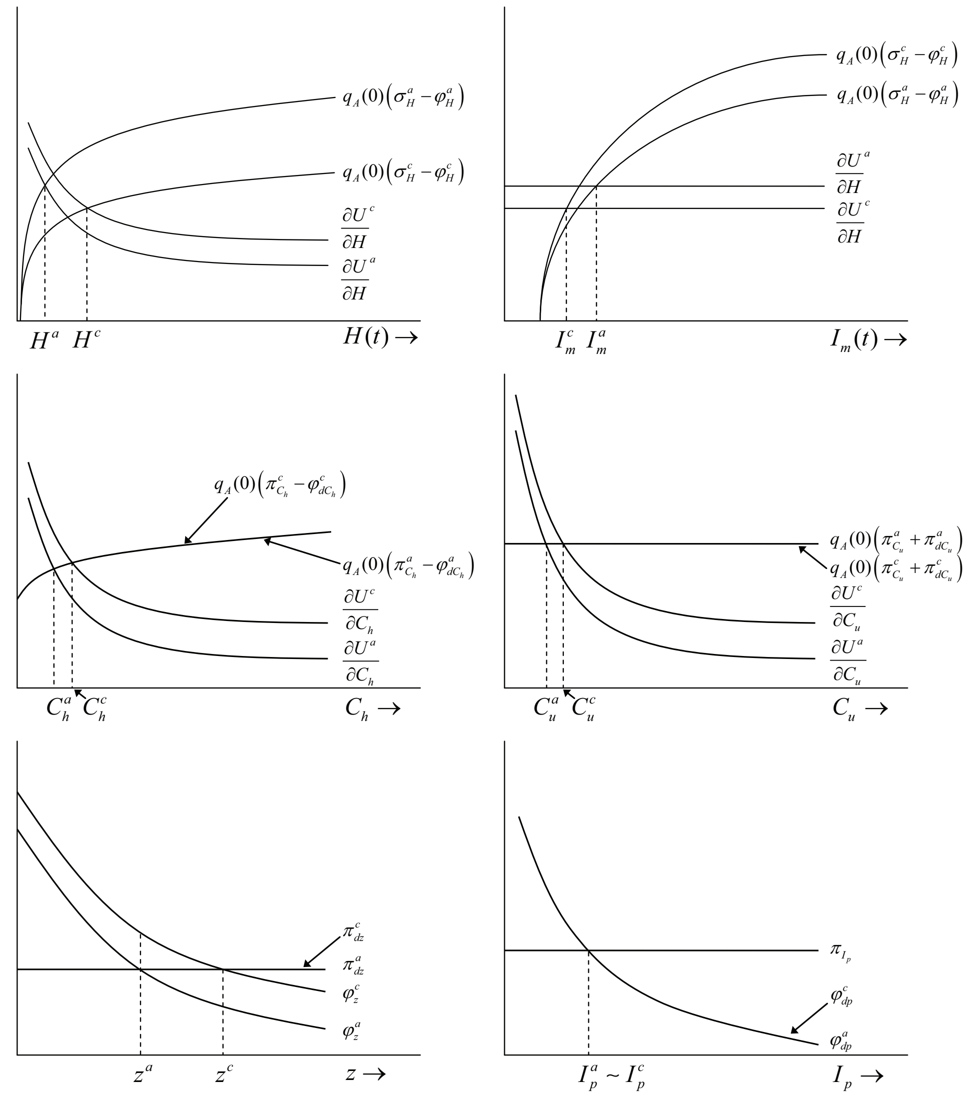

Figure 2: **Differences in Health:** *Marginal consumption $\partial U/\partial H$ and marginal production benefit φ_H of health versus the user cost of health capital at the margin σ_H as a function of health (top left) and as a function of health investment (top right). Marginal utility of healthy consumption $\partial U/\partial C_h$ versus the marginal monetary cost π_{C_h} and the marginal health benefit φ_{dC_h} of healthy consumption $C_h(t)$ (center left); Marginal utility of unhealthy consumption $\partial U/\partial C_u$ versus the marginal monetary cost π_{C_u} and the marginal health cost π_{dC_u} of unhealthy consumption $C_u(t)$ (center right); Marginal health cost π_{dz} versus the marginal production benefit φ_z of job-related health stress $z(t)$ (bottom-left); Marginal monetary cost π_{I_p} versus the marginal health benefit φ_{dI_p} of investment in preventive care $I_p(t)$ (bottom-right). In labeling the curves we have omitted the time varying term with exponent $(\beta - \delta)t$.*

investment, and $I^c_m < I^a_m$ (consistent with our assumption).[33] Hence, a larger health stock $H^c > H^a$ reduces the marginal monetary cost of curative care $\pi^c_{I_m} < \pi^a_{I_m}$ (see equation 16).

Healthy and unhealthy consumption: Changes in health and in the marginal monetary cost of curative care have no direct effect on the marginal monetary cost of healthy consumption $\pi_{C_h}(t)$ (equation 20) and of unhealthy consumption $\pi_{C_u}(t)$ (equation 23), hence both are shown as unchanged in the center-left and center-right panels of Figure 2. A higher level of health increases the health benefit of healthy consumption $\varphi_{dC_h}(t)$ (and the health cost of unhealthy consumption $\pi_{dC_u}(t)$), yet at the same time the health benefit (cost) is reduced through a lower marginal monetary cost of curative care as healthy individuals demand less curative care ($\varphi_{dC_h}(t) \propto \pi_{I_m}(t)H(t)$ and $\pi_{dC_u}(t) \propto \pi_{I_m}(t)H(t)$; see equations 21 and 24). The net effect is ambiguous. To reflect this, we show both the net marginal cost of healthy consumption (solid line labeled $q_A(0)(\pi^c_{C_h} - \varphi^c_{dC_h})$; center-left panel of Figure 2) and the net marginal cost of unhealthy consumption (solid line labeled $q_A(0)(\pi^c_{C_u} + \pi^c_{dC_u})$; center-right panel of Figure 2) as being unchanged (i.e., $\pi^c_{I_m} H^c \simeq \pi^a_{I_m} H^a$; we will return to this point later).

Now turn to the marginal utility of healthy consumption $\partial U(t)/\partial C_h(t)$ and of unhealthy consumption $\partial U(t)/\partial C_u(t)$. The direct effect of higher health $H^c > H^a$, due to complementarity in utility of consumption and health (assumption 6) is to shift both the marginal utility of healthy consumption $\partial U(t)/\partial C_h(t)$ (curve labeled $\partial U^c/\partial C_h$, center-left panel of Figure 2) and of unhealthy consumption $\partial U(t)/\partial C_u(t)$ (curve labeled $\partial U^c/\partial C_u$, center-right panel of Figure 2) upward.[34] The optimal solution for healthy consumption as well as for unhealthy consumption is higher: $C^c_h > C^a_h$ and $C^c_u > C^a_u$.

In the scenario discussed above it was assumed that the net marginal cost of healthy and of unhealthy consumption remain unchanged. Now consider two alternative scenarios. In scenario 1, the direct effect of higher health $H^c > H^a$ exceeds the indirect effect of changes in the marginal cost of curative care (as a result of higher health), i.e., $\pi^c_{I_m} H^c > \pi^a_{I_m} H^a$, and in scenario 2 we explore the opposite, i.e., $\pi^c_{I_m} H^c < \pi^a_{I_m} H^a$.[35] In scenario 1 both the marginal health benefit of healthy consumption $\varphi_{dC_h}(t)$ and the marginal health cost of unhealthy consumption $\pi_{dC_u}(t)$ are higher for individual c.[36] This further increases the level of healthy consumption C^c_h, but lowers

[33]Note that solutions are possible in which higher health leads to greater investment in health, i.e. a positive correlation between health and curative care (which is generally not observed). This requires the indirect effect on the marginal utility of health, operating through a higher level of consumption (as a result of greater health), to be quite substantial. Such solutions cannot be ruled out but appear less plausible.

[34]An indirect effect operates through consumption. Because of substitutability in utility of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption (assumption 7), both the marginal utility of healthy consumption $\partial U(t)/\partial C_h(t)$ and the marginal utility of unhealthy consumption $\partial U(t)/\partial C_u(t)$ shift downward. Assuming that the direct effect dominates the indirect effect the net result is nevertheless an upward shift.

[35]Scenario 1 corresponds to a small elasticity of health investment with respect to health and scenario 2 corresponds to a high elasticity. Assume $I_m(t) \propto H^{-\gamma}$, where γ is the elasticity of health investment with respect to health. Scenario 1 ($\pi^c_{I_m} H^c > \pi^a_{I_m} H^a$) then implies $\gamma < 1/(1-\alpha)$ while scenario 2 implies $\gamma > 1/(1-\alpha)$.

[36]Although these shifts are not depicted it is useful to use Figure 2 for reference. Scenario 1 implies a downward shift

the level of unhealthy consumption C_u^c (compared with the example shown in the center-left hand panel of Figure 2).[37] In scenario 2 we expect to observe the opposite pattern: higher health $H^c > H^a$ decreases the level of healthy consumption C_h^c, and further increases the level of unhealthy consumption C_u^c (compared with the example shown in the center-right hand panel of Figure 2).

Job-related health stress and investment in preventive care: Greater health is potentially associated with a greater marginal production benefit of job-related health stress $\varphi_z(t) = \partial Y(t)/\partial z(t)$ (curve labeled φ_z^c; bottom-left panel of Figure 2) as healthy individuals have higher earnings $Y(t)$ (see equation 9). The marginal monetary cost of preventive care $\pi_{I_p}(t)$ is independent of the level of health (equation 29). The effect on the marginal health cost of job-related health stress and the marginal health benefit of investment in preventive care is once more ambiguous (see equations 26 and 30) and both are shown as unchanged (i.e., $\pi_{I_m}^c H^c \simeq \pi_{I_m}^a H^a$). The resulting optimal level of job-related health stress is higher $z^c > z^a$ (bottom-left panel of Figure 2) and the level of investment in preventive care is unchanged $I_p^c \sim I_p^a$ (bottomright panel of Figure 2).

In scenario 1 (scenario 2) the marginal health cost of job-related health stress π_{dz} and the marginal health benefit of investment in preventive care φ_{I_p} are higher (lower), and the level of job-related health stress decreases (increases) and the level of investment in preventive care increases (decreases) with respect to the case shown in the bottom-left and bottom-right panels of Figure 2.

Summary and discussion – the effect of health on behavior: Table 2 provides a brief overview of the effect of greater health on behavior. Regardless of the scenario, individuals in better health invest less in curative care $I_m(t)$. In scenario 1 individuals consume more healthy consumption $C_h(t)$ and invest more in preventive care $I_p(t)$, while the effect on unhealthy consumption $C_u(t)$ and job-related health stress is ambiguous. In scenario 2 individuals consume more unhealthy consumption $C_u(t)$, engage more in job-related health stress $z(t)$, and invest less in preventive care $I_p(t)$, while the effect on healthy consumption $C_h(t)$ is ambiguous.

If there is no complementarity in utility of consumption and health (assumption 6) the predictions would remain the same, except that the effect of greater health on healthy consumption is negative in scenario 2 and the effect of greater health on unhealthy consumption is positive in scenario 1 (i.e., not ambiguous as shown in Table 2). If the relation were instead one of substitutability (worse health improves the utility of consumption), solutions are possible in which

---

in the net marginal costs of healthy consumption with respect to the curve shown $q_A(0)(\pi_{C_h}^c - \varphi_{dC_h}^c)$ in the center-left panel through increased φ_{dC_h}. Scenario 1 also implies an upward shift in the net marginal cost of unhealthy consumption with respect to the curve shown $q_A(0)(\pi_{C_u}^c + \pi_{dC_u}^c)$ in the center-right panel through increased π_{dC_u}.

[37] These shifts are exacerbated due to substitutability in utility of healthy $C_h(t)$ and unhealthy $C_u(t)$ consumption (assumption 7), as the marginal utility of healthy consumption $\partial U(t)/\partial C_h(t)$ increases for a lower level of unhealthy consumption $C_u(t)$ and the marginal utility of unhealthy consumption $\partial U(t)/\partial C_u(t)$ decreases for a higher level of healthy consumption $C_h(t)$.

Scenario	1	2
$I_m(t)$	$-$	$-$
$C_h(t)$	$+$	$+/-$
$C_u(t)$	$+/-$	$+$
$z(t)$	$+/-$	$+$
$I_p(t)$	$+$	$-$

Table 2: *The effect of higher health $H^c > H^a$ on behavior.*

greater health leads to lower levels of consumption and more investment in curative care. This is generally not observed.

4 Discussion and conclusions

The aim of this paper is to provide a contribution toward a theory of the relation between health and socioeconomic status (SES) over the lifecycle. Our life-cycle model incorporates multiple mechanisms that could explain (jointly) a large part of the observed disparities in health by SES. In our model, lifestyle factors (preventive care, healthy and unhealthy consumption), working conditions (physical and psychosocial health stresses), living conditions (housing, neighborhood social environment), curative care and the constraining effect of health on work are mechanisms through which SES (endowed wealth, life-time earnings and education) and health are related.

The main mechanism through which SES translates into health is by increasing the marginal cost of and the demand for curative care. This in turn increases the health benefit of (and hence demand for) preventive care and healthy consumption, and the health cost of (and hence reduced demand for) unhealthy working and living environments, and unhealthy consumption.

Even without the inclusion of additional potential mechanisms responsible for the SES health gradient (beside utilization of curative care), the theory predicts differences in the "effective" rate of health decline $\dot{H}(t)$ between high- and low-SES individuals due to differences in the level of investment in curative care $I_m(t)$. This addresses the criticism leveled by Case and Deaton (2005). But greater SES also induces healthy lifestyles, encourages investment in preventive care and protects individuals from the health risks of physical working conditions (e.g., hard labor) and/or psychosocial aspects of work (e.g., low status, limited control, repetitive work, etc) that are detrimental to health.

Endowed wealth, life time earnings and education each operate in distinct ways. The effect of greater earnings over the life cycle on health differs from the effect of greater endowed wealth in that the "wealth" effect is moderated by the higher opportunity cost of time. Plausibly, however, the effect of greater earnings over the life cycle dominates the opportunity cost effect. For example, Dustmann and Windmeijer (2000) and Contoyannis et al. (2004) find a positive effect on health from a permanent wage increase and a negative effect from a transitory wage increase. The effect of education on health is similar to that of greater earnings over the life cycle, but with the additional effect of increasing the efficiency of the production and consumption of curative and preventive care.

Irrespective of the SES indicator, for individuals who are initially equally healthy, the health trajectories of high and low SES individuals will begin to diverge. In addition, the higher the health stock, the greater the earnings (e.g., see equation 9) such that reverse causality (from health to SES) could further reinforce the SES health gradient. Results from earlier studies (Ehrlich and Chuma, 1990; Ehrlich, 2000; Galama et al. 2009) suggest that the more rapidly worsening health of low SES individuals could lead to early withdrawal from the labor force and shorter life spans. Early withdrawal from the labor force may contribute to further increasing disadvantage (widening of the SES health gradient) as the associated loss of income disproportionally affects low SES individuals. Mortality selection, i.e. lower SES people are more likely to die early, may result in an apparently healthier surviving disadvantaged population, partially explaining the narrowing of the gradient in late age.

Further, depending on the elasticity of investment in curative care with respect to health, the predicted divergence in health trajectories between low and high SES individuals could be further reinforced (scenario 1; small elasticity) or mitigated (scenario 2; large elasticity). In scenario 2 (see Table 2 [opposite signs for lower health] in section 3.4), over time the rate of divergence slows as subsequent lower levels of health encourages low SES individuals to invest more in health and engage in healthier behavior. Thus the theory predicts an initial widening and potentially a subsequent narrowing of the SES-health gradient. In scenario 1 less healthy individuals engage in less healthy behavior (with the exception of investment in curative care), and the theory predicts a continued widening of the gradient with age (or a weaker narrowing process).

Scenario 1 thus predicts a process of *cumulative advantage* for high SES individuals. 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. In contrast, scenario 2 predicts an economic variant of the *age-as-leveller* hypothesis (House et al. 1994; Elo and Preston, 1996; Beckett, 2000). The age-as-leveler hypothesis maintains that deterioration in health is an inevitable part of aging irrespective of SES with the result that the SES-health gradient narrows at later ages. Relative to the disadvantaged, economically advantaged people may be better able to postpone, but not prevent, declining health status.

Our theory can explain additional stylized empirical facts. The model predicts that individuals in better health invest less in curative care $I_m(t)$. This finding is supported by casual observation (the healthy do not go to the doctor) and by numerous empirical studies that find a strong negative correlation between measures of health and measures of curative (medical) care usage (see Galama and Kapteyn, 2009, for an overview of the empirical literature). Further, as our health declines with age, the demand for curative care increases. If the effect of deteriorating health on investment in curative care dominates the effect of the opportunity cost of time,[38] the model is capable of reproducing the observation that young individuals invest little in curative care, the middle-aged more and the elderly the most.

Another prediction of the theory 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) when compared to lower SES individuals. Greater wealth permits more consumption but also increases the marginal monetary value of health lost. This could provide a plausible 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 for health) than low SES individuals (e.g., Cutler and Lleras-Muney, 2008; Stringhini et al. 2010).

Our theory suggests that the SES health gradient could be strong in countries with universal health care coverage and low deductions, where the price of curative care is low and health care is affordable for everyone, as well as in countries with large uninsured populations. The marginal cost of curative care is largely determined by life-time wealth ($q_A(0)$, i.e., SES) and by the health stock $H(t)$ (see the first-order condition 15). Thus, a low price of curative care $p_m(t)$ does not influence the marginal monetary cost of curative care but increases the demand for curative care $I_m(t)$ (see equation 17 and keeping $\pi_{I_m}(t)$ unchanged). Further, because the marginal cost of curative care is not sensitive to price, the marginal health benefit of healthy consumption and preventive care and the health cost of unhealthy consumption and job-related health stress are unchanged. Thus the price of curative care does not affect choices in consumption, preventive care and in living and working environments directly, and also in countries with universal health care coverage and low deductibles there will be a significant SES health gradient. This is particularly true if medical care is not a large determinant of the SES health gradient (e.g., Adler et al. 1993) and could explain why the observed SES health gradient over the life cycle is strikingly similar between countries with relatively low levels of protection from loss of work and health risks, such as the U.S., and those with stronger welfare systems, such as the Netherlands (e.g., Smith, 1999; 2004; 2007; Case and Deaton 2005; van Kippersluis et al. 2010).

The predictions regarding the effects of SES on health depend on the notion that health has both intrinsic as well as instrumental value.[39] Differences in endowed wealth $q_A(0)$ have no effect on health if health does not provide utility (e.g., in the pure investment model, Grossman, 1972a; 1972b). In this case, the effect of greater earnings over the life cycle would be ambiguous as a higher wage rate increases both the user cost of health capital $\sigma_H(t)$ and the production benefit of health $\varphi_H(t)$. The effect of education would mostly operate through greater efficiency of medical and curative care.

[38]Low at young ages and high in middle and old age as a result of the typical hump-shaped wage profile with age (e.g., Mincer, 1974).

[39]As recognized by, e.g., Mushkin (1962) who noted that *" Health services . . . are partly investment and partly consumption . . . An individual wants to get well so that life for him may be more satisfying. But also when he is well he can perform more effectively as a producer"*

Thus, if health is mostly valued for its production benefit (generating earnings) this could explain the absence of strong evidence for a causal effect of financial indicators of SES on health (e.g., Cutler et al. 2010). Another possible explanation of this finding is that the effect of SES on health accumulates over time. The effect of financial indicators of SES on health is typically estimated contemporaneously (or with a small delay) and the wealth effect may be countered by the opportunity cost of time. Education, on the other hand, is obtained early in life (and hence its effect has had ample time to accumulate) and education potentially increases the efficiency of the production and consumption of curative and preventive care. This may provide an explanation for the strong effect of education on health outcomes observed in empirical studies (e.g., Grossman, 2000; Lleras-Muney, 2005; Silles, 2009). It also suggests that the protective effect of SES on health, in particular education, increases with age (Ross and Wu, 1996; Lynch, 2003).

In order to illustrate the theory and to derive predictions, we have made assumptions about the nature of the relations between functions of interest. The assumptions (1 to 5) of diminishing or constant returns to scale are commonly made in economics. If there is no complementarity in utility of consumption and health (assumption 6) the predictions would remain the same. If the relation were instead one of substitutability (worse health improves the utility of consumption), solutions are possible in which greater SES leads to less investment in curative care. This is generally not observed. If there were no substitutability in utility of healthy and unhealthy consumption (assumption 7) higher SES would not lead to a reduction in unhealthy consumption (except for severaly unhealthy consumption) but would still be associated with a shift toward healthy consumption (a smaller fraction of a larger budget is devoted to unhealthy consumption).

Our model includes major mechanisms identified in a review of the literature as explaining (jointly) a large part of the observed disparities in health by SES. Given the complexity (e.g., Cutler et al. 2010) of the various relations between SES and health, we have focused on potential explanations that a) explain a large part of the gradient and b) are relatively straightforward to include in our theoretical framework.

Compared to Grossman (1972a; 1972b), Ehrlich and Chuma (1990) and Case and Deaton (2005) the model presented in this paper contains several improvements and extensions: (i) A distinguishing feature is our interpretation of the relation (15) as being the first-order condition for optimal health investment $I_m(t)$, conditional on the level of the health stock $H(t)$, rather than the first-order condition for optimal health $H(t)$. This interpretation necessitates the assumption of decreasing-returns-to-scale (DRTS) in the health production function (as in Ehrlich and Chuma, 1990; see also Dustmann and Windmeijer, 2000; and Liljas, 2000), and addresses the indeterminacy problem ("bang-bang" solution) for investment in curative care (Ehrlich and Chuma, 1990), ensures that investment in curative (medical) care is non-negative (for the usual assumptions of functional forms), reproduces the observed negative relation between health and the demand for medical care, finds the health stock to be a function of initial health, past biological aging and past health investments made, and explains differences in the level of health as well as the rate of health decline between low and high SES groups (see Galama, 2010). Addressing these issues has been crucial: unlike alternative life-cycle models of health, medical care, and SES, our formulation can explain the formation of disparities in health by SES with age. (ii) We have included the concept of healthy consumption (as well as unhealthy consumption as in Case and Deaton, 2005) and allow the demand for consumption to be governed both by the direct monetary price of consumption as well as the indirect health benefit (healthy consumption) or indirect health cost (unhealthy consumption). Case and Deaton (2005) on the other hand consider an unhealthy consumption good whose price is only paid in terms of health. (iii) We have broadened the concept of "job-related health stress" to include not only hard/risky labor (as in Case and Deaton, 2005) but also psychosocial aspects of work that are detrimental to health. (iv) We have argued that the effect of housing and neighborhood social environment can be included by extending the definition of healthy consumption as well as exogenous environmental factors to include relevant aspects of housing and neighborhood characteristics. (v) We have introduced the concept of preventive care.

Numerical methods are required to solve the full model, including endogenous retirement decisions and mortality. With regard to mortality, the model provides a natural way to include length of life. In Grossman's original formulation (Grossman, 1972a; 1972b) length of life is determined by a minimum health level H_{min}, below which an individual dies. Endogenous length of life can be incorporated as in Ehrlich and Chuma (1990) and Ehrlich (2000) and simulated and calibrated as in Ehrlich and Yin (2005). With regard to retirement, as emphasized by Smith (2004) and Case and Deaton (2005), reverse causality from health to income through labor force participation could be an important mechanism explaining the SES-health gradient. In our model, this could be incorporated by an endogenous retirement age (as in Galama et al. 2009).

Another important extension of our model would be to incorporate insights from the literature on socioeconomic differences in the evolution of child health (e.g., Case et al. 2002; Currie and Stabile, 2003; Currie et al. 2007; Murasko, 2008), and from the literature on the impact of fetal and early-childhood conditions on health in adulthood (e.g., Barker et al. 1993; Case et al. 2005; van den Berg et al. 2006).[40] This might be feasible by including the production of health by the family (including the health of the child) similar to, e.g., Jacobson (2000) and Bolin et al. (2001; 2002a; 2002b).

We do not explicitly take into account the influence of the wider social context and social relationships of the family or neighborhood on health (House et al. 1988; Robert, 1998; Kawachi and Berkman, 2003). Less affluent areas are more polluted, have lower quantity and quality of municipal services, have higher crime rates, and are associated with unhealthy lifestyles (Robert, 1998). Also, the social isolation induced by poor quality and quantity of social contacts is an important risk factor for health (House et al. 1988). In our model this is partly captured by the exogenous part of the biological aging rate (exogenous environmental factors $\xi(t)$). However, it is

[40]The potential influence of childhood health on education is not included in our formulation as education is treated as being predetermined by the time individuals join the labor-force. Childhood conditions can be accounted for by treating the health status of an individual joining the labor force and investment in human capital prior to adulthood as initial conditions, i.e., we take initial health $H(0)$ and years of schooling E as given. Our model is therefore limited to explaining the formation of disparities in health from early adulthood till old age but not during childhood or the fetal period. As a result, the formulation cannot model the possible joint determination of education and health.

likely that social factors are partly endogenous to socioeconomic status (Robert, 1998). The role of the wider social context, social relationships, and other psycho-social risk factors (House et al. 1988; 1994; Robert, 1998; Kawachi and Berkman, 2003) can partially be captured in our model by extending the definition of healthy consumption to include choice of housing / neighborhood social environment. This might be further extended by including social capital similar to, e.g., Bolin et al. (2003).

We have not explicitly included racial and gender disparities in health. Racial categories importantly capture differences in power, status, and resources (Williams, 1999). Differences in SES between racial groups account for most of the observed racial disparities in health (Williams and Collins, 1995; Lillie-Blanton et al. 1996). Yet, racial differences in health and mortality persist even at "equivalent levels" of SES, and race/ethnicity has an independent effect beyond indicators of socioeconomic status (e.g., House and Williams, 2000). To the extent that racial/ethnic influences act independently of SES, race/ethnicity can be included in our formulation through the exogenous component of the biological aging rate. The same holds for gender disparities in health, if operating independently of SES (Luchenski et al. 2008). However, it has been argued that gender and race potentially moderate the relation between SES and health. It could be that discrimination makes it difficult to translate high SES into good health, or that employer discrimination makes minorities in poor health particularly likely to lose their jobs. The literature is inconclusive to what extent race/ethnicity and gender moderate the relationship between SES and health (Matthews et al. 1999; House and Williams, 2000; Luchenski et al. 2008).

Lastly, insights from the behavioral-economic and psychological literature regarding myopia and lack of self-control (e.g., Blanchflower et al. 2009) might be incorporated following Laibson (1998). Uncertainty (e.g., health shocks) could be included similar to, e.g., Cropper (1977), Dardanoni and Wagstaff (1990), Liljas (1998) and Ehrlich (2000). Joint determination of health and socioeconomic status due to factors such as intelligence, cognitive ability and non-cognitive skills may be incorporated by allowing these factors to raise the efficiency of household production in a similar way as education (e.g., Chiteji, 2010).

Empirical estimation of the model is needed to test the assumptions and the theoretical predictions presented in this work and to assess the relative importance of mechanisms, study interactions between mechanisms, and disentangle the different patterns of causality. This will require developing structural- and reduced-form relations. Model estimates may contribute to improving our understanding of the operational roles of major mechanisms in explaining the SES health gradient, and to simulating the long-term effects of policy interventions.

References

Adler, N.E., Boyce, T., Chesney, M., Folkman, S., & Syme, L. (1993), "Socioeconomic inequalities in health: No easy solution", Journal of the American Medical Association 269: pp. 3140-3145.

Adler, N.E., Boyce, T., Chesney, M.A., Cohen, S., Folkman, S., Kahn, R.L., Syme, S.L. (1994), "Socioeconomic status and health: The challenge of the gradient", American Psychologist 49(1): pp. 15-24.

Almond, D. (2006), "Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 U.S. population" Journal of Political Economy 114: pp. 562-712.

Auld, M.C. and N. Sidhu (2005), "Schooling, cognitive ability, and health", Health Economics, 14(10): pp. 1019-1034.

Barker, D., Gluckman, P.D., Godgrey, K.M., Harding, J.E., Owens, J.A., Robinson, J.S. (1993), "Fetal nutrition and cardiovascular disease in adult life", Lancet 341(8850): pp. 938-941.

Barsky, R.B., F.T. Juster, M.S. Kimball and M.D. Shapiro (1997), "Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement study", Quarterly Journal of Economics, 112(2): pp. 537579.

Beckett, M. (2000), "Converging Health Inequalities in Later Life: An Artifact of Mortality Selection?", Journal of Health and Social Behavior, 41: pp. 10619.

Behrman, J.R., and M.R. Rosenzweig (2004), "Returns to birth weight", Review of Economics and Statistics 86(2): pp. 586-601.

Black, S.E., P.J. Devereux, and K.G. Salvanes (2007), "From the cradle to the labor market? The effect of birth weight on adult outcomes", Quarterly Journal of Economics 122(1): pp. 409-439.

Blanchflower, D.G., Oswald, A.J., van Landeghem, B. (2009), "Imitative Obesity and Relative Utility", Journal of the European Economic Association 7: pp. 528-538.

Bleakley, H. (2007), "Disease and development: Evidence from hookworm eradication in the American South", Quarterly Journal of Economics 122(1): pp. 73-117.

Blundell, R., and T. MaCurdy (1999), "Labor Supply" in Orley Ashenfelter and David Card, eds., Handbook of Labor Economics, Amsterdam: North-Holland.

Bolin, K., Jacobson, L. and Lindgren, B. (2001), "The family as the producer of healthWhen spouses are Nash bargainers", Journal of Health Economics 20: pp. 349-362.

Bolin, K., Jacobson, L. and Lindgren, B. (2002a), "Employer investments in employee health – implications for the family as health producer", Journal of Health Economics 21: pp. 563-583.

Bolin, K., Jacobson, L. and Lindgren, B. (2002b), "The family as the health producer – When spouses act strategically", Journal of Health Economics 21: pp. 475-495.

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 56(12): pp. 2379-2390.

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 51: pp. 1019-1030

Case, A., Lubotsky, D. and Paxson, C. (2002), "Economic status and health in childhood: The origins of the gradient", American Economic Review 92: pp. 1308-1334.

Case A., A. Fertig, and C. Paxson (2005), "The lasting impact of childhood health and circumstance", Journal of Health Economics 24(2): pp. 365-389.

Case, A., Deaton, A. (2005), "Broken Down by Work and Sex: How our Health Declines", in: David A. Wise (ed.) (2005), *A*nalyses in the Economics of Aging, The University of Chicago Press, Chicago, pp. 185-212.

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 100: pp. 200-204.

Contoyannis, P. and Rice, N. (2001), "The impact of health on wages: evidence from the British Household Panel Survey", Empirical Economics 26: pp. 599-622.

Contoyannis, P., Jones, A.M. and Rice, N. (2004), "The dynamics of health in the British Household Panel Survey", Journal of Applied Econometrics, 19(4): pp. 473-503.

Cropper, M.L. (1977), "Health, investment in health, and occupational choice", Journal of Political Economy 85: pp. 1273-1294.

Currie, J. and Madrian, B.C. (1999), “Health, Health Insurance and the labour market”, In: Ashenfelter O, Card, D. (eds.) *Handbook of labour economics* 3: pp. 3309-3415, Amsterdam: Elsevier Science B.V.

Currie, J. and Stabile, M. (2003), “Socioeconomic status and child health: why is the relationship stronger for older children?”, American Economic Review 93: pp. 1813-1823.

Currie, A., Shields, M. and Price, W. (2007), “The Child Health/Family Income Gradient: Evidence from England”, Journal of Health Economics 26: pp. 213-232.

Cutler, D.M. and A. Lleras-Muney (2008), “Education and Health: Evaluating Theories and Evidence”, in Robert F. Schoeni et al., eds., Making Americans Healthier: Social and Economic Policy As Health Policy (National Poverty Center Series on Poverty and Public Policy).

Cutler, D.M., Lleras-Muney, A. and Vogl, T. (2010), “Socioeconomic Status and Health: Dimensions and Mechanisms”, Oxford Handbook of Health Economics, forthcoming.

Dardanoni, V., Wagstaff, A. (1990), “Uncertainty and the demand for medical care”, Journal of Health Economics 9: pp. 23-38.

Deary, I. (2008), “Why do intelligent people live longer?” Nature 456: pp. 175-176.

Deaton, A. (2002), “Policy implications of the gradient of health and wealth”, Health Affairs 21(2): pp. 13-30.

De Serpa, A.C. (1971), “A Theory of the Economics of Time”, The Economic Journal 81(324): 828-846.

Duncan, G.J., and Holmlund, B. (1983), “Was Adam Smith Right After All? Another Test of the Theory of Compensating Wage Differentials”, Journal of Labor Economics 1(4): pp. 366-379.

Dustmann, C., and Windmeijer, F. (2000), “Wages and the Demand for Health – A Life Cycle Analysis”, IZA Discussion Paper 171, Germany.

Ehrlich, I. and Chuma, H. (1990), “A Model of the Demand for Longevity and the Value of Life Extension”, Journal of Political Economy 98(4): pp. 761-782.

Ehrlich, I. (2000), “Uncertain lifetime, life protection, and the value of life saving”, Journal of Health Economics 19(3): pp. 341-367.

Ehrlich, I. and Yin, Y. (2005), “Explaining Diversities in Age-Specific Life Expectancies and Values of Life Saving: A Numerical Analysis”, The Journal of Risk and Uncertainty 31(2): pp. 129-162.

Elo, I.T. and S.H. Preston (1996), “Educational differentials in mortality: United States, 1979-85”, Social Science and Medicine, pp. 42:47-57.

Finkelstein, A., Luttmer, E.F.P. and Notowidigdo, M.J. (2008) “What Good Is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption”, NBER Working Paper 14089.

Forster, M. (2001), “The meaning of death: some simulations of a model of healthy and unhealthy consumption, Journal of Health Economics, 20: pp: 613-638.

Fuchs, Victor (1982), “Time Preferences and Health: An Exploratory Study”, in Economic Aspects of Health. Victor Fuchs, ed. Chicago: U. Chicago Press, pp. 93-120.

Fuchs, Viktor R. (1986), *The Health Economy*, first ed. Harvard University Press, Cambridge, MA.

Galama, T.J., Kapteyn, A., Fonseca, F., Michaud, P.C. (2009), “Grossman’s Health Threshold and Retirement”, RAND Working Paper, WR-658.

Galama, T.J. and Kapteyn, A. (2009), “Grossmans missing health threshold”, RAND Working Paper, WR-684.

Galama, T.J. (2010), “A theory of the demand for health, health investment and longevity”, in preparation.

Glied, S., and Lleras-Muney, A. (2008), “Technological Innovation and Inequality in Health”, Demography, 45(3): pp. 741-761.

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 99: pp. 10929-10934.

Grossman, M. (1972a), “The Demand for Health-A theoretical and Empirical Investigation”, New York: National Bureau of Economic Research.

Grossman, M. (1972b), “On the Concept of Health Capital and the Demand for Health”, Journal of Political Economy 80(2): pp. 223-255.

Grossman, M. (2000), “The Human Capital Model”, in *Handbook of Health Economics*, Volume

1, pp. 347-408, A.J. Culyer and J.P. Newhouse (Eds.), Elsevier Science.

Hanlon, P., Walsh D., and Whyte, B. (2006), Let Glasgowflourish. Glasgow: Glasgow Centre for Population Health.

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 86(1): pp. 5-45.

House, J.S., Landis, K.R., and Umberson, D. (1988), “Social relationships and health”, Science 241(4865): pp. 540-545.

House, J.S., Kessler, R.C., Regula Herzog, A. (1990), “Age, Socioeconomic Status, and Health”, Milbank Quarterly 68(3): pp. 383-411.

House, J.S., Lepkowski, J.M., Kinney, A.M., Mero, R.P., Kessler, R.C., Regula Herzog, A. (1994), “The Social Stratification of Aging and Health”, Journal of Health and Social Behavior 35(3): pp. 213-234.

House, J.S., and D.R. Williams (2000), “Understanding and Reducing Socioeconomic and Racial/Ethnic Disparities in Health”, pp. 81-124, in Smedley, B.D. and S. Leonard Syme (eds.), *Promoting Health: Intervention Strategies from Social and Behavioral Research*, Washington, DC: 284 National Academy Press.

Jacobson, L. (2000), “The family as producer of healthan extension of the Grossman model”, Journal of Health Economics 19: pp. 611-637.

Kawachi, I., and Berkman, L. (Eds.) (2003), *Neighborhoods and health*, Oxford, U.K.: Oxford University Press.

Khang, Y.H., Lynch, J.W., Yang, S., Harper, S., Yun, S.C., Jung-Choi, K., et al. (2009), “The contribution of material, psychosocial, and behavioral factors in explaining educational and occupational mortality inequalities in a nationally representative sample of South Koreans: relative and absolute perspectives”, Social Science & Medicine 68(5): pp. 858-866.

Kunst, A.E. and J.P. Mackenbach (1994), “The size of mortality differences associated with educational level in nine industrialized countries”, American Journal of Public Health 84: pp. 932-937.

Laibson, D. (1998), “Life-cycle consumption and hyperbolic discount functions”, European Economics Review 42: pp. 861-871.

Lantz, P.M., House, J.S., Lepkowski, J.M., Williams, D.R., Mero, R.P., Chen, J. (1998), “Socioeconomic Factors, Health Behaviors, and Mortality – Results From a Nationally Representative Prospective Study of U.S. Adults”, Journal of the American Medical Association 279(21): pp. 1703-1708.

Liljas, B. (1998), The demand for health with uncertainty and insurance”, Journal of Health Economics 17(2): pp. 153-170.

Liljas, B. (2000), “Insurance and imperfect financial markets in Grossman’s demand for health model - a reply to Tabata and Ohkusa”, Journal of Health Economics 19(5): pp. 821-827.

Lillie-Blanton, M., Parsons, P.E., Gayle, H., and Dievler, A. (1996), “Racial Differences in Health: Not Just Black and White, But Shades of Gray”, Annual Review of Public Health 17: pp. 411-448.

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 dEconomie et Statistique, special issue in memory of Zvi Griliches, No 79/80.

Lleras-Muney, A. (2005), “The relationship between education and adult mortality in the United States”, Review of Economic Studies 72(1): pp. 189-221.

Luchenski, S., Quesnel-Vallee, A, Lynch, J. (2008), “Differences between women’s and men’s socioeconomic inequalities in health: longitudinal analysis of the Canadian population, 19942003”, Journal of Epidemiology and Community Health 62:1036-1044.

Lynch, J.W., Kaplan, G.A., Shema, S.J. (1997), “Cumulative Impact of Sustained Economic Hardship on Physical, Cognitive, Psychological, and Social Functioning”, New England Journal of Medicine 337(26): pp. 1889-1895.

Lynch, S.M. (2003), “Cohort and Life-Course Patterns in the Relationship between Education and Health: A Hierarchical Approach”, Demography, 40: pp. 30931.

Mackenbach, J.P., Huisman, M., Andersen, O., Bopp, M., Borgan, J.K., Borrell, C., et al. (2004), “Inequalities in lung cancer mortality by the educational level in 10 European populations”, European Journal of Cancer 40(1): pp. 126-135.

Marmot, M., Bosma, H., Hemmingway, H., Brunner, E., Stansfeld, S. (1997a), “Contribution of job control and other risk factors to social variations in coronary heart disease incidence”, Lancet 350(9073): pp. 235-239.

Marmot, M., Ryff, C.D., Bumpass, L.L., Shipley, M., Marks, N.F. (1997b), "Social inequalities in health: next questions and converging evidence", Social Science & Medicine 44(6): pp. 901-910.

Marmot, M. (1999), "Multi-level approaches to understanding social determinants", in Lisa Berkman and Ichiro Icawachi (eds.), *Social Epidemiology*, Oxford: Oxford University Press.

Matthews, S., Manor, O., and Power, C. (1999), "Social inequalities in health: are there gender differences?", Social Science & Medicine 48(1): pp. 49-60.

Michaud, P.-C. and A. van Soest (2008), "Health and wealth of elderly couples: Causality tests using dynamic panel data models", Journal of Health Economics, 27: pp. 1312-1325.

Mincer, J.A. (1974), "Schooling, Experience, and Earnings", Columbia University Press.

Miguel, E. and Kremer, M. (2004), "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities", Econometrica 72(1): pp. 159-217.

Murasko, J.E. (2008), "An evaluation of the age-profile in the relationship between household income and the health of children in the United States", Journal of Health Economics, 27(6): pp. 1489-1502.

Mushkin, S.J. (1962), "Health as an investment", Journal of Political Economy 70(5): pp. 129-157.

Muurinen, J-M. (1982), "Demand for health: a generalized Grossman model", Journal of Health Economics 1: pp. 5-28.

Muurinen, J-M., and Le Grand, J. (1985), "The Economic Analysis of Inequalities in Health", Social Science & Medicine 20(10): pp. 1029-1035.

Olson, C.A. (1981), "An Analysis of Wage Differentials Received by Workers on Dangerous Jobs", Journal of Human Resources 16(2): pp. 167-185.

Oreopoulos, P. (2006), "Estimating average and local average treatment effects of education when compulsory schooling laws really matter", American Economic Review, 96(1): pp. 152-175

Preston S.H. and Elo, I.T. (1995), "Are educational differentials in adult mortality increasing in the United States?", Journal of Aging and Health, 7: pp. 476-496.

Robert, S.A. (1998), "Community-Level Socioeconomic Status Effects on Adult Health", Journal of Health and Social Behavior 39(1): pp. 18-37.

Ross C.E., and Wu C.L. (1996), "Education, Age, and the Cumulative Advantage in Health", Journal of Health and Social Behavior, 37: pp. 10420.

Royer, H. (2009), "Separated at Girth: US Twin Estimates of the Effects of Birth Weight", American Economic Journal: Applied Economics 1(1): pp. 49-85.

Schrijvers, C.T.M., van de Mheen, H.D., Stronks, K., Mackenbach, J.P. (1998), "Socioeconomic inequalities in health in the working population: the contribution of working conditions", International Journal of Epidemiology 27: pp. 1011-1018.

Seierstad, A. and Sydsaeter, K. (1977), "Sufficient Conditions in Optimal Control Theory", International Economic Review 18(2): pp. 367-391.

Seierstad, A. and K. Sydsaeter (1987), "Optimal Control Theory With Economic Applications", in *Advanced Texbooks in Economics*, Volume 24, Series Editors: C.J. Bliss and M.D. Intriligator, Elsevier, North Holland.

Silles, M. (2009), "The causal effect of education on health: evidence from the United Kingdom", Economics of Education Review, 28(1): pp. 122-128.

Skalicka, V., van Lenthe, F., Bambra, C., Krokstad, S. And Mackenbach, J. (2009), "Material, psychosocial, behavioural and biomedical factors in the explanation of relative socio-economic inequalities in mortality: evidence from the HUNT study", International Journal of Epidemiology 38(5): pp. 1272-1284

Smith, A. (1776), "An inquiry into the nature and causes of the wealth of nations", Reprinted Roy H. Campbell and Andrew S. Skinner, eds. Oxford: Clarendon Press, 1976.

Smith, R.S. (1978), "Compensating Wage Differentials and Public Policy: A Review", Industrial and Labor Relations Review 32(3): pp. 339-352.

Smith, J.P. (1999), "Healthy bodies and thick wallets", Journal of Economic Perspectives 13(2): pp. 145-166.

Smith, J.P. (2004), "Unraveling the SES-health connection", in Aging, Health, and Public Policy: Demographic and Economic Perspectives, a supplement to Population and Development Review 30: pp. 108-132.

Smith, J.P. (2007), “The impact of socioeconomic status on health over the life course”, Journal of Human Resources 42(4): pp. 739-764.

Stringhini, S., Sabia, S., Shipley, M., Brunner, E., Nabi, H., Kivimaki, M. and Singh-Manoux, A. (2010), “Association of Socioeconomic Position With Health Behaviors and Mortality”, Journal of the American Medical Association 303(12): pp. 1159-1166.

Usher, D. (1975), “Comments on the correlation between health and schooling”, In: N.E. Terleckyj, ed., Household Production and Consumption (Columbia University Press for the National Bureau of Economic Research, New York) 212-220.

Van den Berg G., Lindeboom M., Portrait, F. (2006), “Economic Conditions Early in Life and Individual Mortality”, American Economic Review 96(1): pp. 290-302.

Van Kippersluis, H., O’Donnell, O., van Doorslaer, E., Van Ourti, T., (2010), “Socioeconomic Differences in Health over the Life Cycle in an Egalitarian Country”, Social Science and Medicine 70(3): pp. 428-438.

Van Oort, F.V., van Lenthe, F.J. and Mackenbach, J.P. (2005), “Material, psychosocial, and behavioural factors in the explanation of educational inequalities in mortality in The Netherlands”, Journal of Epidemiology and Community Health, 59(3): pp. 214-220.

Viscusi, K.W. (1978), “Wealth Effects and Earnings Premiums for Job Hazards”, Review of Economics and Statistics 60: pp. 408-416.

Viscusi, K.W. (1979), *Employment hazards: An investigation of market performance*, Cambridge: Harvard U. Press.

Viscusi, K.W. (1993), “The Value of Risks to Life and Health”, Journal of Economic Literature 31(4): pp. 1912-1946.

Wagstaff, A. (1986), “The demand for health: Some new empirical evidence”, Journal of Health Economics 5: pp. 195-233.

Williams, D.R., and Collins, C. (1995), “U.S. socioeconomic and racial differences in health: Patterns and explanations”, Annual Review of Sociology 21: pp. 349386.

Williams, D.R. (1999), “Race, Socioeconomic Status, and Health: The Added Effects of Racism and Discrimination”, Annals of the New York Academy of Sciences 896: pp. 173-188.

Wolfe, J.R. (1985), “A Model of Declining Health and Retirement”, Journal of Political Economy 93(6): pp. 1258-1267.

Yen, I.H. and Kaplan, G.A. (1999), “Neighborhood Social Environment and Risk of Death: Multilevel Evidence from the Alameda County Study”, American Journal of Epidemiology 149(10): pp. 898-907.

Zweifel, P., Breyer, F., (1997), *Health Economics*, Oxford University Press, New York.

A First-order conditions

Associated with the Lagrangian (equation 12) we have the following conditions:

$$\begin{aligned}
\dot{q}_A(t) &= -\frac{\partial \Im(t)}{\partial A(t)} \Rightarrow \\
\dot{q}_A(t) &= -\delta q_A(t) \Leftrightarrow \\
q_A(t) &= q_A(0)e^{-\delta t},
\end{aligned} \tag{31}$$

$$\begin{aligned}
\dot{q}_H(t) &= -\frac{\partial \Im(t)}{\partial H(t)} \Rightarrow \\
\dot{q}_H(t) &= q_H(t)d(t) - \frac{\partial U(t)}{\partial H(t)}e^{-\beta t} - q_A(0)\frac{\partial Y(t)}{\partial H(t)}e^{-\delta t} \\
&= q_H(t)d(t) - \frac{\partial U(t)}{\partial H(t)}e^{-\beta t} - q_A(0)\varphi_H(t)e^{-\delta t} \Leftrightarrow \\
q_H(t) &= q_H(0)e^{\int_0^t d(x)dx} - \int_0^t \left[\frac{\partial U(s)}{\partial H(s)}e^{-\beta s} - q_A(0)\varphi_H(s)e^{-\delta s}\right] e^{\int_s^t d(x)dx} ds,
\end{aligned} \tag{32}$$

$$\begin{aligned}
\frac{\partial \Im(t)}{\partial X_h(t)} &= 0 \Rightarrow \\
\frac{\partial U(t)}{\partial C_h(t)} &= q_A(0)\frac{p_{X_h}(t)}{\partial C_h(t)/\partial X_h(t)}e^{(\beta-\delta)t} + q_H(t)\frac{\partial d(t)}{\partial C_h(t)}H(t)e^{\beta t} \\
&\equiv q_A(0)\,\pi_{C_h}(t)e^{(\beta-\delta)t} - q_H(t)\frac{\varphi_{dC_h}(t)}{\pi_{I_m}(t)}e^{\beta t},
\end{aligned} \tag{33}$$

$$\begin{aligned}
\frac{\partial \Im(t)}{\partial \tau_{C_h}(t)} &= 0 \Rightarrow \\
\frac{\partial U(t)}{\partial C_h(t)} &= q_A(0)\frac{w(t)}{\partial C_h(t)/\partial \tau_{C_h}(t)}e^{(\beta-\delta)t} + q_H(t)\frac{\partial d(t)}{\partial C_h(t)}H(t)e^{\beta t} \\
&\equiv q_A(0)\,\pi_{C_h}(t)e^{(\beta-\delta)t} - q_H(t)\frac{\varphi_{dC_h}(t)}{\pi_{I_m}(t)}e^{\beta t},
\end{aligned} \tag{34}$$

$$\begin{aligned}
\frac{\partial \Im(t)}{\partial X_u(t)} &= 0 \Rightarrow \\
\frac{\partial U(t)}{\partial C_u(t)} &= q_A(0)\frac{p_{X_u}(t)}{\partial C_u(t)/\partial X_u(t)}e^{(\beta-\delta)t} + q_H(t)\frac{\partial d(t)}{\partial C_u(t)}H(t)e^{\beta t} \\
&\equiv q_A(0)\,\pi_{C_u}(t)e^{(\beta-\delta)t} + q_H(t)\frac{\pi_{dC_u}(t)}{\pi_{I_m}(t)}e^{\beta t},
\end{aligned} \tag{35}$$

$$
\begin{aligned}
\frac{\partial \Im(t)}{\partial \tau_{C_u}(t)} &= 0 \Rightarrow \\
\frac{\partial U(t)}{\partial C_u(t)} &= q_A(0)\frac{w(t)}{\partial C_u(t)/\partial \tau_{C_u}(t)}e^{(\beta-\delta)t} + q_H(t)\frac{\partial d(t)}{\partial C_u(t)}H(t)e^{\beta t} \\
&\equiv q_A(0)\,\pi_{C_u}(t)e^{(\beta-\delta)t} + q_H(t)\frac{\pi_{dC_u}(t)}{\pi_{I_m}(t)}e^{\beta t}, \qquad (36)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \Im(t)}{\partial m_m(t)} &= 0 \Rightarrow \\
q_H(t) &= q_A(0)\left\{\frac{p_m(t)I_m(t)^{1-\alpha}}{\alpha[\partial I_m(t)/\partial m_m(t)]}\right\}e^{-\delta t} \\
&\equiv q_A(0)\,\pi_{I_m}(t)e^{-\delta t}, \qquad (37)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \Im(t)}{\partial \tau_{I_m}(t)} &= 0 \Rightarrow \\
q_H(t) &= q_A(0)\left\{\frac{w(t)I_m(t)^{1-\alpha}}{\alpha[\partial I_m(t)/\partial \tau_{I_m}(t)]}\right\}e^{-\delta t} \\
&\equiv q_A(0)\,\pi_{I_m}(t)e^{-\delta t}, \qquad (38)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \Im(t)}{\partial z(t)} &= 0 \Rightarrow \\
0 &= q_H(t)\frac{\partial d(t)}{\partial z(t)}H(t) - q_A(0)\frac{\partial Y(t)}{\partial z(t)}e^{-\delta t} \\
&\equiv q_H(t)\frac{\pi_{dz}(t)}{\pi_{I_m}(t)} - q_A(0)\,\varphi_z(t)e^{-\delta t}, \qquad (39)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \Im(t)}{\partial m_p(t)} &= 0 \Rightarrow \\
0 &= q_H(t)\frac{\partial d(t)}{\partial I_p(t)}H(t) + q_A(0)\frac{p_p(t)}{\partial I_p(t)/\partial m_p(t)}e^{-\delta t} \\
&\equiv -q_H(t)\frac{\pi_{dI_p}(t)}{\pi_{I_m}(t)} + q_A(0)\,\pi_{I_p}(t)e^{-\delta t}, \qquad (40)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \Im(t)}{\partial \tau_{I_p}(t)} &= 0 \Rightarrow \\
0 &= q_H(t)\frac{\partial d(t)}{\partial I_p(t)}H(t) + q_A(0)\frac{w(t)}{\partial I_p(t)/\partial \tau_{I_p}(t)}e^{-\delta t}, \\
&\equiv -q_H(t)\frac{\pi_{dI_p}(t)}{\pi_{I_m}(t)} + q_A(0)\,\pi_{I_p}(t)e^{-\delta t},
\end{aligned} \tag{41}
$$

Equation (33) or (34) combined with (37) or (38) provide the first-order condition for maximization of (12) with respect to healthy consumption (equation 19). Similarly, equation (35) or (36) combined with (37) or (38) provide the first-order condition for maximization of (12) with respect to unhealthy consumption (equation 22). Using (37) or (38) to obtain an expression for $\dot{q}_H(t)$ and substituting the results for $q_H(t)$ and $\dot{q}_H(t)$ in (32) we find the first-order condition for maximization of (12) with respect to investment in curative care (equation 15). Combining equations (37) or (38) and (39) to eliminate $q_H(t)$ we find the first-order condition for maximization of (12) with respect to job-related health stress (equation 25). Lastly, combining equations (37) or (38) and (40) or (41) to eliminate $q_H(t)$ we find the first-order condition for maximization of (12) with respect to preventive care (equation 28).