STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's thesis in economics
Academic year 2022–2023

RBC with Endogenous Health

Linnéa Ekberg (24167) and Rasmus Grehn (24137)

Abstract. The purpose of this paper is to contribute to the limited literature on RBC models with endogenous health accumulation. To this end, we develop an infinitely lived agent model with endogenous health, productivity shocks, and health shocks. Health is conceptualized as a capital stock in accordance with the Grossman model. Unlike previous literature, we assume that investments in health are made using medical expenditures and exercise, where time spent on health-improving physical activity is separate from utility-generating leisure. The model successfully replicates key relationships observed in the data - it predicts pro-cyclical medical expenditures, counter-cyclical physical activity, and a negative correlation between health status and medical expenditures. By simulating impulse response functions, we find that the agent responds to a temporary positive shock by reallocating their resources such that more consumption or more leisure can be enjoyed, whilst holding health state constant. This is achieved by substituting between medical expenditures and exercise in the production of health, suggesting that RBC models with endogenous health accumulation should account for a more complex health production than has previously been used.

Keywords: business cycle, health, health investment, time allocation

JEL: E32, I10, I12

Supervisor:	Kelly Ragan
Date submitted:	December 5, 2022
Date examined:	December 16, 2022
Discussant:	Chung Shun Man
Examiner:	Karl Wärneryd

We would like to express our gratitude to our advisor, Kelly Ragan, for her invaluable support and encouragement. We would also like to thank Karl Wärneryd and Chung Shun Man for their feedback and insightful remarks during the defense of our thesis.

Contents

List of Figures

1 Introduction

It is well-documented in the empirical literature that the health of an individual is responsive to business cycles. Whether health varies pro- or counter-cyclically depends on the proxy used for health. For example, suicide rates are found to be higher during recessions (Dos Santos, Tavares and Barros, 2016; Gerdtham and Johannesson, 2005; Ruhm, 2000) whereas Ruhm (2000) shows that total mortality and eight out of ten considered causes of death increase in expansions.

The variation in mortality over the business cycle is argued to come from shifts in the income and time constraints, which alter the behavior of the individual (Bellés-Obrero and Vall Castelló, 2018; Ruhm, 2000). While physical activity is found to be higher in recessions (Colman and Dave, 2013, 2018; Ruhm, 2005; Xu, 2013), smoking and obesity increase in expansions (Colman and Dave, 2018; Giri and Kumaresan, 2021; Ruhm, 2000, 2005; Xu, 2013). As pointed out by Giri and Kumaresan (2021), the overall effect of economic fluctuations on health is then the aggregate of these effects.

Despite the empirical evidence, the inclusion of health in the theoretical literature on real business cycles (RBCs) remains sparse. With this paper, we seek to contribute to the limited literature on RBC models with endogenous health accumulation. We propose an infinitely lived agents model with preferences over consumption, leisure, and healthy time. Healthy time further determines the amount of time available for activities such as work, leisure, and the production of future health. Health is produced using time and medical expenditures following the health capital framework developed by Grossman (1972a). We respond to one of the major criticisms directed at this concept by making the depreciation rate of health a negative function of current health capital, thereby ensuring that health deteriorates slower with improved health status. In addition to aggregate productivity shocks, we introduce a stochastic component in the deprecation rate of health in order to simulate health shocks.

To the best of our knowledge, only five attempts have been made to bridge the gap between theory and empirical literature by combining an RBC model with endogenous health. Failing to include health in such models may lead to conclusions that are subject to error due to missed dynamics. Introducing health into the model adds both an additional objective (given that health generates utility) and additional trade-offs caused by the fact that health and non-health production must compete for resources (Aísa and Pueyo, 2004; Zon and Muysken, 2001).

As pointed out by Scholz and Seshadri (2010), models that do not include health may therefore over-state the extent to which consumption is smoothed across periods. Indeed, Feng and Gomis-Porqueras (2011) show that an RBC model with endogenous health is better able to predict the observed variability in consumption and labor supply than a standard RBC model. They hypothesize that this improvement is driven by the need to smooth both consumption and health, leading to more volatile investments.

In a standard RBC model, consumption is smoothed by moving assets across periods via savings. Similarly, major health fluctuations can be avoided by adjusting health-improving investments. Following the Grossman model, adjustments of health investments imply adjusting the amount of medical expenses and the time allocated to the production of health. An agent who wishes to smooth both consumption and health therefore needs to find the optimal allocation of wealth between savings, current consumption, and current medical expenditures, as well as the optimal allocation of time between work, leisure and health-generating activities.

Whereas existing papers tend to focus on one of the trade-offs described above, we propose a more extensive model where health investments require the use of medical expenses and health-generating time (from now on exercise) and where an agent can divide their available time between work, leisure, and exercise. In other words, our model includes dynamics to analyze both the shift in the income and the time constraints described by the empirical literature. Furthermore, with the exception of Vasilev (2017), previous literature has not made a distinction between leisure and exercise, with the risk of overstating the time spent on health investments and consequently the health capital of an agent. As pointed out by Colman and Dave (2013), exercise is likely to be utility-reducing, given that as few as 5% of the adult American population get the recommended amount of daily physical activity. In contrast to the majority of previous literature, we therefore introduce the additional trade-off between utility-generating leisure and health-generating exercise.

Our model further stays closer to the conceptualization of health found in the Grossman model. Unlike previous literature, we transform the health capital of an agent into healthy time, which is directly valued in the utility function and that determines the amount of time available for various activities. In other words, our model captures health both as a consumption commodity and as an investment commodity in accordance with the reasoning by Grossman (1972a). While Vasilev (2017) and Laun (2012) allow available time to depend on health status, none of the existing models makes utility a function of healthy time. Instead, the agents derive utility directly from the stock of health capital.

We further propose the inclusion of health shocks. Two of the existing RBC models with endogenous health introduce a stochastic component in the health of an individual. First, Vasilev (2017) makes the outcome of the health investment uncertain. In this model, the agent does not know how much of the investment will transform into an actual improvement of the health capital. Second, the model by Laun (2012) allows the shock to directly affect both the utility and the time constraint of the agent. In contrast to these models, we add a stochastic component to the depreciation rate of capital, thus allowing a shock to directly affect the deterioration of health. The effect of the health shock on investment, time, and utility modeled in previous literature are captured indirectly in our model.

The depreciation rate used in our model also responds to a major criticism directed at the Grossman model. Whereas the law of motion of health capital used by Grossman (1972a) implies that health will deteriorate faster for a healthier individual, the literature on gerontology suggests that the opposite is true (Strulik, 2015). In order to capture the reverse relationship, we make the depreciation rate a decreasing function of the health capital. This further implies that it is less costly to maintain a good level of health once achieved.

We find that the proposed model is able to predict the key relationships discussed in the empirical literature. It successfully predicts pro-cyclical medical expenditures, counter-cyclical physical activity, and a negative correlation between health status and medical expenditures. By simulating impulse response functions, we find that a positive productivity shock leads to increased consumption of both medical and non-medical goods, in line with the income channel proposed by Ruhm (1994).

The productivity shock further leads to a decrease in exercise. Our results suggest that this occurs in order to increase time spent on leisure when the health stock can be maintained by substituting towards more medical consumption. The ability of the agent to substitute medical expenditures for exercise when faced with an expanding budget constraint implies that the health status is unaffected by the productivity shock, suggesting that the agent prefers to reallocate resources in order to avoid health fluctuations.

Based on our results, we conclude the the connection between medical consumption, exercise, and leisure is relevant in explaining the behavior of an agent when faced with productivity shocks. Consequently, we propose that an RBC model with endogenous health accumulation should account for a more complex health production than has previously been used.

The rest of the paper is structured as follows: first, section 2 provides an overview of previous literature. Next, section 3 presents our RBC model with endogenous health accumulation. Section 4 provides a calibration of the model to key relationships observed in the data. The main results are presented in section 5. Finally, we provide a discussion of the results and possible avenues for future research in section 6.

2 Literature Review

Our study seeks to contribute to the literature bridging the gap between standard RBC models and the empirical literature showing a correlation between business cycles and health. We begin by providing an overview of the empirical literature justifying such a contribution. Next, we review the literature on endogenous health accumulation. Finally, we conclude the section with an overview of existing RBC models with endogenous health.

2.1 Empirical Evidence

The empirical literature provides evidence to show that various health measures vary over the business cycle. Common proxies for health include mortality, by Ruhm (2000) found to be lower in recessions, and obesity, again found to be lower in economic downturns (Giri and Kumaresan, 2021; Ruhm, 2000, 2005). When attempting to explain these patterns, the literature argues that business cycles shift the income and the time constraints of the individuals, who as a result alter their behavior.

In his seminal paper, Ruhm (1994) presents evidence for the the procyclical nature of alcohol consumption and traffic deaths. He assigns most of the effect to changes in income over the business cycle, reasoning that as income drops during recessions, the individual consumes less of the normal goods alcohol and driving. Building on this finding, Ruhm (2000) shows that the drop in aggregate mortality in recessions can be explained by increased physical activity, improved diet, and reduced obesity and smoking. Unlike the decrease in alcohol consumption and driving, these effects are not attributed to reduced incomes. Instead, he emphasizes the importance of the opportunity cost of time, with Ruhm (2005) finding little evidence for the importance of income in explaining the healthier behaviors of individuals in recessions. He reasons that the increase in non-market time during recessions make it less costly for agents to undertake healthier behaviors.

2.2 Endogenous Health Accumulation

The literature on endogenous health accumulation is heavily influenced by the Grossman model. Grossman (1972a) treats health as a durable good that depreciates over time until it reaches a minimum threshold, at which point the agent dies. The stock of health can be adjusted to a desired level using medical expenses and time. This desired level is reached instantaneously following the assumptions of exogenous input prices and constant returns to scale in health investments. The health stock is transformed into healthy time, directly valued as a consumption good by the agent and used for activities such as working and the generation of future health capital.

Although iconic and widely used, the Grossman model has been subject to criticism. This critique can be divided into two fields - literature that finds flaws in the theoretical framework of the model and literature that finds that empirical transformations of the theoretical model fails to replicate the data.

2.2.1 Theoretical Criticism

Indeterminacy and Inconsistency

Arguably the most harmful critique directed at the Grossman model is the lack of unique solutions in the paths of health capital and health investment. Ehrlich and Chuma (1990) and T. J. Galama (2015) both argue that this is a result of the constant returns to scale (CRTS) in health investment used by Grossman (1972a), which makes the marginal cost of investment a constant function of current health state and the Hamiltonian a linear function of health investment. According to Ehrlich and Chuma (1990), this generally implies that health investment will have no interior equilibrium, giving rise to a "bang-bang" solution in which the agent invests either nothing or the entirety of their resources, where there is no upper bound to the optimal level of investment.

To solve this problem, Ehrlich and Chuma (1990) and T. J. Galama (2015) suggest the use of decreasing returns to scale (DRTS) in health investment. In contrast to CRTS, DRTS implies that the marginal cost of investment is an increasing function of health, ensuring that an interior solution exists. Concurring with this reasoning, T. J. Galama et al. (2012) confirm that the use of DRTS guarantees the existence of a unique solution by pointing out that the Hamiltonian of such a model is strictly concave and that health investment has a first-order condition strictly increasing in the investment.

However, the necessity of DRTS for the existence of a unique solution has been challenged. Ried (1998) not only criticizes Ehrlich and Chuma (1990) for failing to prove that CRTS leads to indeterminacy, he also provides evidence to show that optimal paths for health capital and health investment can be found given that a unique solution for the optimal length of life exists (see Ried (1998, p. 387) for the proof).

It is however debated whether the Grossman model provides a unique prediction for optimal longevity. Grossman (1972a) includes endogenous longevity by having agents die when their health capital drops below some given threshold, thereby allowing agents to choose their length of life by choosing the corresponding path of health capital. The question is whether this dynamic is sufficient to guarantee that agents choose a finite life. Ehrlich and Chuma (1990) argue that it does not suffice, pointing out that the lack of a transversality condition results in a model that fails to predict consistent choices. Solutions to this problem have been proposed by Laporte (2015), who suggests specifying an upper bound for the length of life whilst maintaining the endogenous characteristics of longevity by allowing agents to choose health capital such that they die at an age at or below this value, or alternatively by introducing a stochastic survival rate where the probability of survival depends on the health capital.

In contrast, Grossman (1998) argues that the Grossman model is capable of predicting a finite, endogenous length of life. He shows that the agent chooses optimal longevity through a process of iteratively maximizing lifetime utility with increasingly longer horizon and checking whether the health stock in the final period of life is less than or equal to the specified minimum threshold at which the agent dies (see Grossman (1998) for the proof). From this reasoning, he concludes that the model guarantees finite life given that the depreciation rate of health capital increases with age.

This discussion is summarized by Case and Angus (2005), who conclude that eternal life may not be unattainable in the Grossman model with CRTS, but it is not optimal given age-dependent depreciation rate of health capital. Given advanced age and the associated decline of health from one period to the next, the cost of health investment will eventually grow sufficiently large to make death preferable.

The Deterioration of Health

A second criticism directed at the Grossman model concerns the counter-intuitive representation of declining health. The law of motion of health capital used by Grossman (1972a), formulated as

$$H_{i+1} = (1 - \delta_i)H_i + I_i, \tag{1}$$

implies a positive relationship between the health capital of an individual and the amount of health lost in each period. In other words, the Grossman model predicts that the health of an individual will deteriorate faster the healthier the individual is. Because the stock of health capital tends to be largest at the beginning of life and decline at a rate slowed by investments until death, this aspect of the Grossman model has been interpreted as young individuals aging at a faster rate, contradicting empirical evidence from gerontology (Strulik, 2015).

A common solution to this problem, used by Grossman (1972a) himself, is to let the depreciation rate of the health capital depend on the age of the agent. Under this assumption, an agent will face an increasingly higher depreciation rate in each subsequent period of life, allowing the model to replicate an aging process under which it is increasingly costly for an agent to maintain the same level of health. As mentioned above, an age-dependent depreciation rate also provides a way of ensuring that agents choose a finite life. There are however two disadvantages with this approach. First, Dalgaard and Strulik (2014) point out that the form of the function is unknown and that an incorrect specification would lead to an inaccurate representation of human aging. Second, the problem of a healthy agent losing more health than an unhealthy one of the same age remains.

Grossman (1972a) mentions the option of making the depreciation rate of health capital a function of current health state. An alternative solution is proposed by Dalgaard and Strulik (2014), who suggest using a model of health deficit rather than the health accumulation found in Grossman (1972a). The model of health deficit builds on the work by Gavrilov and Gavrilova (1991), who introduced the reliability theory concerned with failure rates from engineering into biology. According to this reasoning, a person is born with more functioning capacity than is needed for survival. As the person ages, this redundancy gradually declines, allowing the model by Gavrilov and Gavrilova (1991) to replicate a positive relationship between aging and death rates.

In the model by Dalgaard and Strulik (2014), the health of an individual is the best obtainable health less the health deficit. Taking inspiration from the work on health deficits by Mitnitski et al. (2002), they arrive at a formulation for the change in health from one moment to the next defined as

$$\dot{H}(t) = \mu E - \mu(\bar{H} - H(t)) \tag{2}$$

where μ is the force of aging, E is the health investment, $\bar{H}$ is the best attainable health and H is current health.

In contrast to the law of motion specified by Grossman (1972a) in equation 1, Dalgaard and Strulik (2014) build a model in which the lost health is smaller for a healthier individual. To see this, consider equation 2. The health is positively impacted by investment

E. This positive effect is reduced by the difference between the best attainable health ($\bar{H}$) and current health (H). If the agent enjoys the best attainable health, this difference is zero and they subsequently capture the full potential of the investment, adjusted for the force of aging. If the agent is very unhealthy, this difference will be large and the positive effect of the investment will be largely or fully reduced.

The Deterioration of Health and Socio-Economic Status

The Grossman model has further been criticized for failing to predict that health declines with lower socio-economic status, as shown by Case and Angus (2005). Using data on self-reported health status broken down by age and by income quartile, they show that the health of the bottom quartile is significantly worse than that of the top quartile. They also show that the health of the bottom quartile declines faster with age. Interpreting these findings as evidence for the effect of manual labor on health, Case and Angus (2005) suggest that the depreciation rate of health capital in the original Grossman model does not sufficiently capture external, non-biological components in the process of declining health.

Building on the work by Muurinen and Le Grand (1998), they point out that an individual has three types of capital - health capital, human capital, and physical capital. Considering these as substitutes in ways to earn an income, Case and Angus (2005) reason that an agent with little education (human capital) and little assets (physical capital) may to a larger extent be forced to depend on their health capital by taking strenuous manual jobs avoided by the educated and the wealthy. As a result, the non-biological part of the depreciation of health capital is expected to be higher for these individuals, explaining the faster decline in health over the life cycle. To capture this effect, they propose adjusting the Grossman model to make earnings a function of health capital and physical effort. They also make depreciation rate a function of physical effort, allowing the agent to increase their income at the expense of a faster decline their health capital.

In contrast, Laporte (2015) argues that the Grossman model is capable of producing a scenario where the health status of higher-income agents will decline more slowly than that of poorer agents. Specifying a simplified Grossman model in which longevity has an exogenous upper bound and where income is constant and independent of health, she shows that the higher-income individual faces a lower opportunity cost of investing in health, caused by the fact that a higher income implies that more consumption can be enjoyed for any given level of health investment. Given the differences in opportunity costs, the poorer agent is expected to invest less in their health and consequently suffer a faster deterioration.

2.2.2 Empirical Criticism

The accuracy of the Grossman model has been tested by transforming the theoretical model into an empirical model and evaluating how well it is able to predict observed trends. As pointed out by Wagstaff (1993), there are two possible reasons for why empirical tests of the Grossman model leads to outcomes that contradicts the data. The transformation from a theoretical to an empirical model has either introduced incorrect assumptions, meaning that the test itself is flawed, or the transformation is correct and the theoretical model needs to be adjusted to better match the data. Consequently, only the latter should be regarded as criticism against the theoretical Grossman model.

Health Investment and Current Health Status

Zweifel, Breyer and Kifmann (2009) criticize the positive relationship between health status and health investment predicted by Grossman (1972b)'s empirical formulation, pointing out that this relationship consistently is found to be negative. For example, He and Huang (2022) find that life expectancy and medical expenditures move in opposite directions over the business cycle.

The original empirical formulation of the Grossman model circumvents the problem of unobservable health capital by estimating reduced-form demand functions for medical care. Wagstaff (1986) compares this approach with the use of structural equations where health capital is a latent variable, and shows that the coefficient on health is negative and statistically significant at the 90% level for equations on GP visits, hospital stays, and medicine.

Entertaining both the possibility of an inaccurate and an accurate empirical transformation, Wagstaff (1986) hypothesizes that the negative coefficient in the latter case could be caused by non-instantaneous adjustment from the actual to the optimal level of health capital. However, in a later paper, Wagstaff (1993) argues that non-instantaneous health adjustment is unlikely to explain the negative coefficient on health and instead concludes that an alternative empirical formulation is needed. In providing one, including but not fully relying on non-instantaneous health adjustment, he shows that this alternative formulation brings the results closer to the data.

After this finding, Wagstaff (1993) concludes that the empirical formulation used in previous literature is inaccurate. In other words, it is possible that the inability of the empirical Grossman model to replicate the negative relationship between health status and health investment reflects weaknesses in the empirical model, not the theoretical model. Simultaneously, Wagstaff (1993) points out that his results suggest that individuals do not adjust their health capital instantaneously, as assumed by Grossman (1972a).

Testing the implications of non-instantaneous health adjustments, T. Galama and Kapteyn (2011) relax the assumption of strictly positive health investments imposed by Grossman (1972a) to allow for corner solutions in which the investment equal zero. Under the assumption that investments cannot be negative (agents cannot purchase harmful goods to reduce their health) this creates a scenario under which an agent with initial health capital higher than the optimal level is unable to immediately move to the desired level. Instead, the agent must choose to invest nothing and wait for the gradual decline in the health stock caused by the depreciation rate. As a result of the non-instantaneous adjustment of the health stock, T. Galama and Kapteyn (2011) find that the demand for health investment is zero until health drops to the optimal level. From there on, the agent demands investments in health and follows the optimal path described by Grossman

(1972a). In other words, T. Galama and Kapteyn (2011) show that a Grossman model with non-instantaneous adjustment is able to produce a negative relationship between health and health investment, in which a healthy individual demands no medical care until the health deteriorates to sufficiently low levels.

Health Investment and Returns to Scale

In addition to the criticism reviewed in section 2.2.1, the assumption of CRTS in health investment used by Grossman (1972a) has further been criticized in the empirical tests of the Grossman model. First, T. J. Galama et al. (2012) argue that DRTS is a more realistic assumption. They also show that DRTS leads to an empirical model that is better at replicating stylized facts from the data. In particular, they show that a Grossman model with DRTS in health investment is capable of replicating the observed relationships between the deterioration of health and socio-economic status as well as the observed relationships between health investment and health status, thereby improving two of the sources of criticism discussed above.

Despite the reviewed evidence, it is not undisputed that the theoretical Grossman model should be adjusted to include DRTS instead of CRTS in health investment. When using IV methods to control for endogeneity in health, T. J. Galama et al. (2012) find that their coefficients become statistically insignificant. Given that their results when not accounting for endogeneity are in line with previous literature, and that this literature do not take endogeneity into account, this finding leads T. J. Galama et al. (2012) to question the robustness of previous empirical tests of the relationship between health status and health investment.

2.3 RBC and Endogenous Health

To the best of our knowledge, the literature on RBC models with endogenous health accumulation is limited to five papers. Each paper builds on the Grossman model in their conceptualization of health, with the common deviation from Grossman (1972a) that utility depends on the health capital instead of healthy time. Apart from this common feature, the existing models differ in terms of simplifying assumptions and focus.

First, Laun (2012) builds a model to investigate the interaction between economic fluctuations and health. To this end, the model includes both shocks to productivity and to health, where the latter affects available time via the transformation of health capital and well-being through the utility-generating health capital. In addition to health capital, the utility function depends on consumption and leisure, the only activity besides work that is available to the agent. The health stock depreciates over time and can be maintained using investments. Unlike Grossman (1972a), health investments do not use specific inputs. After calibrating the model to Swedish data, Laun (2012) finds that an increase in productivity leads to higher savings and better health, while improved health leads to higher savings.

With slightly different focus, Vasilev (2017) examines whether health shocks is a source of business cycles. In contrast to Laun (2012), health shocks is introduced as uncertain outcomes of the health investment. Unlike shocks to aggregate productivity, which shocks the supply side of the economy, Vasilev (2017) reasons that health shocks affect both the supply and the demand sides. In the model, the agent derives utility from consumption, leisure, and their health capital. A shock to the latter may therefore allow the agent to change their consumption and still maintain the same level of utility. In addition, the health capital stock is transformed into healthy time, following the Grossman model. Since healthy time is allocated between work, leisure, and exercise, an agent may respond to a shock to the health capital by reallocating time between these activities, thereby altering their labor supply. Unlike Grossman (1972a), Vasilev (2017) assumes that investments in health only require exercise time.

By calibrating the model to US data, Vasilev (2017) presents impulse response functions to an increase in the TFP and to an increase in the productivity of health investment, both effects considered separately. He finds that an increase in the TFP leads to a decrease in exercise time and consequently a decline of health capital and healthy time in the subsequent period. According to Vasilev (2017), the decrease in exercise time is caused by the perfect substitutability between time spent on work, exercise, and leisure. As productivity and wage rate increase, the agent responds by increasing work hours at the expense of time spent exercising. When marginal productivity and wage rate starts to fall over time, the agent begins to shift their time back towards exercise, temporarily increasing it above pre-shock levels before returning to steady-state. Since health investments affect health status according to the law of motion of health capital provided by Grossman (1972a), shown in equation 1, health capital and healthy time follow the response of exercise time with a lag.

In contrast to the aggregate productivity shock, Vasilev (2017) finds that an individual responds to a positive health shock by drastically increasing their exercise time. The effect is short-lived and soon falls below steady-state before returning to pre-shock levels. Despite this, health capital never falls below steady-state levels - it increases in the periods following the shock and then falls back to steady-state. When comparing the effect of the two shocks, Vasilev (2017) concludes that the effect of the health shock is substantially lower than that of the productivity shock. The effect of the former is so small that he considers the aggregate output and consumption to be unchanged. In addition, the health shock is less persistent than the productivity shock. For these reasons, he concludes that health shocks are unlikely to be an important driver of business cycles.

Unlike Vasilev (2017), the models by He and Huang (2022), Feng and Gomis-Porqueras (2011), and He and Hung (2011) follow Grossman (1972a) more closely in the production of health, using both medical expenses and leisure time as inputs in the investment function. The latter does however not depend on health state - the agent is endowed with an exogenous amount of available time, which is then divided between work and leisure. The effect that health plays on labor supply in He and Hung (2011) and He and Huang (2022) is instead modeled with a health-dependent production function. This leads to a slightly different interpretation on why health matters for the income of an agent - in Grossman (1972a) and Vasilev (2017), a healthy agent can provide more work hours, whereas a healthy agent in He and Hung (2011) and He and Huang (2022) is more productive.

The purpose of the paper by He and Huang (2022) is to examine possible causes for the observed cyclical fluctuations of national expenditures and life expectancy. Unlike Vasilev (2017), they neither model health shocks nor provide impulse response functions. Building on the work by Ruhm (2000), they argue for three main channels through which endogenous health matters for the observed behaviors - the utility channel (health is valued by the individual), the production channel (good health increases productivity but also depreciates with production), and the time channel (investments in health require leisure). Based on their results, they conclude that all three channels are important for matching the model to the cyclical fluctuations of health observed in the data. Using an identical model, the relative importance of each channel is examined by He and Hung (2011). They find that the utility channel is the most important, the production channel the second most important, and that the time channel is the least important.

The importance of the utility channel is emphasized in the model by Feng and Gomis-Porqueras (2011). In their model, health determines neither available time nor affects productivity, but matters only as a consumption good in the utility function. Using this set-up, they seek to examine the impact of including endogenous health in an RBC model. They find that this model performs better than a standard RBC model in terms of predicting the observed variability in consumption and labor supply than a standard RBC model. They hypothesize that this improvement is driven by the need to smooth both consumption and health, leading to more volatile investments.

3 Model

We build an infinitely lived agents model in which the agent has preferences over consumption, leisure and healthy time. Healthy time is a function of the endogenous health status, which depreciates over time. The depreciation rate is in turn a decreasing function of the health status. The decline in health can partially or fully be offset by investing in the health capital using medical care and exercise time.

In each period, the agent optimally allocates their wealth between consumption, capital investment, and medical care. In addition, they allocate their healthy time between work, leisure, and exercise. The economy is subject to aggregate productivity shocks and the agent faces health shocks. For simplicity, the economy is assumed to consist of a single individual.

3.1 Production

In any period t, the agent produces an aggregate consumption good according to a standard Cobb-Douglas production function by combining capital (k_t) and labor (n_t) according to equation 3.

$$A_t f(k_t, n_t) = A_t k_t^{\eta} n_t^{1-\eta} \tag{3}$$

The productivity of the agent is given by the term A_t and depends on the state of the world. The economy draws a new state in each period t where the probability of the realization of a given state depends on the realization in the previous period $t-1$. This stochastic process is discussed in more detail in section 3.5.

Capital consists of aggregate goods produced in previous periods saved by the agent and is assumed to depreciate at the constant rate $\delta \in (0,1)$. There is no market for borrowing or lending. Consequently, we have the non-negativity constraint and the law of motion for capital presented in equations 5 and 4,

$$k_{t+1} = (1-\delta)k_t + i_t \tag{4}$$

$$k_{t+1} \geq 0 \tag{5}$$

where i_t denotes the investment in new capital in period t.

3.2 Health capital

Following Grossman (1972a), we conceptualize health as a capital stock, denoted by H_t. The health stock evolves over time according to the law of motion shown in equation 6,

$$H_{t+1} = (1-\rho(\epsilon_t))H_t + g(m_t, e_t) \tag{6}$$

where $\rho(\epsilon_t)$ is a function for the depreciation rate of the health capital and $g(m_t, e_t)$ is a function for the production of new health. In Grossman (1972a), agents die when their health stock falls below a minimum level. Since we assume an infinitely lived agent, we do not specify a minimum level of H_t below which the agent exits the economy. Health capital is therefore permitted to be an element of the non-negative real numbers, $H_t \in \mathbb{R}_+$.

The health stock depreciates over time according to the depreciation rate $\rho(\epsilon_t)$. In the original model by Grossman (1972a), the depreciation rate of the health capital is an increasing function of age. As discussed in section 2.1, this specification ensures that an agent in a finitely lived agents model chooses a finite life, but fails to correct the counter-intuitive positive relationship between health status and health decline.

In order to bring the model closer to the observed negative relationship, we assume that $\rho(\epsilon_t)$ is a decreasing function of the health capital. We also make the assumption that the depreciation rate is a decreasing function of an i.i.d. stochastic component to incorporate health shocks in our model.

Let $X_H \subseteq \mathbb{R}_+$ denote the set of all possible values for H, and let $Z_\epsilon \subseteq (\underline{\epsilon}, \bar{\epsilon})$ denote the set of all possible values for ϵ. We want the function $\rho : X_H \times Z_\epsilon \mapsto D$ to be decreasing in H_t and ϵ_t, bounded below by 0 and above by 1. Based on these considerations, we assume that the function takes the form of equation 7,

$$\rho(\epsilon_t, H_t) = \tau + \exp(-\phi(H_t + \kappa) \times \exp(\lambda\epsilon_t)) \tag{7}$$

where τ is the lower limit of the function, ϕ is the growth rate, κ shifts the function, λ regulates the impact of the shock and ϵ_t represents the health shock. Note that we will usually write $\rho(\epsilon_t)$ to save on notation, but it should be remembered that the depreciation rate is also an explicit function of the health capital. We keep the explicit dependence on the shock in our notation however, because it will make it more convenient when we later define probability measures and the value function.

As shown in equation 7, an unhealthy agent will face a higher depreciation rate than a healthier one. This can be thought of as a sick individual seeing a faster decline of their health due to an already compromised immune system. Because we model an infinitely lived agent without any focus on decisions over the life cycle, we do not make $\rho(\epsilon_t)$ a function of age. One can however imagine a combination of the two approaches in a life cycle model, where the depreciation rate increases with age but falls with health status.

The depreciation rate of health also captures the idiosyncratic health shock. Unlike Vasilev (2017), who assumes that health shocks enters the model as uncertain outcomes of health investments, and Laun (2012), who assumes that health shocks affect utility and available time directly, we shock the rate at which health deteriorates from one period to the next. This can be envisioned as an individual who catches a disease or experiences an accident and subsequently face a sharper decline in their health from one period to the next. Note that this approach indirectly captures the effect of the health shock on both utility and time as modeled by Laporte (2015), via the specified utility function in equation 11 and the transformation of health capital into healthy time in equation 9. The stochastic process of the health shock is described in detail in section 3.5.

In each period, the stock of health capital can be maintained using investments. Following Grossman (1972a), investments in health can be made using expenditures on medical care (m_t) and the use of exercise time (e_t), where medical care is composed of the aggregate good produced according to equation (3). For simplicity, it is assumed that one unit of the aggregate good gives one unit of medical care. Our model therefore differs from the previous RBC models with endogenous health accumulation discussed in section 2.3, with Vasilev (2017) using only exercise in the production function and He and Hung (2011), He and Huang (2022), and Feng and Gomis-Porqueras (2011) all using medical care and leisure. A further discussion on our decision to separate leisure and exercise is given in section 3.3.

Like Feng and Gomis-Porqueras (2011), we assume that new health capital is produced according to a CES production function, shown in equation 8,

$$g(m_t, e_t) = B(am_t^\theta + (1-a)e_t^\theta)^{\frac{1}{\theta}} \tag{8}$$

where B denotes productivity in health production.

3.3 Time

The total amount of time, T, in a given period is divided between healthy time and sick time, where healthy time is a function of the health capital in line with Grossman (1972a). For any level of health capital, the total amount of healthy time enjoyed by the agent is given by the function $h_t = \psi(H_t)$. The function $\psi : \mathbb{R}_+ \to [0, T)$ is assumed to be continuous, differentiable and strictly increasing. Given these assumptions on ψ, we choose an exponential function as our mapping from health capital to healthy time,

$$h_t = \psi(H_t) = T[1 - \exp(-gH_t)] \tag{9}$$

where g is the growth rate of the function.

In other words, we deviate from the linear function assumed by Grossman (1972a). This is motivated as a more realistic choice, given that a one unit improvement in health is reasoned to bring more benefit to a very sick individual than to a very healthy one.

Healthy time can be spent on activities according to equation 10.

$$h_t = n_t + l_t + e_t \tag{10}$$

Unlike previous RBC models with endogenous health reviewed in section 2.3, with the exception of Vasilev (2017), we assume that leisure and exercise are two separate activities. Whereas the models by He and Hung (2011), He and Huang (2022), and Feng and Gomis-Porqueras (2011) all assume that leisure is both health- and utility-generating, Colman and Dave (2013) point out that exercise is likely to be utility-reducing given that only 5% of Americans get the recommended amount of daily physical activity[1].

3.4 Preferences

The agent derives utility from consumption (c_t), leisure (l_t), and healthy time (h_t). Because healthy time is the total available time T less sick time, the utility function implies that the agent derives disutility from being sick. Furthermore, given that leisure is the

[1] A possible extension to our model would be to introduce habit formation. Rather than assuming that all agents derive disutility from exercising, as is implied in our model, it may be more realistic to assume that agents who exercise often will over time derive less disutility, or even utility, from this activity.

healthy time remaining after work and exercise according to equation 10, the agent derives disutility from working and from exercising.

We assume the additively separable period utility function shown in equation 11,

$$u(c_t, l_t, h_t) = \frac{c_t^{1-\alpha} - 1}{1-\alpha} + \log(l_t) + \log(h_t) \tag{11}$$

where future utility is discounted at the rate $\beta \in (0, 1)$.

3.5 Stochastic processes

We have already introduced the assumptions that the productivity parameter in the production function for consumption and medical goods, and the depreciation rate for the health stock are stochastic variables. In this section, we describe these variables in greater detail.

Both A_t and ϵ_t are assumed to be realized at the beginning of the period. Let $Z_A \subseteq \mathbb{R}_{++}$ be the (bounded) set of all possible values for A_t and let $Z_\epsilon \subseteq (\underline{\epsilon}, \bar{\epsilon})$ be the set of all possible values for ϵ_t. $(Z_A, \mathcal{Z}_A)$ and $(Z_\epsilon, \mathcal{Z}_\epsilon)$, where $\mathcal{Z}_A$ and $\mathcal{Z}_\epsilon$ denote the Borel σ-algebras of subsets of Z_A and Z_ϵ respectively, are then measurable spaces (Stokey, Lucas Jr and Prescott, 1989, pp. 168–169). We denote the stochastic state variables in period t by $z_t = (A_t, \epsilon_t) \in Z$, where $Z = Z_A \times Z_\epsilon$.

In section 3.1 we also introduced the assumption that the stochastic process $\{A_t\}_{t=0}^{\infty}$ exhibit persistence. In particular, the natural logarithm of productivity is assumed to follow an AR(1) process

$$\log(A_{t+1}) = \gamma \log(A_t) + \xi_{t+1} \tag{12}$$

where ξ_{t+1} is assumed to be i.i.d with the marginal density function $p_\xi(\xi_{t+1})$. In order to ensure that A_{t+1} is bounded, we will further assume that ξ_{t+1} is bounded from above. With these assumptions, we may define a transition function $\mu_A : \mathcal{Z}_A \rightarrow [0, 1]$ such that for each $A_t \in Z_A$, $\mu_A(A_t, \cdot)$ is a probability measure on $(Z_A, \mathcal{Z}_A)$ with $\mu(a, B) = Pr\{A_{t+1} \in B | A_t = a\}, B \in \mathcal{Z}_A$ (Stokey, Lucas Jr and Prescott, 1989, pp. 210–212). The stochastic component ϵ_t of the depreciation rate for health capital $\rho(\epsilon_t)$, defined in equation (7), is on the other hand not assumed to explicitly depend on the realization in the past period, but is assumed to be i.i.d with the marginal density function $p_\epsilon(\epsilon_t)$. We further assume that A_{t+i} and ϵ_{t+j} are independent for all i and j. Because of this independence assumption, the joint probability of observing the pair $(A_{t+1}, \epsilon_{t+1})$, conditional on the state in period t, (A_t, ϵ_t), is then given by $\mu(z_t, z_{t+1}) = \mu_A(A_t, A_{t+1}) p_\epsilon(\epsilon_{t+1})$.

3.6 The Maximization Problem

A rational agent faces the problem of choosing sequences $\{y_t\}_{t=0}^{\infty} = \{c_t, i_t, m_t, n_t, l_t, e_t\}_{t=0}^{\infty}$ of consumption, capital investment, medical care utilization, labor supply, leisure, and exercise. Since the stochastic state variables are assumed to be realized at the beginning of each period, the agent knows the productivity and the depreciation rate when making their decision.

In each period, the agent faces a budget constraint and a time constraint. Given that the agent enters a period with a certain level of k_t and H_t, and given that the agent is rational and will waste no resources, the total value of c_t, i_t, and m_t must equal the output given by equation (3) in the same period. Furthermore, it is neither possible for the agent to consume negative amounts, nor to utilize negative amounts of medical care. Thus, the budget constraint and non-negativity constraints the agent faces are:

$$c_t = A_t f(k_t, n_t) - i_t - m_t \tag{13}$$

$$c_t, m_t \geq 0 \tag{14}$$

There are T units of time available in each period and given the health stock of the agent, healthy time is given by equation (9). Time lost due to sickness is then $s_t = T - h_t$. Healthy time is what remains of the total time endowment after sick time has been accounted for, and the agent allocates the remaining time between working (n_t), leisure (l_t), and producing new health capital (e_t) according to equation 10. Since the agent is assumed to be rational and hence will waste no time, the time constraint and non-negativity constraints that must hold in each period are:

$$h_t = n_t + l_t + e_t, \tag{15}$$

$$n_t, l_t, e_t \geq 0 \tag{16}$$

Succinctly stated, the maximization problem the agent solves is:

$$\max_{\{y_t\}_{t=0}^{\infty}} \quad E_t \sum_{t=0}^{\infty} \beta^t u(c_t, l_t, h_t) \tag{17}$$

subject to

$$\begin{aligned}
c_t &= A_t f(k_t, n_t) - i_t - m_t, & t &= 0, 1, 2, ... & (18)\\
h_t &= n_t + l_t + e_t, & t &= 0, 1, 2, ... & (19)\\
k_{t+1} &= (1-\delta)k_t + i_t, & t &= 0, 1, 2, ... & (20)\\
H_{t+1} &= (1-\rho(\epsilon_t))H_t + g(m_t, e_t), & t &= 0, 1, 2, ... & (21)\\
h_t &= \psi(H_t), & t &= 0, 1, 2, ... & (22)\\
c_t, m_t, k_{t+1}, n_t, l_t, e_t &\geq 0, & t &= 0, 1, 2, ... & (23)\\
k_0, H_0, A_0 &\quad \text{given} & & & (24)
\end{aligned}$$

We adopt the strategy of writing the problem in such a way that the state and control variables in period t do not enter the transition equations for the state variables in period $t+1$ by making the latter choice variables in themselves in period t. In particular, we rewrite equations 6 in terms of m_t and equation 4 in terms of i_t as follows:

$$\begin{aligned} m_t = G(I_t, e_t) &= \left(\frac{1}{a} \left[\left(\frac{H_{t+1} - (1-\rho(\epsilon_t))H_t}{B} \right)^{\theta} - (1-a)e_t^{\theta} \right] \right)^{\frac{1}{\theta}} \\ i_t &= k_{t+1} - (1-\delta)k_t \end{aligned} \tag{25}$$

The expressions for m_t and i_t are then substituted into the budget constraint shown in equation 13. Finally, we define the new function F as the sum of output and the capital that remains after depreciation according to:

$$F(k_t, n_t) = A_t f(k_t, n_t) + (1-\delta)k_t \tag{26}$$

After rewriting the constraints, we obtain the value function shown in equation 27. Note that the time subscripts have been dropped in this reformulation of the problem. Instead, we adopt the convention of denoting an arbitrary variable today by x and the same variable tomorrow by x'.

$$V(x,z) = \sup_{y \in \Gamma(x,z)} \left\{ u(c,l,h) + \beta \int_Z V(x',z')\mu(z,z')dz' \right\} \tag{27}$$

where

$$\begin{aligned} x &= (k, H), \\ z &= (A, \epsilon), \\ c &= F(k,n) - G(I,e) - k' \\ y &= (k', H', n, e), \\ I &= H' - (1-\rho(\epsilon))H, \\ h &= \psi(H) \\ l &= \psi(H) - n - e, \end{aligned} \tag{28}$$

Let $X = X_k \times X_H$ denote the set of possible values for the pair (k, H), and let $S = S_n \times S_e$ denote the set of possible values for the pair (n, e). We may then define the sublevel set $B(G)(a)$ as

$$B(G)(a) = \{(H', e) \in X_H \times S_e : G(H' - (1-\rho(\epsilon))H, e) \leq a\} \tag{29}$$

The feasible set correspondence $\Gamma(x,z) : X \times Z \twoheadrightarrow X \times S$ is then given by

$$\begin{aligned}
\Gamma(x,z) = \{y \in X \times S : & \\
(H', e) &\in B(G)(F(k,n) - k') \\
I &\in [0, g(F(k,n) - k', e)] \\
k' &\in [0, F(k,n)] \\
n, e &\in [0, \psi(H)] \\
n + e &\in [0, \psi(H)]\}
\end{aligned} \tag{30}$$

4 Calibration

4.1 Stochastic Processes

A_t is approximated using a discrete state Markov process, following the methods developed by Tauchen (1986), where the probability of the realization of a given state in period t depends on the realization in the previous period $t-1$. In order to approximate the continuous distribution for A_{t+1} by a discrete state Markov process, we assume for simplicity that $\xi_{t+1} \sim N(0, \sigma_\xi^2)$, because the likelihood that we will have issues with ξ_{t+1} taking on values close to infinity in simulations is likely to be close to zero.

In each period, the realization of A_t may take values in the set

$$A_t \in Y_A = \{A_t^1, A_t^2, ..., A_t^w\} \tag{31}$$

where $A_t^1 < A_t^2 <, ..., < A_t^s$.

Let Θ_A denote the Markov matrix with elements $\vartheta_{i,j}^A$, $i,j \in \{1,2,...,w\}$, describing the probability of transitioning from a productivity level A_{t-1}^i yesterday to a productivity level A_t^j today.

$$\Theta_A = \begin{bmatrix} \vartheta_{1,1}^A & \vartheta_{1,2}^A & . & . & . & \vartheta_{1,w}^A \\ \vartheta_{2,1}^A & \vartheta_{2,2}^A & . & . & . & \vartheta_{2,w}^A \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ \vartheta_{w,1}^A & \vartheta_{w,2}^A & . & . & . & \vartheta_{w,w}^A \end{bmatrix} \tag{32}$$

Since the elements of the matrix Θ_A are probabilities, we require that $\vartheta_{i,j}^A \geq 0$, for all $i,j \in \{1,2,...,w\}$. We also require that each row of the matrix sum to 1, that is, $\vartheta_{i,1}^A + \vartheta_{i,2}^A + ... + \vartheta_{i,w}^A = 1$, all $i \in \{1,2,...,w\}$.

Unlike the productivity shocks, the realization of the error term in the depreciation rate of health capital is assumed to be independent of the realization in the previous period. In order to make it easier to define a transition matrix specifying the probability of ending up in a certain combination of productivity and health states, we adopt the strategy of approximating also the health shock using a discrete state Markov process by assuming, for similar reasons as above, that $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. The difference is that we set the persistence to zero.

In each period, the realization of ϵ_t may take values in the set

$$\epsilon_t \in Y_\epsilon = \{\epsilon_t^1, \epsilon_t^2, ..., \epsilon_t^s\} \tag{33}$$

where $\epsilon_t^1, \epsilon_t^2, ..., \epsilon_t^s \in (0,1)$ and $\epsilon_t^1 < \epsilon_t^2 <, ..., < \epsilon_t^s$.

Consequently, we have the following Markov matrix, where the absence of persistence implies that all values in a given column are identical.

$$\Theta_\epsilon = \begin{bmatrix} \vartheta_{1,1}^\epsilon & \vartheta_{1,2}^\epsilon & . & . & . & \vartheta_{1,s}^\epsilon \\ \vartheta_{2,1}^\epsilon & \vartheta_{2,2}^\epsilon & . & . & . & \vartheta_{2,s}^\epsilon \\ . & . & . & & & . \\ . & . & & . & & . \\ . & . & & & . & . \\ \vartheta_{s,1}^\epsilon & \vartheta_{s,2}^\epsilon & . & . & . & \vartheta_{s,s}^\epsilon \end{bmatrix} \tag{34}$$

Given that the stochastic process for ϵ_t is assumed to be independent from the shocks to the economy, the conditional probability of observing the pair $(A_{t+1}, \epsilon_{t+1}) = (a, b)$ given that the state today is $(A_t, \epsilon_t) = (c, d)$, is simply the product of the transition probabilities given by the Markov matrices,

$$Pr\{(A_{t+1}, \epsilon_{t+1}) = (a, b) | (A_t, \epsilon_t) = (c, d)\} = \vartheta_{c,a}^A \vartheta_{d,b}^\epsilon$$

Due to the curse of dimensionality, we are limited to solving the model with only three possible values for A_t and ϵ_t respectively - they may both take a low, medium or high value. The variances σ_ξ^2 and σ_ϵ^2, the persistence parameter γ in equation (12) and the parameters τ, ϕ, κ and λ in equation (7) are chosen such that all elements in Y_A and Y_ϵ are finite, and such that $\rho(\epsilon_t) \in D \subseteq (0,1)$ for all H_t and ϵ_t.

4.2 Exogenous Parameters

Given that the purpose of this paper is to present an RBC model with endogenous health accumulation that captures dynamics missed by previous contributions, the calibration mainly seeks to replicate the general direction of key relationships emphasized by the empirical literature. After solving the model using value function iteration and the calibration presented in Table 1, we find that the model successfully predicts most of these observed behaviors. However, the results should be interpreted with the caveat the we are restricted to using only 15 grid points for the state and control variables when we solve for the value and policy functions due to the high dimensional state space in our model.

First, the criticism directed at Grossman (1972a) for predicting a positive relationship between health status and investments in health is discussed in section 2. In response to this criticism, we assume that the depreciation rate of health is a decreasing function of the health capital, shown in equation 7. As a result, a healthy individual needs less resources to maintain their level of health. After obtaining the policy function specifying the optimal choice of medical expenditures for each combination of productivity, realization of the health shock, capital and health capital, we find that the optimal medical expenditure is indeed decreasing in health capital. For example, when both productivity and the health shock take their intermediate values, whilst capital is fixed at the level which is used as the starting point for the time paths generated by our impulse response functions in section 5.3, the correlation between medical expenditures and health capital is found to be -0.8464. In other words, an agent making optimal decisions will use less medical care as they get increasingly healthier. Evaluated at the same point, the correlation between the health capital and the optimal choice of exercise, the second input in the production of health in our model, is found to be -0.9172.

As discussed in section 2, the data further shows that medical expenditures vary procyclically, whereas time spent on physical activities varies counter-cyclically. When evaluating the policy functions at what appear to be locally stable levels of capital and health capital, which are used as points of departure when we generate impulse response functions in section 5.3, we find that the optimal amount of medical expenditures increase whereas the optimal time spent on exercise decreases when moving from a low productivity state to a high productivity state. As mentioned before, we are limited to solving the model with only three productivity states - low productivity, medium productivity, and high productivity. As a result, we are carefully optimistic about the ability of the model to replicate the effects of a productivity shock on optimal medical expenditures and exercise time, but realize that the model presents a highly simplified representation of business cycles.

Finally, He and Huang (2022) argue that a good model must be able to predict the counter-cyclical life expectancy seen in the data. Most famously, Ruhm (2000) shows that mortality is lower in recessions. Since our infinitely lived agents model does not account for life expectancy, we consider the health stock instead. When examining the policy function for health capital in the next period, given a certain capital and health capital state today, and comparing the optimal choice in each productivity state, we find no effect. This feature of our model is discussed in more detail in section 5.3. Even though this contradicts the findings by Ruhm (2000), mortality represents only one of the commonly used proxies for health. In fact, after providing an extensive review of empirical literature on the relationship between health measures and business cycles, Bellés-Obrero and Vall Castelló (2018, p. 1) conclude that "[t]he only well-established finding is that mental health deteriorates during economic slowdowns".

As mentioned, the model is calibrated to predict the general direction of key relationships. In other words, the model is not calibrated to the data of a specific country and period, meaning that the subsequent analysis and discussion in section 5 is limited to general trends, subject to potentially significant error margins. For example, our model estimates

Parameter	Value	Description
a	0.5	Share of m_t relative e_t in health production
α	2	Relative risk aversion
B	1	Health production productivity
β	0.9722	Discount rate
δ	0.3	Depreciation rate of capital
η	0.72	Capital production output elasticity
g	0.1	Growth rate of ψ
γ	0.586	Persistence of productivity shock
κ	0.15	Shift of ρ function
ϕ	0.4	Growth rate of ρ
λ	0.3	Impact of health shock
T	3	Total time
σ_ξ	0.082	Standard deviation of productivity shock
σ_ϵ	0.4	Standard deviation of health shock
τ	0.05	Lower limit of ρ
θ	0.5	Health production output elasticity

Table 1: Calibration

that the time spent working is just below 50% of total available time. In contrast, we would expect that a full-time employee spends approximately one third of their time working. In order to compare the results of our model to that of previous models, and in order to use our model for policy evaluations, a more detailed calibration is needed. We leave this venture for future research.

5 Main Results

5.1 Value Function

The value function at the intermediate levels of productivity and the health shock is presented in figure 1. In order to facilitate the analysis, we also present the value of having different levels of capital (k_t) given a fixed level of health capital (H_t) in figure 2.1. Similarly, figure 2.2 shows the value of having different levels of health capital given a fixed level of capital.

As expected, the value function is increasing monotonously in both capital and health capital. When comparing the relative importance of capital and health capital for the obtained value, one can see that the additional value gained from increasing health capital, holding capital fixed in figure 2.2, is greater than the additional gain from increasing capital holding health capital fixed in figure 2.1.

This result can be explained by the different uses of capital and health capital. While capital can be used for consumption, which is utility-generating, and medical expenses, which is used for health investments, the health capital is transformed into healthy time. Healthy time is in turn both directly utility-generating and can be used for utility-generating leisure.

Figure 1: Value Function Under No Shocks

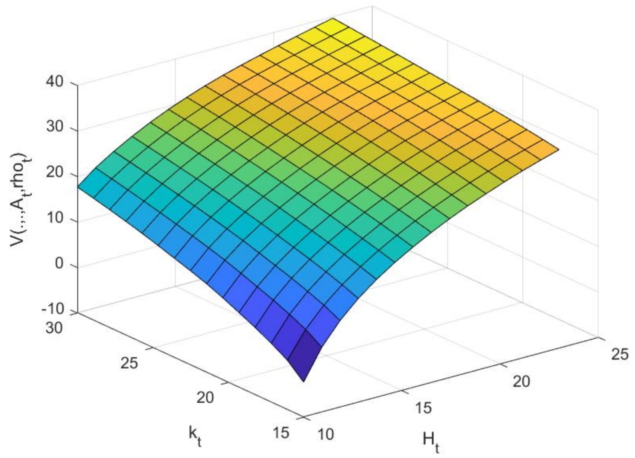
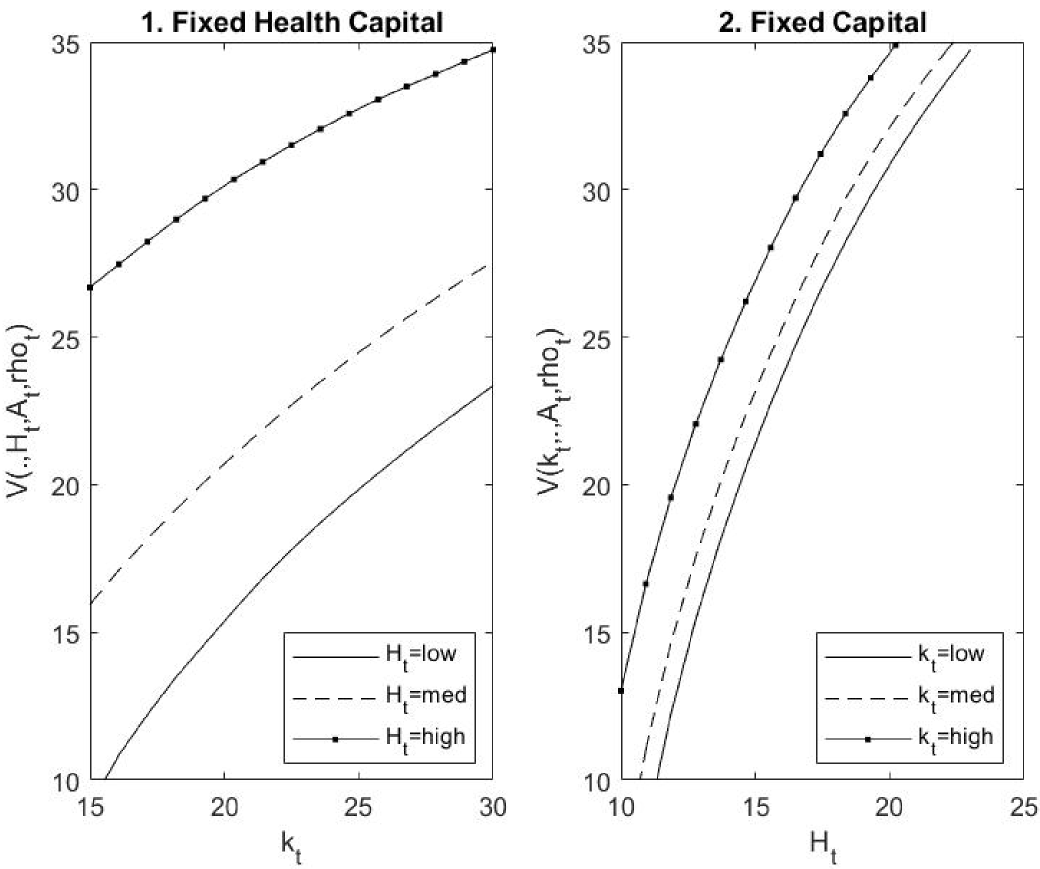
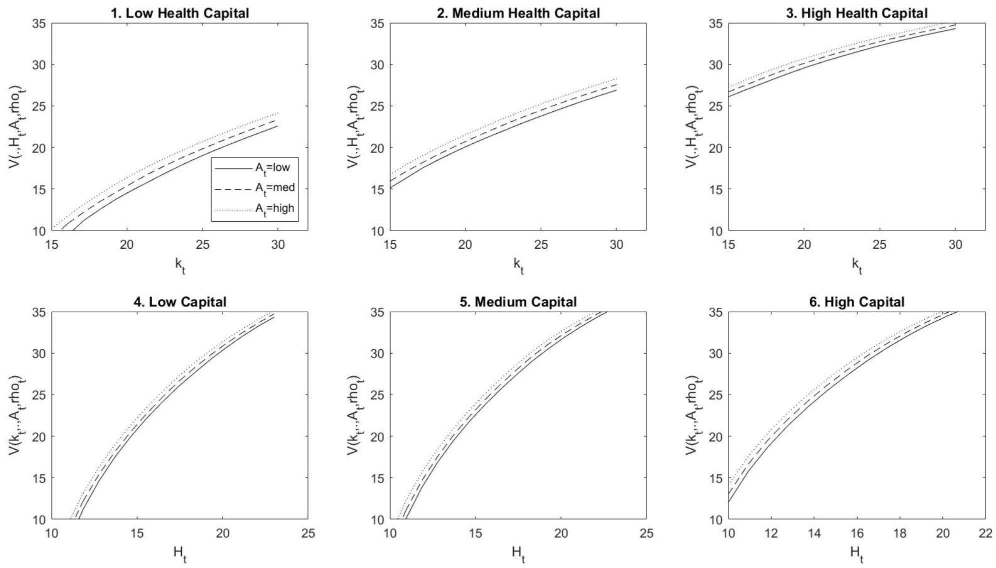


Figure 2: Value Function Holding Capital Stocks Fixed



After analyzing the value function with both shocks fixed at their intermediate levels, we turn the analysis to the value function under productivity shocks, where the realized level of productivity may take either a low, medium or high value. The value for different levels of capital given each realization of the shock is shown for fixed levels of health capital in figures 3.1-3. Figures 3.4-6 show the value associated with different levels of health capital given each realization of the shock, for fixed levels of capital.

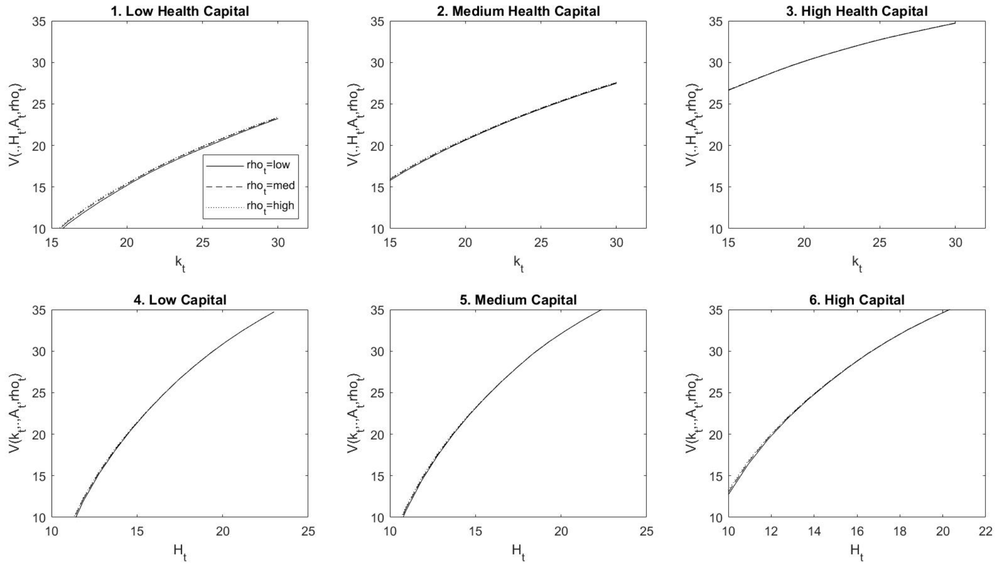
Figure 3: Value Function Over Productivity States



Figures 3.1-3 show that the distance between the curves shrinks as health capital is held fixed at an increasingly higher level. This implies that for any current capital state, the productivity shock will have a larger effect on the value function given a low health capital compared to the same productivity shock to an agent with a higher level of health capital. A reverse but smaller effect is observed when holding the capital stock fixed in figures 3.4-6 - the productivity shock has a larger impact on the value function when capital is fixed at a high level. In other words, the impact of the shock depends on the current combination of capital and health capital states.

Finally, we consider the effect of a health shock. Like the productivity shock, the health shock is assumed to take either a low, medium, or high value. The shock enters the function of the depreciation rate for health capital according to equation 7. As such, a positive shock reduces the depreciation rate whereas a negative shock increases the deprecation rate. Rather than analyzing the effect of the stochastic component ϵ_t of the depreciation rate for health capital in isolation, we consider the effect of the full term $\rho(\epsilon_t)$, which shifts with the shock.

Figure 4: Value Function Over Health States

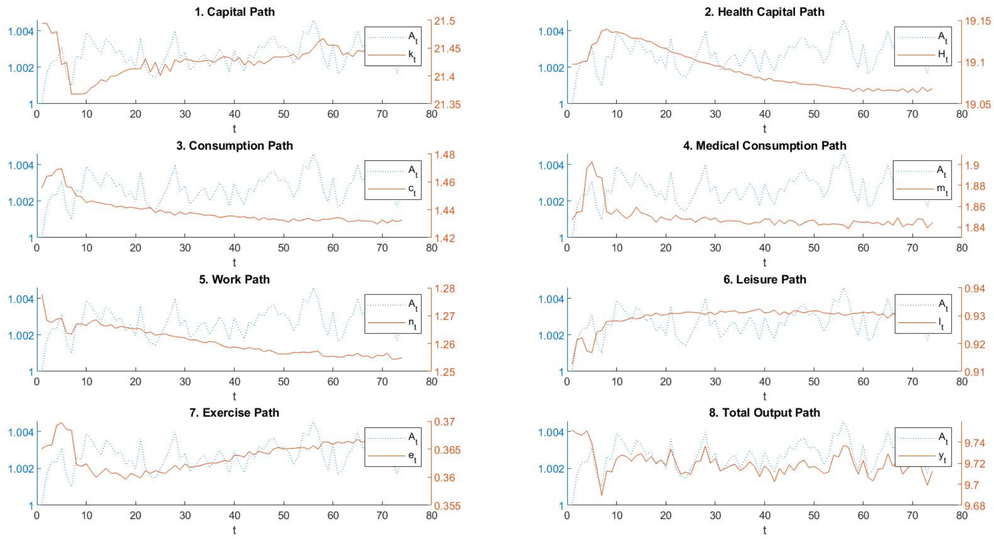


For the majority of figure 4, the effect of the health shock is too small to observe. This is not surprising given the combination of the health capital grid points used for solving the model and the fact that the depreciation rate of health is a decrasing function of the health state according to equation 7. For the levels of health capital considered, the depreciation rate has already begun to approach its lower limit. As a result, we do not see much movement in the value when changing the depreciation rate via a shock. In contrast, we would expect to see more noticeable effects for lower values of the health capital. Indeed, figure 4.1, where health capital is held fixed at a low level, shows a larger difference in value between high and low values for $\rho(\epsilon_t)$ compared to figure 4.3, where health capital is fixed at a higher level. Alternatively, figures 4.4-6 show that the difference between obtained value from high and low values of $\rho(\epsilon_t)$ is larger at the lower range of health capital.

5.2 Simulated Shocks

Figure 5 shows how the economy evolves over time. The paths are obtained by running 10,000 simulations of the economy, subject to both productivity and health shocks in each of the 75 periods. Because of the minimal effect of the health shock on higher levels of health, as discussed in section 5.1, the fluctuations in $\rho(\epsilon_t)$ is not explicitly shown.

Figure 5: Simulated Shocks



As expected, total output follows the productivity shock closely. The exception can be found in the first couple of periods, where the negative shock is paired with a high level of output. This can be explained by the high level of capital and high level of labor supply in these periods. Over time, the capital stock and time spent working decline, and total output is to a larger extent determined by fluctuations in productivity.

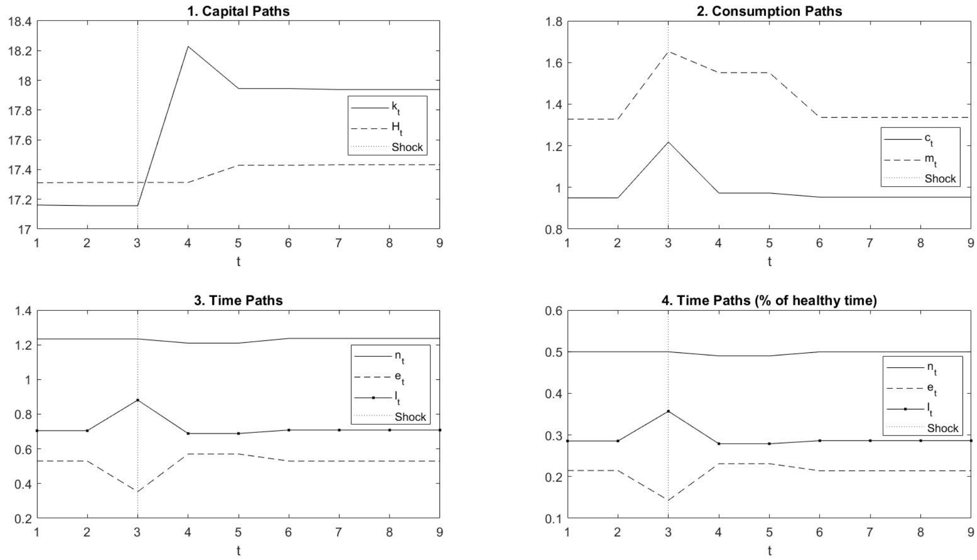
Whether the other variables of the model fluctuate pro- or counter-cyclically is more difficult to discern. Both the capital stock and medical consumption appear to follow the productivity shocks, the former with the expected lag given the law of motion of capital given in equation 4. In order to facilitate the analysis of how the agent changes their behavior in response to the shifting constraints brought about by the shocks, we next turn to impulse response functions.

5.3 Impulse Response Functions

5.3.1 Positive Productivity Shock

The impulse response functions to a positive productivity shock is presented in figure 6. The temporary shock occurs in period 3 and shifts the economy from a low productivity state to a high state, before returning back to low productivity in period 4. The health state is held fixed at no shock in all simulated periods. As shown by the flat curves in periods 1 and 2, all variables follow their steady state paths before the shock. Consequently, all observed movements are caused by the temporary increase in productivity.

Figure 6: Impulse Response Functions to a Positive Productivity Shock



The positive productivity shock in period 3 shifts the budget constraint outwards, according to equation 13. As a result, the agent increases their consumption of both non-medical and medical goods. They also increase their level of investment in capital, as illustrated by the increase in the capital stock in period 4. This lagged effect of the productivity shock on the capital stock is expected given the law of motion specified in equation 4.

In contrast to capital, the health capital stock does not increase in period 4 following the higher level of medical consumption. This is explained by the reduction in exercise time, perfectly matched to offset the increase in medical expenditures. This feature illustrates the point made by Giri and Kumaresan (2021) - economic fluctuations shift the constraints faced by the agent, who subsequently alters their behavior. The overall effect of business cycles on health is the aggregate of these responses. We find that the agent chooses to hold their health state constant and to increase the time spent on leisure, by substituting towards more medical consumption and less exercise in the production of health.

Two interesting points regarding these choices can be emphasized. First, the decision of the agent not to increase their health capital stock indicates that the benefit from additional leisure is greater than the benefit from an outward shift in their time constraint, shown in equation 15. In turn, this indicates that additional health capital, given its already high level, will generate little extra time according to the transformation in equation 9. Second, the agent chooses to reallocate their resources to enjoy a moderate increase in consumption and a moderate increase in leisure. This decision, in contrast to keeping medical consumption and exercise constant to instead enjoy a large increase in consumption, reflects the diminishing marginal utility in the expressions in the value function.

These findings only partially conform to the impulse responses found by Vasilev (2017).

Like us, he finds that exercise time declines in response to a positive productivity shock. In contrast to our results, this is caused by a shift of time from exercise to work in response to the higher marginal wage rates. We hypothesize that the absence of an effect of the productivity shock on the labor supply in our model is again explained by diminishing marginal utility - the agent is better off increasing leisure than enjoying the additional consumption from providing more labor. It could also be that our crude approximations of the policy functions, due to the low number of grid points we are restricted to use, are unable to capture the effect.

Furthermore, Vasilev (2017) finds that the health capital declines following the shock. This discrepancy can be explained by the differences in our models. In contrast to our model, where both medical consumption and exercise can be used to maintain the health stock, Vasilev (2017) includes only the latter. As a result, the agent in his model does not have the option of substituting exercise time for medical consumption in response to the shifting budget constraint. When time spent on exercise is reduced in favor of work, it must by definition happen at the expense of the health capital.

The effects of the productivity shock are also found to be far less persistent in our model compared to Vasilev (2017). Whereas he finds that the paths return to pre-shock levels after around 60 periods, the paths generated by our model reach their new stable levels within two time periods. This is likely to be the result of the few gridpoints used when solving our model and approximating the policy functions, a limitation imposed by the number of dimensions in the maximization problem. With few gridpoints, the change in value when moving from one choice of a control variable to the next may be large. As a result, even though the presented impulse response functions are simulated using interpolated grids, it is possible that we would see a slower return to pre-shock levels if the agent were given more options to choose from at the stage of solving for the policy functions.

The consumption path for medical care presents the exception to the low degree of persistence. Because exercise increases to and even slightly above its pre-shock level in period 4,[2] the slower decline of medical consumption implies that the agent increases their health investment in this period. As a result, we see an increase in the health capital stock in period 5. Noticeably, this is paired with an end to the decline in the capital stock, which began to fall again following the return to the low productivity state. It appears that as a result of the shock, both the capital and the health capital rise to a higher steady state. Once reached, the budget and the time constraints both shift outwards, explaining why consumption and activities return to near but not exactly pre-shock levels.

The limitation of few gridpoints could provide an alternative explanation to the sustained higher levels of capital and health capital after the productivity shock. In line with the reasoning above, it is possible that the agent would choose to return to pre-shock levels if provided with more options. Consequently, we are cautious about interpreting figure 6 as evidence for the existence of two steady states. Instead, we focus our attention on the

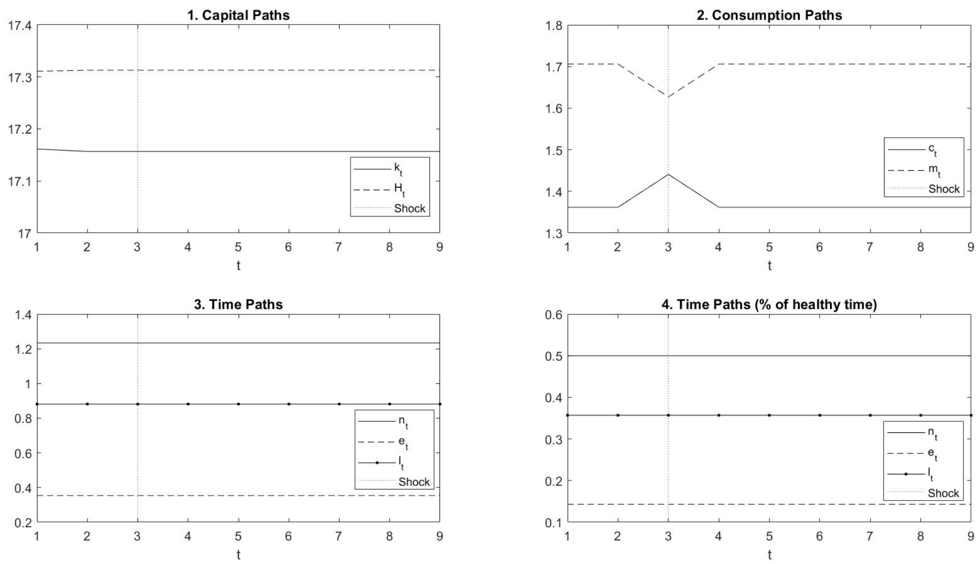
[2]The attentive reader will notice that the increase in exercise time above the pre-shock level is achieved by reducing the time spent working.

dynamics following immediately after the productivity shock.

5.3.2 Positive Health Shock

The impulse response to a positive health shock is presented in figure 7. The temporary shock occurs in period 3 and shifts the agent from a high depreciation rate of health capital to a low one, before returning back to a high rate in period 4. The productivity state is held fixed at no shock in all simulated periods.

Figure 7: Impulse Response Functions to a Positive Health Shock



As shown in figure 7, only the paths of medical and non-medical consumption respond to the health shock[3]. This is in stark contrast to the results by Vasilev (2017) who, despite concluding that the effects of the health shock are smaller and significantly less persistent than those of the productivity shock, finds that exercise increases substantially in response to the shock. This effect spills over to other variables, including capital and health capital. The different effects, including the fact that Vasilev (2017) finds that a positive health shock increases health investment via more exercise whereas we find that health investment falls following lower medical consumption, can again be explained by differences in the models.

Vasilev (2017) defines a positive health shock as increased efficiency of the uncertain health investment. As such, the agent increases their investment when the shock improves the probability of a successful outcome. In contrast, we define the health shock as an

[3] The capital stock and the health capital stock appear to be non-flat between period 1 and 2. This is the result of rounding when solving the model - the optimal choices in period 1 and 2 are in fact found in the same gridpoint. Consequently, this does not reflect any actual movement of the stocks

error term on the depreciation rate of health capital. Consequently, a positive health shock reduces the depreciation rate and thus the amount of health lost in that period. As a result, it is less costly for the agent to maintain their health stock at the pre-shock level. They subsequently reduce their medical expenditures to keep health capital constant, and instead increases their consumption. Because exercise remains unchanged, the allocation of time to other activities is not affected by the shock. In period 4, the agent again faces the higher depreciation rate, and returns to steady state levels of medical and non-medical consumption.

These decisions reflect the fact that the gain from additional consumption is greater than the gain from increasing the health stock and thereby shifting the time constraint outwards. Similarly, but in contrast to the impulse response functions found in section 5.3.1, the gain from decreasing medical expenditures in order to increase consumption is greater than the gain from decreasing exercise to increase leisure. The different behaviors observed in response to the two shocks can again be explained by the diminishing marginal utility of consumption. Whereas consumption has already risen in response to a shift in the budget constraint following the productivity shock, consumption is not directly affected by the health shock. Consequently, the gain from increasing consumption is greater under the health shock compared to the gain from increasing it further under the productivity shock.

6 Discussion

The impulse response functions in section 5.3 indicate that the connection between medical expenditures, exercise, and leisure is relevant in explaining the behavioral choices by an agent in an economy where health is endogenous. As a result, the simplifying assumptions regarding the production of health in previous RBC models with endogenous health accumulation may leave them at risk of overlooking relevant dynamics.

The differences between our model and the model by Vasilev (2017) has already been discussed in detail in section 5.3. In short, by assuming that health is produced using only exercise, Vasilev (2017) denies the agent the opportunity of substituting between exercise and medical expenditures in response to changing circumstances. In our analysis, this possibility lies at the core of the observed behavior. As a consequence of a single input in the health production, the health status is by Vasilev (2017) found to be very responsive to shocks. In contrast, our results suggest that an agent prefers to reallocate their resources in order to avoid health fluctuations.

In contrast to Vasilev (2017), the models by He and Huang (2022), He and Hung (2011), and Feng and Gomis-Porqueras (2011) assume two inputs in the production of health. Whereas this enables the agent to substitute between the two inputs, all three models assume that health is produced using leisure in addition to medical expenditures. As emphasized in section 5.3, the separation of time spent on utility-generating leisure and time spent on health-improving exercise incentivizes the agent to shift time from the latter to the former when an expanding budget constraint enables them to do so and when the marginal utility of leisure exceeds the marginal utility of consumption. As a result, the model successfully accounts for the counter-cyclical nature of physical activity.

Whereas the empirical literature tend to suggest that the counter-cyclical fluctuations of exercise is caused by increased opportunity cost of time in expansions due to rising wage rates, we find that the agent reduces exercise in favor of leisure. In other words, our results suggest an alternative mechanism through which business cycles affect physical activity and subsequently health. This potential mechanism is overlooked in the models by He and Huang (2022), He and Hung (2011), and Feng and Gomis-Porqueras (2011).

This also provides a natural direction for future research seeking to refine the RBC model with endogenous health accumulation further. Possible extensions to the model presented here could for example draw inspiration from the empirical papers by Xu (2013), who divides activities into those that are time intensive and those that require little time, and Colman and Dave (2018), who divide behaviors into either utility-generating or utility-reducing and into either health-improving or health-destroying.

7 Conclusion

The purpose of this paper is to contribute to the limited literature of RBC models with endogenous health accumulation. To this end, we develop an infinitely lived agents model with endogenous health, productivity shocks, and health shocks. The conceptualization of health is based on the model by Grossman (1972a) but adjusts for some existing criticism. Specifically, we assume that the depreciation rate of health capital is a decreasing function of the health state, thereby ensuring that health deteriorates slower with improved health and that it is cheaper for a healthier individual to maintain their health state.

After solving the model, we find that it is able to predict the key relationships discussed in the empirical literature. In particular, the model predicts pro-cyclical medical expenditures, counter-cyclical physical activity, and a negative correlation between health status and medical expenditures. By simulating impulse response functions, we find that a positive productivity shock has a larger effect than a positive health shock, in line with the findings by Vasilev (2017). Our results also show increased consumption of both medical and non-medical goods in response to the productivity shock, in accordance with the income channel proposed by Ruhm (1994).

In contrast to the results by Vasilev (2017), we do not find that the counter-cyclical trend in exercise is explained by the increased opportunity cost of time as marginal wage rates rise in expansions. Instead, our results suggest that the agent decreases exercise in favor of leisure when the health stock can be maintained by substituting towards more medical consumption. The ability of the agent to substitute medical care for exercise when faced with an expanding budget constraint further leads to no effect of a productivity shock on the health status, suggesting that the agent prefers to reallocate their resources in order to avoid health fluctuations.

Based on our results, we conclude that the connection between medical expenditures, exercise, and leisure is relevant in explaining the behavioral choices by an agent in an economy where health is endogenous. Consequently, the assumption by Vasilev (2017) to use only exercise as an input in the production of new health risks overstating the fluctuations in health as the agent is unable to substitute towards medical consumption in response to changing circumstances. Furthermore, our results suggest that the agent prefers to decrease exercise and increase medical consumption in favor of leisure when the marginal benefit of leisure exceeds the marginal benefit of consumption. This dynamic is missed in the formulations by He and Huang (2022), He and Hung (2011), and Feng and Gomis-Porqueras (2011), who do not distinguish between leisure and exercise. Overall, we propose that an RBC model with endogenous health accumulation should account for a more complex health production than has previously been used.

References

Aísa, Rosa and Fernando Pueyo (2004). "Endogenous longevity, health and economic growth: a slow growth for longer life?" In: *Economics Bulletin* 9.3, pp. 1–10.

Bellés-Obrero, Cristina and Judit Vall Castelló (2018). "The Busienss Cycle and Health". In: *Oxford Research Encyclopedia, Economics and Finance*. Oxford University Press USA.

Case, Anne and Deaton Angus (2005). "Broken Down by Work and Sex: How Our Health Declines". In: *Analyses in the Economics of Aging*. University of Chicago Press. Chap. 6, pp. 185–212.

Colman, Gregory and Dhaval Dave (2013). "Exercise, physical activity, and exertion over the business cycle". In: *Social Science Medicine* 93, pp. 11–20.

– (2018). "Unemployment and Health Behaviors over the Business Cycle: A Longitudinal View". In: *Southern Economic Journal* 85.1, pp. 93–120.

Dalgaard, Carl-Johan and Holger Strulik (2014). "Optimal Aging and Death". In: *Journal of the European Economic Association* 12.3, pp. 672–701.

Dos Santos, João Pereira, Mariana Tavares and Pedro Pita Barros (2016). "More than just numbers: Suicide rates and the economic cycle in Portugal (1910-2013)". In: *SSM - population health* 2, pp. 14–23.

Ehrlich, Isaac and Hiroyuki Chuma (1990). "A model of the demand for longevity and the value of life extension". In: *Journal of Political Economy* 98.4, pp. 761–782.

Feng, Zhigang and Pedro Gomis-Porqueras (2011). "Consequences of valuing health: A macroeconomic perspective". In: *8th World Congress on Health Economics,Toronto, 10 July 2011- 13 July 2011*, pp. 1–16.

Galama, Titus and Arie Kapteyn (2011). "Grossman's Missing Health Threshold". In: *Journal of Health Economics* 30, pp. 1044–1056.

Galama, Titus J (2015). "A contribution to health-capital theory". In: *CESR-Schaeffer Working Paper No. 2015-004*.

Galama, Titus J et al. (2012). "Is there empirical evidence for decreasing returns to scale in a health capital model?" In: *Health Economics* 21, pp. 1080–1100.

Gavrilov, Leonid A and Natalia S Gavrilova (1991). *The Biology of Human Life Span: A Quantitative Approach*. Harwood Academic.

Gerdtham, Ulf-G and Magnus Johannesson (2005). "Business cycles and mortality: results from Swedish microdata". In: *Social science & medicine* 60.1, pp. 205–218.

Giri, Jeeten K. and Talitha Kumaresan (2021). "The Business Cycle, Health Behavior, and Chronic Disease: A Study Over Three Decades". In: *Economics and Human Biology* 43.

Grossman, Michael (1972a). "On the Concept of Health Capital and the Demand for Health". In: *The Journal of Political Economy* 80.2, pp. 223–255.

– (1972b). *The Demand for Health – A theoretical and Empirical Investigation*. National Bureau of Economic Research.

– (1998). "On Optimal Length of Life". In: *Journal of Health Economics* 17, pp. 499–509.

He, Hui and Kevin X.D Huang (2022). "'Healthy' business cycle". In: *SSRN*. URL: https://ssrn.com/abstract=3860641 (visited on 07/11/2022).

He, Hui and Sheng-Ti Hung (2011). "Are Recessions Good for Your Health? A Macroeconomic Analysis". In: *Society for Economic Dynamics* 2011 Meeting Papers 1178, pp. 1–21.

Laporte, Audrey (2015). "Should the Grossman Model of Investment in Health Capital Retain its Iconic Status". In: *CCHE/CCES Working Paper No: 2014:04*.

Laun, Tobias (2012). "Essays on Macroeconomics and Health". PhD thesis. Stockholm School of Economics.

Mitnitski, Arnold et al. (2002). "The Accumulation of Deficits with Age and Possible Invariants of Aging". In: *Scientific World* 2, pp. 1816–1822.

Muurinen, Jana Marja and Julian Le Grand (1998). "The Economic Analysis of Inequalities in Health". In: *Social Science and Medicine* 20.10, pp. 1029–1035.

Ried, Walter (1998). "Comparative Dynamic Analysis of the Full Grossman Model". In: *Journal of Health Economics* 17, pp. 383–425.

Ruhm, Christopher J (1994). "Economic Conditions and Alcohol Problems". In: *NBER Working Paper Series* 4914.

– (2000). "Are recessions good for your health?" In: *Quarterly Journal of Economics* 115.2, pp. 617–650.

– (2005). "Healthy Living in Hard Times". In: *Journal of Health Economics* 24, pp. 341–363.

Scholz, John Karl and Ananth Seshadri (2010). "Health and Wealth in a Life Cycle Model". In: *Michigan Retirement Research Center Research Paper No. 2010-224*.

Stokey, Nancy L., Robert E. Lucas Jr and Edward C. Prescott (1989). *Recursive methods in economic dynamics*. Harvard University Press.

Strulik, Holger (2015). "A closed-form solution for the health capital model". In: *Journal of Demographic Economics* 81, pp. 301–316.

Tauchen, George (1986). "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions". In: *Economic Letters* 20, pp. 177–181.

Vasilev, Aleksandar (2017). "US Health and Aggregate Fluctuations". In: *Bulletin of Economic Research* 69.2, pp. 147–163.

Wagstaff, Adam (1986). "The Demand for Health: Some New Empirical Evidence". In: *Journal of Health Economics* 5, pp. 195–233.

– (1993). "The Demand for Health: An Empirical Reformulation of the Grossman Model". In: *Health Economics* 2.2, pp. 189–198.

Xu, Xin (2013). "The Business Cycle and Health Behaviors". In: *Social Science Medicine* 77, pp. 126–136.

Zon, Adriaan van and Joan Muysken (2001). "Health and endogenous growth". In: *Journal of Health Economics* 20, pp. 169–185.

Zweifel, Peter, Friedrich Breyer and Mathias Kifmann (2009). "Empirical Studies of the Production of Health". In: *Health Economics*. Springer, Berlin, Heidelberg. Chap. 4, pp. 119–153.