Reference Health and Investment Decisions [1]

Chunli Cheng Christian Hilpert Alexander Szimayer Peter Zweifel

Abstract

Reference points influence economic decisions. This paper considers how health reference points and their adaptation to decreasing health influence medical spending, consumption, and investment in a dynamic model. A static reference point implies an aspiration to offset health losses already at a high initial level. In contrast, the case of reference adaptation entails much lower lifetime healthcare expenditure that concentrates late in life. A projection bias, i.e., the agent's failure to anticipate the reference adaptation, induces behavior that initially resembles the static reference case. With decaying health, choices approach, but remain distinct, to those derived for adaptive reference health.

JEL classification: G11, G40, I10

Keywords: Investment, medical spending, optimal control, behavioral economics, reference adaptation.

---

[1]Cheng: Lingnan College, Sun Yat-sen University, Guangzhou 510275, PR China, Email: chengchli@mail.sysu.edu.cn. Hilpert: Lingnan College, Sun Yat-sen University, Xingang Xi Road 135, 510275 Guangzhou, China. Email: martin@mail.sysu.edu.cn. Szimayer: Universität Hamburg, Department of Economics and Social Science, Von-Melle-Park 5, 20146 Hamburg, Germany. Email: alexander.szimayer@wiso.uni-hamburg.de. Zweifel: Emeritus, University of Zürich, Department of Economics, Email: peter.zweifel@uzh.ch.
We thank Marcel Freyschmidt (discussant), Evgeny Lyandres, Junqing Kang, Tarik Umar (discussant) and seminar participants at Sun Yat-sen University for helpful comments and discussions. Furthermore, we thank participants of the 50th Seminar of the European Group of Risk and Insurance Economists, the 2023 Econometric Society Australasian Meeting, the 27th Asia-Pacific Risk and Insurance Association, the Asian Meeting of the Econometric Society in East and Southeast Asia, the 2022 SusTech Workshop Financial Engineering, and the New Zealand Finance Meeting 2022 for helpful comments.

1 Introduction

Benchmark values, also called reference points, influence the economic decisions of individuals. They form a fundamental building block of descriptive theories of choice under risk, such as prospect theory and cumulative prospect theory by Kahneman and Tversky (1979) and Tversky and Kahneman (1992) and the models of reference dependence by Köszegi and Rabin (2006, 2007). Typically, such reference points are not static but evolve over time (Baucells, Weber, and Welfens, 2011; Thakral and To, 2021).

The objective of this paper is to show how reference in health and its adaptation to decreasing health influence medical spending as well as consumption and investment. Hugonnier, Pelgrin, and St-Amour (2013) model and empirically confirm the interdependence of health and asset holdings. Yogo (2016) documents that stochastically decaying health and out-of-pocket health expenditure significantly affect household investment decisions. Gottlieb (2012) applies prospect theory with its reference point to puzzling choices of life insurance and annuities.

The first contribution of this paper is a dynamic model of medical spending, resulting health development, consumption and investment choices incorporating static and adaptive health reference levels. An agent faces a series of health shocks which can be offset in part by medical care. Yet, the associated out-of-pocket expenditure limits the wealth available for future expenditure on health and consumption. Additionally, investment in a risky and risk-free asset needs to be decided upon anew. Thus, this paper builds on a consumption and portfolio choice problem (Merton, 1969) and a model of health care expenditure (Grossman, 1972). However, contrasting Hugonnier, Pelgrin, and St-Amour (2013), it takes a reference-dependence perspective instead of a recursive preference one.

The second contribution is the updating of health reference points. As health decays with every shock, the agent either still judges her health against the previous value or adjusts the reference point downward to a new level. In the case of perfect adaptation, this does not entail a loss of utility. In the case of less-than-perfect adaptation, the agent suffers a loss in the guise of a permanently lower quality of life. Reference points evolve over time, (Baucells, Weber, and Welfens, 2011), but retain a dependency on the original reference level (De Giorgi and Post, 2011). This adaptation influences consumption choices if the reference point adapts to the consumption history (van Bilsen, Laeven, and Nijman, 2020), and mitigates the strength of loss aversion in the context of portfolio choice (He and Strub, 2022). This paper, in contrast, introduces a reference level for health.

However, despite reference points being a key driver of individual decision making, a projection bias, that is, the agent's failure to anticipate the adaptation of the reference point after the next health shock to a new, lower level, affects utility. Loewenstein, O'Donoghue, and Rabin (2003) capture projection bias in a model of consumer choice and explore its impact on consumption decisions. Acland and Levy (2015) use a combination of projection bias and naiveté concerning habit formation to explain why people fail to commit to attending their gym. Kliger and Levy (2008) infer a sizable projection bias of US investors in derivatives by transforming stochastic discount factors. Strub and Li (2020) analyze the time inconsistency projection bias introduces to dynamic portfolio choice and study the impact of various commitment strategies on optimal risk taking.

The third contribution is the specification of the utility function. At every point in time, consumption and the investment portfolio composition after adjustment determine the agent's instantaneous utility. On the one hand, good health enhances utility derived from consumption; on the other hand, a high reference health level lowers it because the current health status is perceived as a loss. Finally, instantaneous utility needs to be balanced against the after-shock utility stream which includes a bequest utility and depends on the adaptation of the health reference level which in turn affects current choices.

The utility function is designed to capture the empirical properties of the connection between health reference points and resulting health and consumption found by Harris and Kohn (2018) in a continuous-time optimal control framework. Most importantly, the utility function is specified in such a generic form that it can accommodate the special cases of terminal illness and absence of health shocks in semi-closed form, and characterize the optimal controls and expected lifetime utility for the general case as an ordinary differential equation.

The numerical analysis shows that a static reference point implies a strong desire to offset health losses already at a high initial level, calling for much health care in early life. The introduction of a static health reference value causes substantial deviations from the classical Merton (1969) solution. As health deteriorates, both consumption utility and expected lifetime utility decrease; the utility function exhibits an increasing degree of risk aversion; the optimal consumption sharply increases in response to a hike in marginal utility of consumption; the optimal investment becomes less risky in a way that depends on wealth in a non-linear way. However, in the case of terminal illness, the aversion reverts to some degree, causing the share invested in the risky asset to soar.

Compared to the case of a static reference health value, the case of adaptation entails much less change in optimal behavior. The realization that previous health levels become unattainable, and the focus on achieving more realistic health levels enhance the utility of consumption and limits medical spending. Also, lifetime health care expenditure is less and occurs late in life. Risk aversion does not increase as before, resulting in a higher share of risky investment compared to the case of static reference health. In all, optimal consumption and investment choices approximate the Merton (1969) solution.

Finally, from behavioral economics it is known that agents suffer from biases in their decision making. In the present context, projection bias may cause them to fail to correctly anticipate that the health shock impacts utility, resulting in optimal choices that are intermediate between the cases of static and adaptive reference health values in case the agent anticipates the reference adaptation. Given loss aversion, a health shock causes substantial disutility (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Hence, when the agent fails to anticipate adaptation of the health reference value, the health loss after the next shock appears sizable, inducing behavior similar to the static reference case. Over time, however, the gap between actual health and its reference value decreases, and consumption and investment choices approach the ones induced by fully adaptive reference health. Still, projection bias alters the dynamic choice problem compared to the case without projection bias.

The remainder of this paper has the following structure. The elements of the behavioral model are expounded in Section 2, followed by a characterization of the optimal control problem in Section 3. The special cases of terminal illness and no health shocks are analyzed in Section 4, while Section 5 is devoted to the numerical solution of the optimal control problem in the case of a static health reference value, a fully adaptive one with anticipation, and of projection bias. Section 6 contains conclusions and suggestions for overcoming the limitations of the present work.

2 Stochastic health shocks, medical spending, consumption and investment, and impacts on utility

2.1 Health shocks and medical spending

Health H is at an initial level $H_0 > 0$ at starting time $t = 0$. In random intervals, it suffers a random shock θ occurring with probability $\lambda(H)$ in the interval $[t, t+dt]$ that decreases in H. The random shocks have a support $\Theta = [\underline{\theta}, \overline{\theta}]$ with $0 < \underline{\theta} < \overline{\theta}$, which are independent and identically distributed with a distribution F_θ. A point process N_t documents the health shocks with T_n indicating the timing of the shocks and θ_n denoting their magnitude. Death occurs if health falls below a critical value H_D, with $0 < H_D < H_0$, which occurs at the stopping time T_D given by

$$T_D = \inf\{t \geq 0 : H_t \leq H_D\} \; = \; \inf\{T_n : H_{T_n} \leq H_D\}\,. \tag{1}$$

Health shocks can be offset to a degree $V(k,\theta,h) \geq 0$ with $V(0,\theta,h) = 0$ by medical treatment costing $k \geq 0$. The resulting change in health therefore equals $V(k,\theta,h) - \theta$. Medical spending has the following effects: $\lim_{k\to 0} V_k(k,\theta,h) > 0$, $V_{k,k}(k,\theta,h) < 0$, $\lim_{k\to\infty} V_k(k,\theta,h) = 0$, and $V_{k,\theta}(k,\theta,h) > 0$. Thus, marginal effectiveness is strictly positive at zero medical spending, but is decreasing to become zero for infinite medical spending, while increasing with the severity of the shock. Finally, the medical treatment effectiveness depends on health through some lasting loss of health $\varepsilon(h) > 0$ left by health shocks such that

$$\theta - V(k,\theta,h) \geq \varepsilon(h)\,. \tag{2}$$

$\varepsilon(h)$ is decreasing in health: The lower the level of health, the more lasting damage the agent incurs. To be specific, the following functional form having the properties cited above is assumed,

$$V(k,\theta,h) = (\theta - \varepsilon(h))\left(1 - e^{-\frac{k}{\theta}}\right). \tag{3}$$

with a simple functional form $\varepsilon(h) = \frac{\varepsilon}{1+h}$, $\varepsilon > 0$.[1] To ensure a positive relationship between medical spending and health improvement at any health level h and health shock θ, ε is set to satisfy $\frac{\varepsilon}{1+H_D} < \underline{\theta}$.

The net health shock equals

$$\theta - V(k,\theta,h) = \theta\, e^{-\frac{k}{\theta}} + \frac{\varepsilon}{1+h}\left(1 - e^{-\frac{k}{\theta}}\right). \tag{4}$$

The health process therefore evolves as follows,

$$H_t = H_0 - \sum_{n=1}^{N_t}\left(\theta_n - V(k_n,\theta_n,H_{T_n-})\right). \tag{5}$$

The shocks cause initial health H_0 to deteriorate until it hits the death value T_D.

[1] Alternative complex functional forms of $\varepsilon(h)$ can be considered, but do not materially change our results.

2.2 Consumption and financial investment

Besides managing health and medical expenditure, the agent spends money on consumption at a rate c_t and decides on how to invest his or her savings. Between health shocks, wealth is invested in a portfolio X that contains a risk-free and a risky asset as in Merton (1969). With the risky asset following a geometric Brownian motion, the portfolio dynamics is given by

$$dX_t = (a_t(\mu - r) + r)X_t dt - c_t dt + \beta H_t X_t dt + a_t \sigma X_t dW, \qquad X_0 = x_0 > 0. \tag{6}$$

Here, a_t denotes the investment strategy which splits the assets between the one with the risk-free return r and the one with the risky expected return $\mu > r$. The agent receives a continuous health-dependent income stream of $\beta H_t X_t$, where $\beta \geq 0$ represents that high health allows for higher earnings as in Hugonnier, Pelgrin, and St-Amour (2013). Risk is represented by the Brownian motion W scaled by the volatility $\sigma > 0$. Accounting for medical spending in response to the health shocks, he or she obtains for net wealth, with x_0 symbolizing its initial value,

$$X_t = x_0 + \int_0^t (X_s\,[a_s\,(\mu - r) + r + \beta H_s] - c_s)\,\mathrm{d}s + \int_0^t X_s\, a_s\, \sigma\, \mathrm{d}W_s - \sum_{n=1}^{N_t} k_n\,. \tag{7}$$

Since the issue of debt is excluded for simplicity, the constraint $X_t > 0$ is imposed.

2.3 Reference health and a utility function

The agent draws utility from consumption and health over his or her lifespan.[2] Health H_t is valued both in absolute terms and relative to a reference level B_t, according to the utility function $u_t(c_t, H_t, B_t)$ at time t. Its specfication shall absorb the empirical findings on reference health impacts on utility by Harris and Kohn (2018). Staying in line with Harris and Kohn (2018), the health reference value reflects past health realizations, giving rise to the health reference

[2]The lifespan depends on medical spending; conceivably, it could also depend on spending on a healthy lifestyle that decreases the frequency of health shocks.

updating rule

$$B_{t\in[T_n,T_{n+1})} = (1-\omega)B_{T_{n-1}} + \omega H_{T_n}, \text{ with } T_0 = 0 \text{ and } B_{t\in[0,T_1)} = H_0. \tag{8}$$

Here, the reference starts at the initial health level, $B_0 = H_0$, and updates at the arrival of health shocks. $\omega \in [0,1]$ measures the speed of the updating of the health reference value, with $1-\omega$ denoting the weight of the previous reference value and ω that of the after-shock health level. At one extreme, $\omega = 0$, the lifetime reference point remains at $B_t = H_0$, which is referred to as the static health reference. With $\omega = 1$, the reference value immediately adjusts to the after-shock health level such that $B_t = H_t$. Therefore, the agent's reference health level at time t is given by

$$B_t = (1-\omega)^{N_t} H_0 + \sum_{n=1}^{N_t} (1-\omega)^{N_t-n} \omega H_{T_n}. \tag{9}$$

Next, taking the signs of the marginal cross-utilities of consumption, health, and reference health empirically identified by Harris and Kohn (2018), which are $u_{CH} < 0$[3], $u_{CB} > 0$ and $u_{HB} > 0$, we specify the utility function

$$u(c,h,b) = \frac{(ce^{h-b})^{1-\gamma}}{1-\gamma},\ \gamma > 1, \tag{10}$$

satisfying usual consumption and health derivative conditions $u_C > 0$, $u_H > 0$, $u_{CC} < 0$, and $u_{HH} < 0$. With $u_{CH} < 0$, consumption and health are substitutes: When the agent is at a worse health stage, he or she feels more compensated by additional consumption, and at the same time, his or her additional health improvement matters more when he or she already has less to consume, complying with correlation aversion in consumption and health (Eeckhoudt, Rey, and

[3]In Harris and Kohn (2018), $u_{CH} < 0$ was identified for high consumption, while for low consumption, they found $u_{CH} > 0$. To be comparable to Hugonnier, Pelgrin, and St-Amour (2013) which assumes health and consumption are substitutes, we adopted $u_{CH} < 0$ to define the utility function in this paper.

Schlesinger, 2007). The health reference, which reveals the agent's health history, also affects the marginal benefits from additional consumption and health: With $u_{CB} > 0$ and $u_{HB} > 0$, a higher health trajectory lets the agent enjoy additional consumption and health more. Nevertheless, the health reference is a benchmark based on which the agent evaluates his or her current health. A higher reference level has a negative effect on the agent's utility, that is $u_B < 0$. γ symbolizes the risk-aversion parameter. With $\gamma > 1$, the utility function (10) implies constant relative risk aversion of $\gamma > 1$ w.r.t consumption and constant absolute risk aversion of $\gamma - 1 > 0$ w.r.t. health.

2.4 Optimal control problem and projection bias

The agent manages his or her health by spending wealth on medication in face of health shocks and forms consumption and financial investment decisions to maximize his or her utlity $u(c,h,b)$ over his or her lifespan. With (H,B,X) being a controlled process and (a,c,k) the control, lifetime utility equals

$$J(h,b,x;a,c,k) = \mathbb{E}_{(h,b,x)}\left[\int_0^{T_D} e^{-\rho t} u_t(c_t,H_t,B_t)\,\mathrm{d}t\right] + \mathbb{E}_{(h,b,x)}\left[e^{-\rho T_D} U_D(X_{T_D})\right], \tag{11}$$

with $\rho > 0$ as the rate of time preference and conditional on the current status of the state space (h,b,x). The first term describes the expected lifetime utility in present value terms, while the second term denotes the expected utility of wealth at death X_{T_D} left as the bequest also in present value terms. All utility terms vary with both current health H_t and health reference value B_t.

The death utility $U_D(x)$ captures the agent's bequest motive given by his or her expected discounted utility with an infinite time horizon after the death time T_D assumed for the heirs,

$$U_D(x) = \mathbb{E}_{(h,b,x)}\left[\int_{T_D}^{\infty} e^{-\rho\,(t-T_D)} u(c_D(x),h,b)\,\mathrm{d}t\right]. \tag{12}$$

Note that consumption $c_D(x)$ depends exclusively on bequest wealth x which at the time of death is known. After T_D a risk-free value-preserving strategy is implemented by assumption such that

$$0 \stackrel{!}{=} \mathrm{d}X_t = (rX_t - c_D(X_t))\,\mathrm{d}t \quad \Longleftrightarrow \quad c_D(x) = rx\,.$$

Thus, death utility is the utility of a certain consumption stream financed by the interest on the risk-free asset assumed for the heirs. By rewriting the utility function $u(c,h,b)$ as

$$u(c,h,b) = \frac{c^{1-\gamma}}{1-\gamma} K(h,b,\gamma), \tag{13}$$

where $K(h,b,\gamma) = e^{(h-b)(1-\gamma)}$, and substituting into the death utility equation (12), one obtains

$$U_D(x) = \mathbb{E}_{(h,b,x)}\left[\int_{T_D}^{\infty} e^{-\rho\,(t-T_D)}\frac{(rx)^{1-\gamma}}{1-\gamma} e^{(h-b)(1-\gamma)}\,\mathrm{d}t\right] = -e^{(h-b)(1-\gamma)}\frac{x^{1-\gamma}}{\frac{\rho}{r}\,r^{\gamma}(\gamma-1)}\,.$$

Hence, the lifetime utility in equation (11) becomes

$$J(h,b,x;a,c,k) = \mathbb{E}_{(h,b,x)}\left[\int_0^{T_D} e^{-\rho\,t}u_t(c_t,H_t,B_t)\,\mathrm{d}t\right] - \frac{e^{(h-b)(1-\gamma)}}{\frac{\rho}{r}\,r^{\gamma}(\gamma-1)}\mathbb{E}_{(h,b,x)}\left[e^{-\rho\,T_D}X_{T_D}^{1-\gamma}\right] \tag{14}$$

The stochastic health decay, the health reference level, and the assets all are Markovian. Additionally, the health shock times follow a point process; consequently, the health decay is memoryless. The expected time until the next health shock only depends on the intensity $\lambda(H)$. The actual time elapsed since the last health shock carries no information. This structure implies that the agent's decisions are also Markovian, reflecting exclusively current health, health reference level, and financial wealth, i.e., the current status of the state space (H,B,X). This further implies that age does not play a role, leaving H_t to reflect the proximity to death, in line with Zweifel, Felder, and Meier (1999) and Zweifel, Felder, and Werblow (2004), who found healthcare expenditure to increase with proximity to death rather than age.

Solving directly the optimal control problem (2.4) with $\omega > 0$ in the health reference updating equation (8) assumes that the agent anticipates his or her health reference updating in response to future health shocks, that is, adapting to future health shocks is assumed to be foreseeable. However, presented by Loewenstein, O'Donoghue, and Rabin (2003) (see Section 2), this is usually not the case. People fail to forecast their future adaptation, a behavior which the authors coin ”projection bias”. As a consequence, people make decisions based on how they have adapted to health shocks experienced until now rather than based on their foreseen adaptations to future health shocks. In other words, people make decisions based on their reference health level available at that time with no anticipation of any change of it for the future. As time passes with future health shocks materializing, people update their health reference level according to equation (8), however, they were not aware of this updating before hand. To include projection bias in the agent's decision making problem, we update his or her lifetime utility by including a projection bias indicator $\hat{\omega}$:

$$J(h,b,x;a,c,k) = \mathbb{E}_{(h,b,x)}\left[\int_0^{T_D} e^{-\rho t} u_t(c_t, H_t, (1-\hat{\omega})b + \hat{\omega} B_t)\,\mathrm{d}t\right] - \frac{e^{(h-b)(1-\gamma)}}{\frac{\rho}{r} r^{\gamma}(\gamma-1)} \mathbb{E}_{(h,b,x)}\left[e^{-\rho T_D} X_{T_D}^{1-\gamma}\right], \tag{15}$$

with

$$\hat{\omega} = \begin{cases} 1, & \text{without projection bias,} \\ 0, & \text{with projection bias.} \end{cases}$$

With $\hat{\omega} = 1$, the lifetime utility in equation traces back to that in equation , by which the agent forms his or her decisions based on the foreseen updating health reference B_t, that is the agent does not incur projection bias. Otherwise, the agent makes decisions based on his or her current health reference b, implying that the agent has projection bias. Note for people with $\omega = 0$, the reference value is fixed at the initial health level, $B_t = H_0$, for which there is no health reference updating, and automatically, there is no projection bias. In this case, the indicator $\hat{\omega}$ loses its value in distinguishing projection bias and both values of $\hat{\omega}$ lead to the same decisions.

3 Optimal medical spending, consumption, and investment

The model depicts two choices, how much to spend on health care in response to health shocks, and how to adjust consumption and the structure of investment. The two choices are interconnected: Medical spending restores health at the expense of future consumption, yet a higher health level expands the lifetime to enjoy consumption. Consumption brings immediate utility now but limits both future medical spending and consumption utility due to a reduced lifespan. The investment choice trades off a higher return on the risky asset which can be spent on healthcare and consumption against the risk of low returns in some cases.

Consider the medical spending choice first. At T_n the agent knows the health level h and wealth x just before the next health shock with observable magnitude θ_n arrives. The amount k_n spent on health care results in a remaining lifetime utility of

$$J(h-\theta_n+V(k_n,\theta_n,h),(1-\hat{\omega})b+\hat{\omega}B_{T_n},x-k_n;a,c,k)\,. \tag{16}$$

Health is restored by a degree $V(k_n,\theta_n,h)$ but reduces wealth by k_n which lowers future spending ability by the same amount. B_{T_n} is the new health reference after the shock, which represents the weighted average of the health reference level before the shock and the updated new after-shock health value such that $B_{T_n}=(1-\omega)b+\omega(h-\theta_n+V(k_n,\theta_n,h))$, $\omega\in[0,1]$. Depending on whether the agent has projection bias in predicting his or her health reference change, he or she makes decisions based either on the new health reference B_{T_n}, that is without projection bias ($\hat{\omega}=1$), or on the health reference before the shock b, that is with projection bias ($\hat{\omega}=0$).

Turning to the optimization, denote by $k^{\star}(h,b,x,\theta_n,\omega,\hat{\omega};a,c,k)$ the argument maximizing (16) at (h,b,x) given the shock size θ_n, the value of ω and $\hat{\omega}$, and subsequent strategy

(a,c,k). Suppose that an optimal strategy $(a^\star, c^\star, k^\star)$ exists with value function

$$U(h,b,x) = \sup_{(a,c,k)} J(h-\theta+V(k,\theta,h),(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega(h-\theta+V(k,\theta,h))),x-k;a,c,k)\,. \tag{17}$$

Then, $k^\star$ depends on (h,b,x), θ, ω, and $\hat{\omega}$ and satisfies

$$k^\star(h,b,x,\theta,\omega,\hat{\omega}) = \\ \arg\max_{0\le k\le x} U(h-\theta+V(k,\theta,h),(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega(h-\theta+V(k,\theta,h))),x-k),$$

where $B_{T_n} = (1-\omega)b+\omega(h-\theta_n+V(k_n,\theta_n,h))$, $\omega \in [0,1]$, $\hat{\omega} = 1$ is without projection bias, and $\hat{\omega} = 0$ is with projection bias. Since $X \ge 0$, we have that $k \le x$, and thus $k^\star(h,b,x,\theta,\omega,\hat{\omega})$ exists, provided U and V are sufficiently smooth, however, it may not be unique.[4]

Proposition 1. *Under stated assumptions, the (possibly not unique) optimal medical expenditure choice exists and is characterized by*

$$k^\star(h,b,x,\theta,\omega,\hat{\omega}) = \\ \arg\max_{0\le k\le x} U(h-\theta+V(k,\theta,h),(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega(h-\theta+V(k,\theta,h))),x-k), \tag{18}$$

where

$$\hat{\omega} = \begin{cases} 1, & \textit{without projection bias}, \\ 0, & \textit{with projection bias}, \end{cases}$$

and $\omega \in [0,1]$.

[4]In case $k^\star(h,b,x,\theta,A)$ is not unique, a criterion has to be postulated in order to select the maximizing argument. Arguably, this is maximum medical spending because it yields the highest post-shock health level and hence lifespan.

Next, the value function U needs to be specified. At any time t where no health shock takes place and for a starting value (h,b,x), it can be written as

$$U(h,b,x)=\mathbb{E}_{(h,b,x)}\left[\int_0^{T_1} e^{-\rho t}u(c_t,h,b)\,\mathrm{d}t\right]+\mathbb{E}_{(h,b,x)}\Big[e^{-\rho T_1}U(h-\theta_1+V(k^\star(h,b,X_{T_1},\theta_1,\omega,\hat{\omega}),\theta_1,h),$$
$$(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega(h-\theta_1+V(k^\star(h,b,X_{T_1},\theta_1,\omega,\hat{\omega}),\theta_1,h))),X_{T_1}-k^\star(h,b,X_{T_1},\theta_1,\omega,\hat{\omega}))\Big]. \tag{19}$$

The first term is the agent's expected stream of utility derived from current consumption, health level, and health reference value, discounted by the time preference parameter up to the next health shock. The second term is the expected value of future utility valued at the next health shock given the optimal medical spending choice. For any potential magnitude of the health shock, the agent selects the optimal medical spending that trades off future consumption against recovering from this particular shock.

Equation (19) can be simplified, since T_1 is random with intensity $\lambda(h)$ and the distribution of the next health shock θ_1 is given by F_θ, both of which are independent of the Wiener process W driving the wealth process, one can integrate T_1 out to obtain

$$U(h,b,x)=\mathbb{E}_{(h,b,x)}\left[\int_0^{\infty} e^{-[\rho+\lambda(h)]t}u(c_t,h,b)\,\mathrm{d}t\right]+\mathbb{E}_{(h,b,x)}\Big[\int_0^{\infty} e^{-[\rho+\lambda(h)]t}\lambda(h)\int U(h-\theta+$$
$$V(k^\star(h,b,X_t,\theta,\omega,\hat{\omega}),\theta,h),(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega(h-\theta+V(k^\star(h,b,X_t,\theta,\omega,\hat{\omega}),\theta,h))),$$
$$X_t-k^\star(h,b,X_t,\theta,\omega,\hat{\omega}))\,\mathrm{d}F_\theta(\theta)\,\mathrm{d}t\Big]. \tag{20}$$

Also, given the optimal medical spending choice, expected future utility at the next health shock is given by integrating over the health shock distribution F_θ,

$$\mathbb{U}(h,b,x) = \int U(h-\theta+V(k^\star(h,b,X_t,\theta,\omega,\hat{\omega}),\theta,h),$$
$$(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega(h-\theta+V(k^\star(h,b,X_t,\theta,\omega,\hat{\omega}),\theta,h))),X_t-k^\star(h,b,X_t,\theta,\omega,\hat{\omega}))\,\mathrm{d}F_\theta(\theta)\,. \tag{21}$$

The quantity $\mathbb{U}(h,b,x)$ is calculated based on $(U(h',b',\cdot))_{\{h-\varepsilon(h)\geq h'\geq H_D,\, b'=(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega h')\}}$. Since h and b act as parameters, one can characterize $U(h,b,x)$ for fixed values h and b and variable x using optimal control arguments, that is, the Hamilton-Jacobi-Bellman formalism.

Fix h and b and let $x\geq 0$ be variable, then assume that $U(h',b',x)$ is given for $h'\leq h-\varepsilon(h)$ and $b'=(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega h')$, where $\varepsilon(h)$ is the minimum effect of a health shock after the best possible treatment, see Equation (2). From there and using (18), the optimal healthcare expense $k^\star(h,b,x,\theta,\omega,\hat{\omega})$ can be derived for $h'\leq h-\varepsilon(h)$ and $b'=(1-\hat{\omega})b+\hat{\omega}((1-\omega)b+\omega h')$, and all $x\geq 0$ as well as θ in the support of the health shock distribution F_θ. Hence $\mathbb{U}(h,b,x)$ can be computed for all $x\geq 0$ based on (21).

The value function can now be written as

$$U(h,b,x) = \mathbb{E}_{(h,b,x)}\left[\int_0^\infty e^{-[\rho+\lambda(h)]t}\,(u(c_t,h,b)+\lambda(h)\,\mathbb{U}(h,b,X_t))\,\mathrm{d}t\right]. \tag{22}$$

The first term represents the stream of the discounted utility derived from consumption and a given health level and a given health reference value. The second term is the expected utility at the next health shock, weighted by its instantaneous probability. The optimal share in the optimal consumption $c^\star$ and risky asset $a^\star$ are characterized using the Hamilton-Jacobi-Bellman equation

$$0 = \max_{(a,c)} \mathscr{A}U(h,b,x)+u(c_t,h,b)+\lambda(h)\,\mathbb{U}(h,b,x)-[\rho+\lambda(h)]\,U(h,b,x)\,. \tag{23}$$

The generator of the wealth process in between health shocks $\mathscr{A}$ is

$$\mathscr{A}g(x) = (x\,[a\,(\mu - r) + r + \beta h] - c)\,g_x(x) + \frac{1}{2}\,a^2\,\sigma^2\,x^2\,g_{xx}(x)\,. \tag{24}$$

The first order conditions determine the optimal strategies

$$c^\star(h,b,x) = \left(\frac{U_x(h,b,x)}{K(h,b,\gamma)}\right)^{-\frac{1}{\gamma}}, \qquad a^\star(h,b,x) = \frac{\mu - r}{\sigma^2}\,\frac{U_x(h,b,x)}{-x\,U_{xx}(h,b,x)}.$$

The value function U, for fixed h, satisfies

$$0 = \frac{1}{2}\,\frac{(\mu - r)^2}{\sigma^2}\,\frac{U_x(h,b,x)^2}{-U_{xx}(h,b,x)} + (r + \beta h)\,x\,U_x(h,b,x) + \frac{\gamma}{1-\gamma}\,U_x(h,b,x)^{\frac{\gamma-1}{\gamma}}\,K(h,b,\gamma)^{\frac{1}{\gamma}} \\ + \lambda(h)\,\mathbb{U}(h,b,x) - [\rho + \lambda(h)]\,U(h,b,x)\,.$$

Proposition 2. *Under stated assumptions, the optimal consumption and investment choices are*

$$c^\star(h,b,x) = \left(\frac{U_x(h,b,x)}{K(h,b,\gamma)}\right)^{-\frac{1}{\gamma}}, \qquad a^\star(h,b,x) = \frac{\mu - r}{\sigma^2}\,\frac{U_x(h,b,x)}{-x\,U_{xx}(h,b,x)}\,. \tag{25}$$

The value function is characterized by the ordinary differential equation

$$0 = \frac{1}{2}\,\frac{(\mu - r)^2}{\sigma^2}\,\frac{U_x(h,b,x)^2}{-U_{xx}(h,b,x)} + (r + \beta h)\,x\,U_x(h,b,x) + \frac{\gamma}{1-\gamma}\,U_x(h,b,x)^{\frac{\gamma-1}{\gamma}}\,K(h,b,\gamma)^{\frac{1}{\gamma}} \\ + \lambda(h)\,\mathbb{U}(h,b,x) - [\rho + \lambda(h)]\,U(h,b,x)\,. \tag{26}$$

This proposition relates the optimal consumption and investment choices to the value function, that is, the expected lifetime utility. Specifically, the consumption is characterized by the expected marginal consumption utility, whereas the share of risky investment depends on the inverse ratio of the first and second derivative of the value function. Effectively, this ratio poses the expected relative risk aversion over the lifetime.

4 Two special cases: Terminal illness and no health shocks

In this section, two special cases are analyzed. They provide useful boundaries and benchmark values for the general case discussed in Section 3.

4.1 Terminal illness

Terminal illness is defined as the health level at which the next health shock implies immediate death regardless of medical treatment. For analyzing utility and optimal choice in this state, fix h and note that for $h \leq H_D$ the health level is below the death threshold and hence $T_D = 0$ giving $U(h,b,x) = U_D(x)$ for all $x \geq 0$, as we assume $\gamma > 1$. Both utility of death and the expected future utility after any health shock for a critically ill agent are fixed. Hence, for a critical health level h with

$$H_D < h \leq H_D + \varepsilon_H \,, \tag{27}$$

one has $\mathbb{U}(h,b,x) = U_D(x,h,b)$, for all $x \geq 0$ regardless of medical spending. In this case, the ordinary differential equation has a semi-closed-form solution.

Proposition 3. *In the case of terminal illness, given that*

$$\frac{\rho + \lambda(h)}{\gamma} - \frac{1}{2}\frac{(\mu - r)^2}{\sigma^2}\frac{1-\gamma}{\gamma^2} - (r + \beta h)\frac{1-\gamma}{\gamma} - > 0$$

holds, the value function $U(h,b,x)$ is of the form

$$U(h,b,x) = G(h,b)\,x^{1-\gamma}, \tag{28}$$

with $G(h,b)$ determined, setting $L(h,b) = ((1-\gamma)\,G(h,b))^{-\frac{1}{\gamma}}$ (noting that $G < 0$), with $L > 0$, by

$$0 = \frac{1}{2}\frac{(\mu - r)^2}{\sigma^2}\frac{(1-\gamma)}{\gamma^2} + \frac{(r+\beta\, h)\,(1-\gamma)}{\gamma} - \frac{[\rho + \lambda(h)]}{\gamma} + L(h,b)K^{-\frac{1}{\gamma}} + \frac{\lambda(h)\,r}{\rho\gamma r^{\gamma}}e^{(1-\gamma)(h-b)}\,L(h,b)^{\gamma},$$

which has a unique solution. The optimal consumption and investment choices become

$$c^{\star} = \left(\frac{(1-\gamma)G(h,b)}{K(h,b,\gamma)}\right)^{-\frac{1}{\gamma}} x, \qquad a^{\star} = \frac{\mu - r}{\sigma^2\gamma}.$$

Medical spending equals $k^{\star} = 0$ because it is ineffective.

The proof of this proposition is presented in Appendix I.1.

4.2 No health shocks

The case of no health shocks amounts to a generalization of the famous Merton (1969) investment problem, however with a utility function which has also health and reference health among its arguments. It provides an upper bound of the agent's utility because medical spending is not necessary, permitting maximization of consumption over an infinite lifetime. It also yields a measure of the degree to which medical spending achieves this utopic case.

In a world without any health shocks, i.e, for $\lambda(h) = 0$, for all h, the utility function $U(h,b,x)$ attains an upper bound $\overline{U}(h,b,x)$. However, then the standard setting applies, implying that the upper bound is of the form $\overline{U}(h,b,x) = \overline{G}(h,b)\,x^{1-\gamma}$, for $x \geq 0$, with $\overline{G}(h,b)$ depending on health h and reference level b. We obtain the following result.[5]

[5] Here, $\mathbb{U}$ is not important as its weight $\lambda(h) = 0$.

Proposition 4. *In absence of health shocks, and under stated assumptions, the consumption and investment choices are*

$$c^{\star}(h,b,x) = \left(\frac{\rho}{\gamma} - \frac{1}{2}\left(\frac{\mu - r}{\sigma}\right)^2 \frac{1-\gamma}{\gamma^2} - (r+\beta h)\frac{1-\gamma}{\gamma}\right)x, \qquad a^{\star}(h,b,x) = \frac{\mu - r}{\sigma^2}\frac{1}{\gamma}. \quad (29)$$

The value function equals $\overline{U}(h,b,x) = \overline{G}(h,b)\,x^{1-\gamma}$ *for* $x \geq 0$ *with*

$$\overline{G}(h,b) = \frac{K(h,b,\gamma)}{(1-\gamma)\left(\frac{\rho}{\gamma} - \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1-\gamma}{\gamma^2} - (r+\beta h)\frac{1-\gamma}{\gamma}\right)^{\gamma}}. \quad (30)$$

The proof of this proposition is in Appendix I.2.

5 Numerical analysis

To obtain insight into the model's empirical prediction, we calibrate the model to the Health and Retirement Study (HRS) data, ranging biennially from 1996 - 2020.[6]

To calibrate the financial parameters of the model for the period of 1996 to 2020, the 3-Month Treasury Bill Secondary Market Rate, Discount Basis (TB3MS) delivers an annualized risk-free rate of $r = 2.1\%$.[7] The for the risky asset, the S&P 500 index as provided by the Wall Street Journal[8] features an annualized return of $\mu = 7.1\%$ and a volatility of $\sigma = 19.5\%$. Consequently, the calibrated equity premium equals 5%[9] Finally, to determine the participants' assets, we calibrate initial wealth to the individuals' financial wealth. The analysis is restricted to participants with positive financial wealth during their involvement in the study. Initial

[6]This paper employs the RAND HRS Longitudinal File 2020 (V1) from 1992 -2020 (Early Release, March 2023). Because in the first two waves, relevant health variables are not standardized to their later definitions, our analysis takes the third wave in 1996 as the starting point.

[7]The data is provided by the Board of Governors of the Federal Reserve System (US), 3-Month Treasury Bill Secondary Market Rate, Discount Basis [TB3MS], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/TB3MS, May 28, 2023.

[8]See Wall Street Journal online at https://www.wsj.com/market-data/quotes/index/SPX/historical-prices.

[9]The equity premium slightly exceeds the estimate of Yogo (2016) and Cocco (2004), presumably because of the extended time horizon featuring low interest rates.

wealth equals the median financial wealth at the time the participant enters the survey, that is, $x_0 = 34000$ USD. The wealth grid ignores the most extreme observations, that is, $x_{\min} = 600$ USD and $x_{\max} = 278800$ USD, respectively.

To determine the health-dependent income component of the asset process, we regress the participant's labour income as well as capital income relative to their financial wealth scaled by the participant's health, winsorised at one percent, resulting in $\beta = 0.36\%$.

The calibration of the health process and medical spending relies on the HRS data. Excluding individuals below 50 and above 95 (Jung, 2022) excludes biased non-representative tail observations. We consider females only in our sample because the health and mortality statistics for males and females deviate. The analysis for males is relegated to an online appendix.

The HRS data contains the subjective health status of each respondent across interview waves. Estimating the distribution and likelihood of the health shocks from this data is possible (Yogo, 2016), however, our approach, due to reference dependence, requires a viable procedure to identify objective health.[10] Weighing objective health information, for example, diagnoses by doctors and direct measurements of physical and mental health available in the HRS data in the spirit of Sprangers et al. (2000) delivers an objective health measure,[11] which then determines the health reference value via Equation (8).

The calibration of the health process requires to determine the intensity of the health shocks $\lambda(H)$ and the support $\Theta = [\underline{\theta}, \overline{\theta}]$ and distribution F_θ of the health shocks. Additionally, the functional form for the impact of medical spending in Equation (2) requires the minimal health damage, which we set to the average health damage of the highest 99.5%-quantile of the medical

[10] Most previous work in the life-cycle literature either considers the subjectively stated health as the objective health (Hugonnier, Pelgrin, and St-Amour, 2013; Rosen and Wu, 2004; Yogo, 2016) or, alternatively, imputes the reference health from the subjective health ad hoc (Harris and Kohn, 2018). However, concerns about biased assessment in presence of reference values (Groot, 2000) as well as framing effects (Crossley and Kennedy, 2002) exists.

[11] Specifically, we consider hypertension, diabetes, cancer, chronic respiratory problems, heart problems, strokes, psychological diseases, arthritis, and memory issues. Because of limited data, we eliminate back problems. Additionally, for a number of participants the medical history is disputed due to time inconsistent statements. Whenever imputed values from interviews correct the inconsistencies, we follow these. If this information is unavailable, we remove the disputed history from our sample.

spending distribution, that is, $\varepsilon_H = 0.05$. Calibrating the structure of the impact of medical spending on health shocks requires the determination of shock sizes θ, which cannot be readily observed, as observed health adjustments already include the impact of medication. The HRS data contains the medical expenditure k directly. As previously, we winsorise the medical expenditure at one percent. The observation of $\Delta H_{i,t} = H_{i,t} - H_{i,t-1}$ implies

$$\Delta H_{i,t} = \theta_i - V(k_i, \theta_i) = \theta_i - (\theta_i - \varepsilon_H)\left(1 - e^{-k_i/\theta_i}\right)$$

Here, the before medication shock size θ_i is a solution to the above equation for each $\Delta H_{i,t}$.

The preference parameters of the agent remain the final set of parameters to be calibrated. The time preference ρ is challenging to estimate directly, however, the life-cycle literature consistently provides values ranging from 5% (Hugonnier, Pelgrin, and St-Amour, 2013) over 4% (Cocco, 2004; Yogo, 2016) to 3% (Ameriks et al., 2011). Consequently, we take $\rho = 4\%$ to be our base case value.The relative risk aversion γ with respect to financial risks needs to reflect a low proportion of stock investments. We initially select $\gamma = 5$ as in Yogo (2016) but verify alternative estimates ranging from $\gamma = 2.5$ to $\gamma = 10$.Unlike related work by Hugonnier, Pelgrin, and St-Amour (2013), Yogo (2016), and Harris and Kohn (2018), our base case includes a bequest motive that (Ameriks et al., 2011) documents as a key part of individual's decision.

The calibration of the speed of the reference adaptation ω and the existence of projection bias remain pose remaining challenge of the calibration. First, the reference point requires the identification of the health reference value in the data. The first-wave observed health H_0 implies a plausible initial health reference value B_0. To ensure the health reference value has not adjusted already, the analysis considers participants without initial health issues only. The health reference value then updates according to evolution of the health process for a given ω, either with or without projection bias. To determine the adaptation speed ω, we optimize the fit an ordered probit model in terms of explaining the subjective health stated by the objective and reference health values of the survey participants over the adaptation speed.

Table 1: *Model parameters for the base case.*

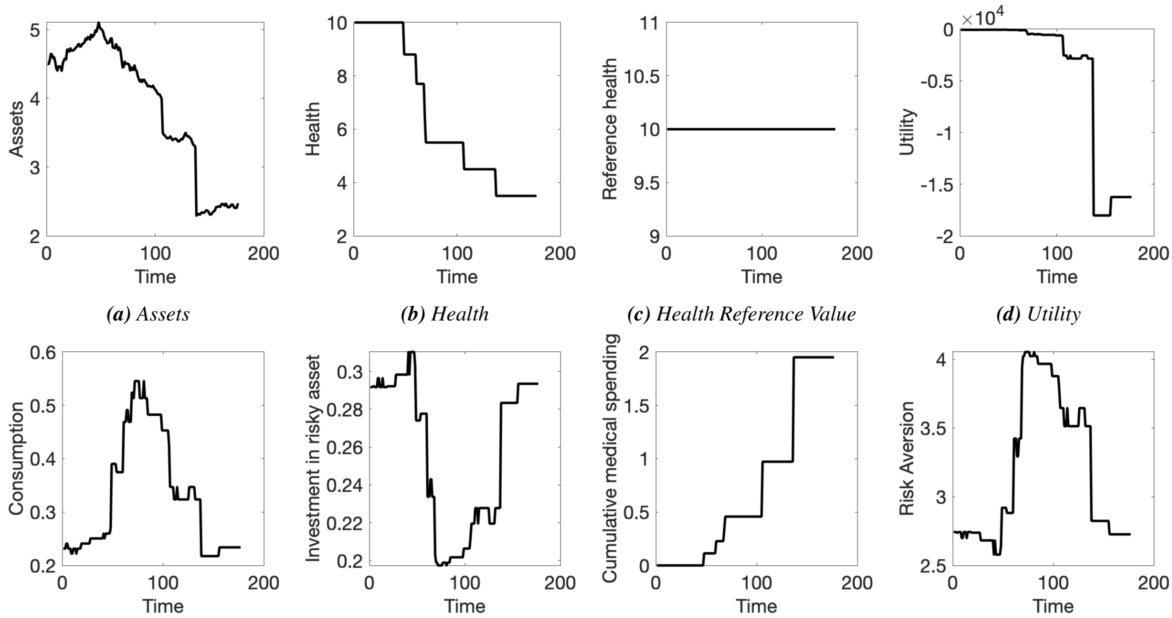
Parameter	Value	Parameter	Value
Risk-free rate	$r = 2.1\%$	Death threshold	$H_D = 0$
Mean risky return	$\mu = 7.1\%$	Initial health	$H_0 = 5$
Volatility (annualized)	$\sigma = 19.5\%$	Health shock frequency p.a.	$\lambda = 7.2\%$
Relative risk aversion	$\gamma = 7.5$	Minimum health shock	$\underline{\theta} = 1$
Initial health reference point	$B_0 = 5$	Maximum health shock	$\overline{\theta} = 5$
Discount parameter	$\rho = 4\%$	Minimum health loss	$\varepsilon_H = 0.05$

Since the second-order ordinary differential equation (26) in Proposition 2 does not have an analytical solution, it is solved using the Markov chain approximation along the lines of Kushner (1999) and Kushner and Dupuis (2001). Appendix II contains the technical details of the discretization and derivation of the necessary boundary conditions. Table 1 displays the economic parametrization of the base case. The numerical parametrization is presented in Table II.1 in Appendix II.

5.1 Static health reference value

Consider first the case of a static health reference value similar to the analysis by Harris and Kohn (2018). The development of health is evaluated against this initial value, resulting in an accumulation of health losses. Figure 1 displays a sample path of assets (Panel 1a), health (Panel 1b), and health reference values (Panel 1c) developments, accompanied by the agent's utility stream (Panel 1d). Optimal consumption (Panel 1e), share of risky investment (Panel 1f), accumulated medical spending (Panel 1g) and expected relative risk aversion over the lifetime (Panel 1h) follow. The expected relative risk aversion over the lifetime is defined below Proposition 2.

With the static reference health level equaling initial health, there is a strong desire to offset shocks already at high health levels. Consequently, health deteriorates rather slowly at the price of a fast accumulation of medical spending (see Panel 1g). To compensate the utility losses for falling health, the agent increases consumption (Panel 1e) until assets diminish sharply (Panel



(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

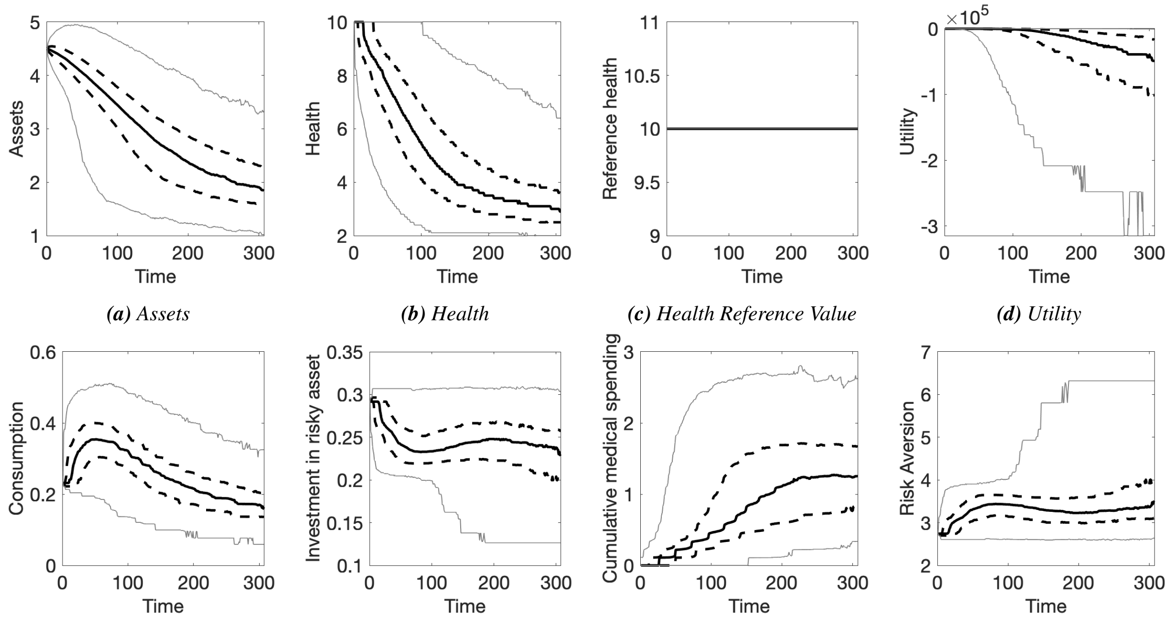
(e) Consumption *(f) Investment* *(g) Medical spending* *(h) Expected risk aversion*

Figure 1: Sample path, static health reference value. *This figure shows a sample development of the state variables, i.e., assets (Panel 1a) health, (Panel 1b), and reference health (Panel 1c), as well as the utility (Panel 1d) in the first row . The second row shows the optimal choices in consumption (Panel 1e), investment (Panel 1f), and accumulated medical spending (Panel 1g) along with the expected relative risk aversion over the lifetime (Panel 1h). The reference point does not update, that is* $\omega = 0$. *All other parameters are as in Table 1.*

1a), forcing constancy of medical spending sufficient to ensure survival albeit at a low level, and with little consumption.

Finally, as health deteriorates, the agent decreases the share of risky investments but increases it again late in life (Panel 1f). This development mirrors that of the expected lifetime risk aversion, which first rises but then falls (Panel 1h) due to two opposing influences. On the one hand, the drop in wealth causes marginal utility to increase, but on the other hand, assets drop sharply at the end of life.

While Figure 1 illustrates a specific example, Figure 2 presents a more general perspective derived from a Monte Carlo sample of 10,000 paths. Its panels depict the quantiles of the three state space variables (assets, health, and health reference value) as well as utility in the top row, and the three control variables (consumption, investment in risky assets, and medical spending)



(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

(e) Consumption *(f) Investment* *(g) Medical spending* *(h) Expected risk aversion*

Figure 2: State space and optimal choice quantiles, static health reference value. *This figure shows the quantiles of the state variables, i.e., assets (Panel 1a) health, (Panel 1b), and reference health (Panel 1c), as well as the utility (Panel 1d) in the first row. The second row displays the optimal choices in consumption (Panel 1e), investment (Panel 1f), and accumulated medical spending (Panel 1g) along with the expected relative risk aversion over the lifetime (Panel 1h) for a Monte Carlo sample. Each plot shows, for every time point, the median (solid black line), the 25- and 75 percent quantiles (dashed black lines), and the 1- and 99 percent quantiles (solid gray lines). The reference point does not update, that is* $\omega = 0$*. All other parameters are as in Table 1.*

along with expected lifetime risk aversion in the bottom row. The median is drawn as a solid black line, the 25- and 75 percent quantiles as dashed black lines, and the 1- and 99 percent quantiles as solid gray lines.

The distributions generally confirm the previous results with three exceptions. Median utility in Panel 2d does not fall as sharply as in Panel 1d, as is true for the share of risky investment (Panels 2f and 1f), and the spike in expected risk aversion disappears. (Panels 2h and 1h). However, in both Figure 1 and 2, the agent's consumption and investment behavior deviates substantially from that predicted by Merton (1969). Consumption is not a constant fraction of wealth, and the share of risky investment also varies with the individual's expected relative risk aversion which is itself endogenous here. The reason for these deviations lie in the fact that both health and the health reference value shape utility, and hence risk preference.

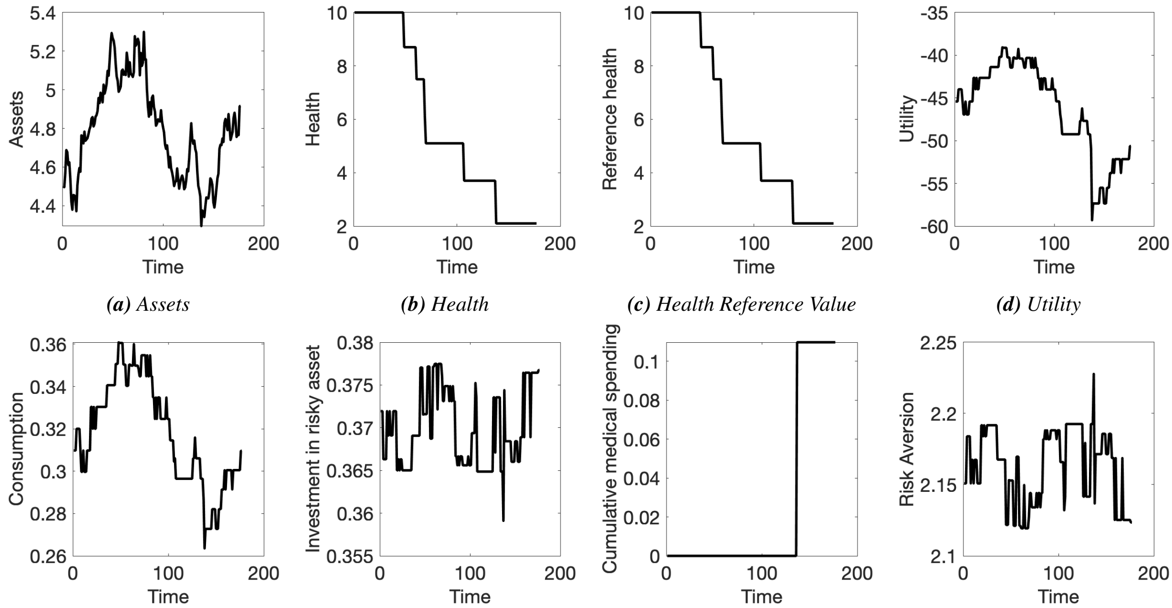
As to optimal consumption, Proposition 2 relates it to the shape of the value function, i.e., the marginal impact of consumption on both instantaneous and expected lifetime utility. At the initial level, the agent has perfect health and the health reference value coincides with it. Decaying health causes in increasing disutility both directly and indirectly through the increasing shortfall from the health reference value. The associated increase in marginal utility entails a sharp shift away from the solution of Merton (1969).

Proposition 2 also relates the optimal share of risky investment to the shape of the value function. Here, the inverse ratio of the first and second derivative of the utility function basically determine expected relative risk aversion over the lifetime. Initially, when health is perfect and its reference value coincides with it, the solution obtained by Merton (1969) prevails. As health decays, the agent's expected risk aversion increases, causing the share of risky investment to fall.

5.2 Adaptive health reference value

Here, the agent's health reference value fully adapts to the after-shock health level. Specifically, the agent knows the speed of the reference adaptation ω. The change in future utility is anticipated, causing optimal choices to be adapted. Figure 3 presents the same sample path as in the previous section, but with $\omega = 1$. Likewise, Figure 4 presents the quantiles for the Monte Carlo sample.

In Panel 3c of Figure 3, the reference health value now tracks the decay of health, in contradistinction with Panel 1c of Figure 1. Medical spending remains very low at first until it jumps to a maximum when health hits its terminal value (Panels 3h and 3b), an outcome prevented by earlier medical spending in the static case (Panels 1h and 1b). The faster deterioration of health causes utility to be lower throughout (Panels 3d and 1d), reinforced by lower consumption (Panels 3e and 1e), whose profile now follows that of assets very closely (Panels 3a and 1a), similar to the Merton (1969) model. The focus on the treatment of health shocks late in life



(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

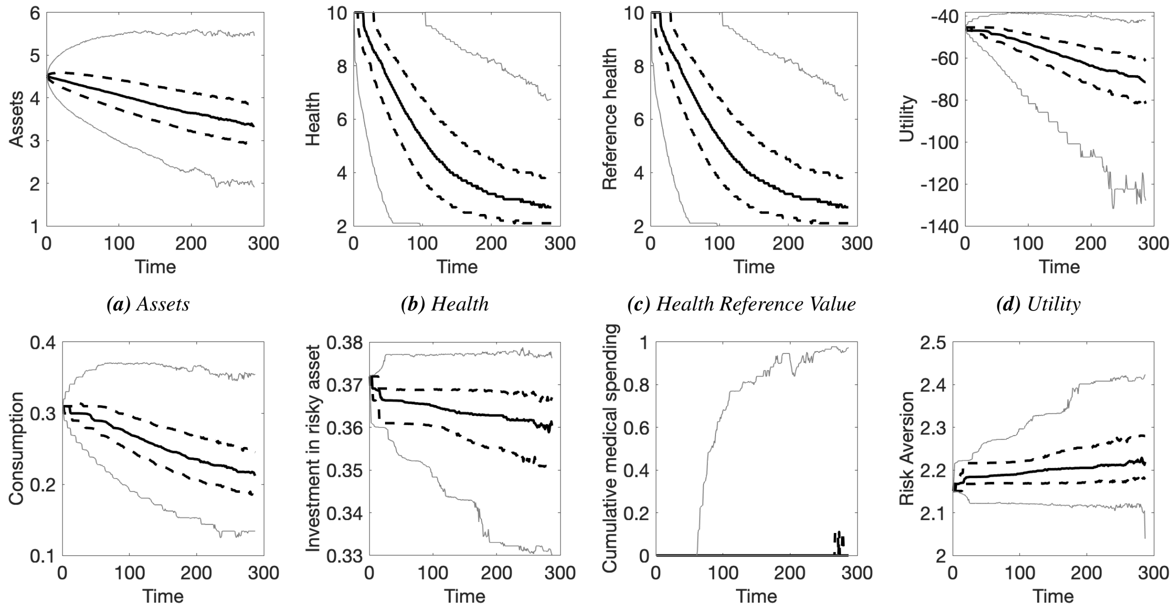
(e) Consumption *(f) Investment* *(g) Medical spending* *(h) Expected risk aversion*

Figure 3: Sample path, fully adaptive reference health with full anticipation. *This figure shows a sample development of the state variables, i.e., assets (Panel 3a) health, (Panel 3b), and reference health (Panel 3c), as well as the utility (Panel 3d) in the first row . The second row shows the optimal choices in consumption (Panel 3e), investment (Panel 3f), and accumulated medical spending (Panel 3g) along with the expected relative risk aversion over the lifetime (Panel 3h). The reference health fully adapts to health changes, that is* $\omega = 1$*. All other parameters are as in Table 1.*

results in substantially lower cumulative healthcare expenditure, at the price of a shorter lifespan. Finally, the share of risky investment is higher (Panels 3f and 1f) and remains approximately constant as in Merton (1969). The cause is an almost constant expected risk aversion (Panels 3h and 1h), which in turn is the consequence of a stable marginal utility of wealth since the health reference value instantaneously adapts to the after-shock level.

5.3 Projection bias with a fully adaptive health reference value

In this section, the health reference value immediately adapts ($\omega = 1$) as in Section 5.2. However, now the agent fails to anticipate that the health reference value adjusts to the after-shock level. Consequently, the utility reduction following a health shock is overestimated. As before,



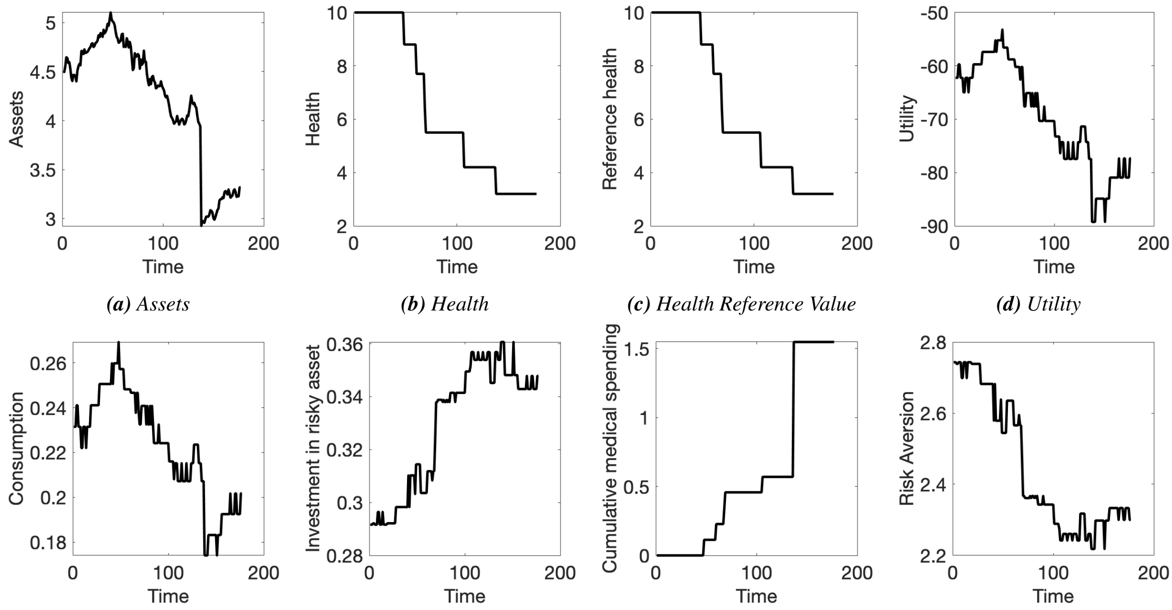
(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

(e) Consumption *(f) Investments* *(g) Medical spending* *(h) Expected risk aversion*

Figure 4: State space and optimal choice quantiles, fully adaptive reference health with anticipation. *This figure shows the quantiles of the state variables, i.e., assets (Panel 3a) health, (Panel 3b), and reference health (Panel 3c), as well as the utility (Panel 3d) in the first row. The second row displays the optimal choices in consumption (Panel 3e), investment (Panel 3f), and accumulated medical spending (Panel 3g) along with the expected relative risk aversion over the lifetime (Panel 3h) for a Monte Carlo sample. Each plot shows, for every time point, the median (solid black line), the 25- and 75 percent quantiles (dashed black lines), and the 1- and 99 percent quantiles (solid gray lines). The reference health fully updates, that is* $\omega = 1$. *All other parameters are as in Table 1.*

Figure 5 presents the case with projection bias for a sample path. Similarly, Figure 6 displays the quantiles.

Overall, the projection bias reduces both the instantaneous and the life-time utility (Panel 6d). Although the actually experienced utility for every point in the state space, that is, the same assets, health, and reference level approximately equals the one without projection bias, the agents biased extrapolation of his current preferences to the future leads to health care and consumption choices between the ones of static and adaptive reference health without the projection bias. In particular, the agent spends more on health care (Panel 6g), limits consumption (Panel 6e), and boosts health (Panel 6b) compared to the case without projection bias. Investment risk rises over the lifetime but remains below the case without projection bias (Panel 6f).

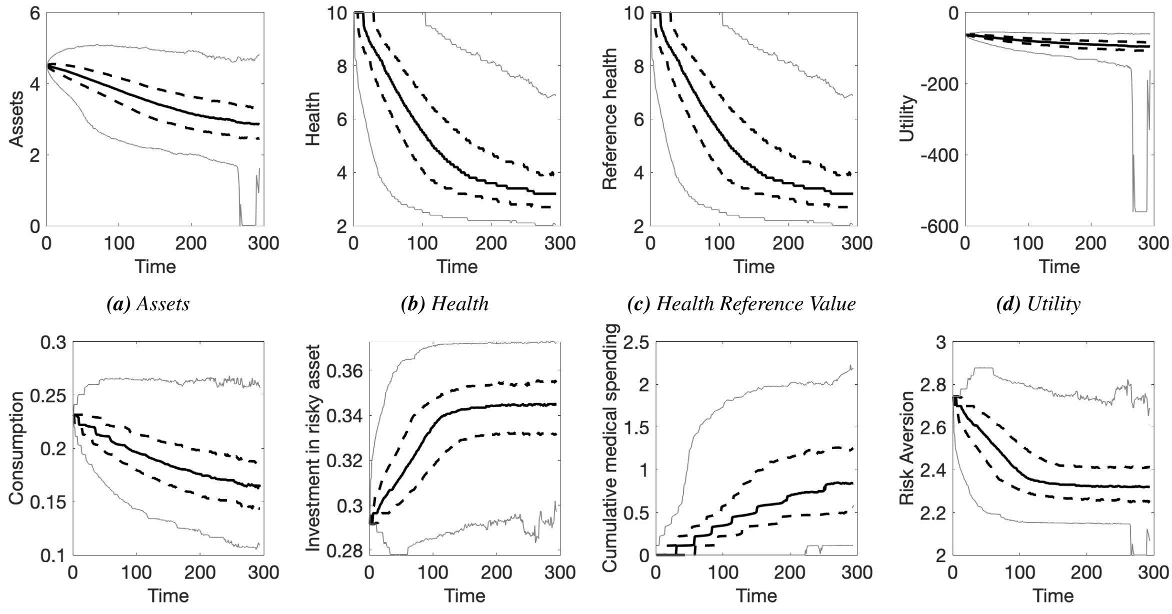


(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

(e) Consumption *(f) Investment* *(g) Medical spending* *(h) Expected risk aversion*

Figure 5: Sample path, fully adaptive reference health without anticipation. *This figure shows a sample development of the state variables, i.e., assets (Panel 5a) health, (Panel 5b), and reference health (Panel 5c), as well as the utility (Panel 5d) in the first row . The second row shows the optimal choices in consumption (Panel 5e), investment (Panel 5f), and accumulated medical spending (Panel 5g) along with the expected relative risk aversion over the lifetime (Panel 5h). The reference health fully adapts to health changes, that is* $\omega = 1$. *All other parameters are as in Table 1.*

The key difference between the adaptive reference health with and without projection bias lies in how the agent expects to feel after the health shock: With projection bias, the agent postulates to feel worse than he actually does. This reshapes the utility in a similar way as a static reference point. The agent considers the utility with a substantial scaling from the disutility from health losses relative to the projected reference health. As before, Proposition 2 implies that the optimal choices deviate from the Merton (1969) solution. This deviation structurally differs from the one for the static reference health: the expected utility shape is affected by the current health and reference health, that is, for the agent with projection bias, after a few health shocks, the reference health value gets considerably closer to the current health level. Hence, consumption and investment choices approach the ones from the case for an adaptive reference health value without projection bias, although they remain different. Medical spending, although



(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

(e) Consumption *(f) Investments* *(g) Medical spending* *(h) Expected risk aversion*

Figure 6: State space and optimal choice quantiles, fully adaptive reference health without anticipation. *This figure shows the quantiles of the state variables, i.e., assets (Panel 5a) health, (Panel 5b), and reference health (Panel 5c), as well as the utility (Panel 5d) in the first row. The second row displays the optimal choices in consumption (Panel 5e), investment (Panel 5f), and accumulated medical spending (Panel 5g) along with the expected relative risk aversion over the lifetime (Panel 5h) for a Monte Carlo sample. Each plot shows, for every time point, the median (solid black line), the 25- and 75 percent quantiles (dashed black lines), and the 1- and 99 percent quantiles (solid gray lines). The reference health fully updates, that is* $\omega = 1$*. All other parameters are as in Table 1.*

similar to the static reference case, displays a focus on late-in-life spending, that is, it is in accordance with the red herring hypothesis.

Comparing the cases of the adaptive reference health with projection bias and the static reference health, the medical spending of the adaptive reference health with projection bias resembles the static reference health one, although on a lower level. The reason for this perhaps surprising result lies in the mechanics of the projection bias: The projection bias causes the agent to systematically underestimate the utility after a health shock, which labels anything past the health shock incorrectly as a strong loss. Because of the diminishing sensitivity of our utility function, both the utility with projection bias and the static reference health case have a similar shape, although the diminishing sensitivity only partly removes the impact of the reference adaption.

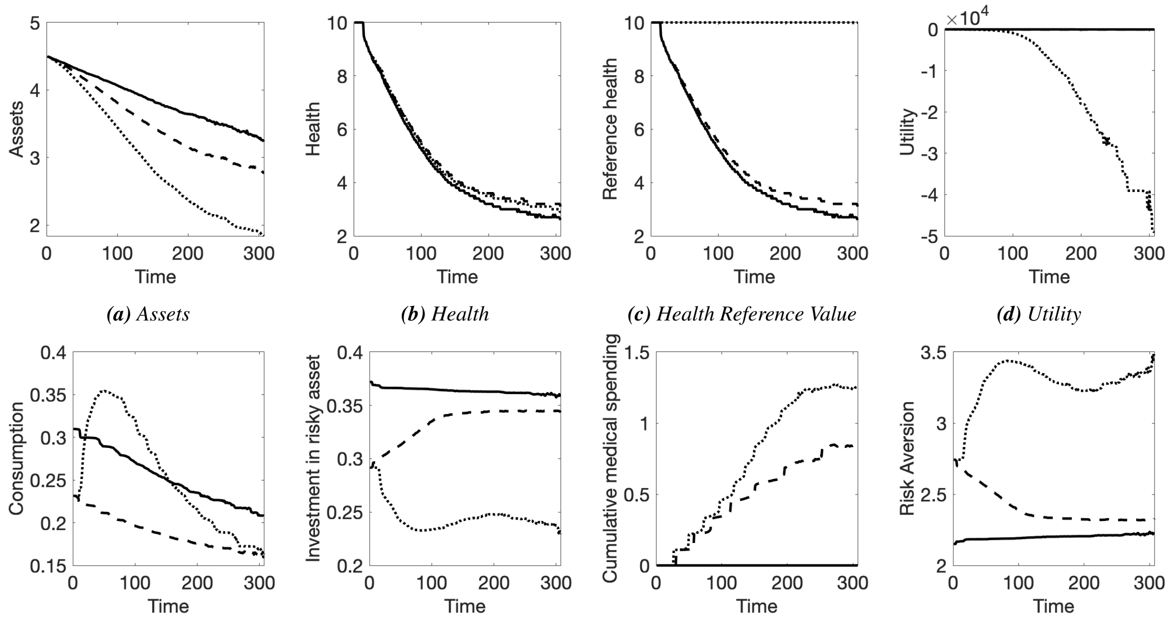
5.4 Projection bias with a partly adaptive health reference value

Whereas in the previous section, the health reference value instantaneously adapts to the after-shock level ($\omega = 1$), it adjusts only with a lag here ($\omega = 0.5$) because it is influenced by both past and present health levels.

Figure 7 ($\omega = 1$) and Figure 8 ($\omega = 0.5$) only present the median values in the case of (a) full anticipation (solid line), (b) projection bias (dashed line), and (c) static health reference value (dotted line, with $\omega = 0$). A comparing of the two figures reveals that the model's results are robust to the choice of ω. As one would expect, the profiles pertaining to (b) lie between the pertaining to (a) and (c) in both figures, with the only exception of consumption (Panels 7e and 8e). Therefore, the findings of Section 5.1 and 5.2 are confirmed. Regardless of the type of projection bias, optimal trajectories differ mainly according to whether the health reference level is static of fully adaptive.

6 Conclusion

This paper analyzes the impact of stochastic health shocks on medical spending, life health, risk aversion, consumption, and investments in the presence of a health reference point. If static, the reference point causes optimal behavior to deviate from the Merton (1969) solution in several aspects. Since the disutility of a shock is substantial, the shock triggers a spike of early medical spending designed to maintain health, an increase in risk aversion, a drop in consumption to finance health care expenditure, and a restructuring of investment in favor of the riskless asset.



(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

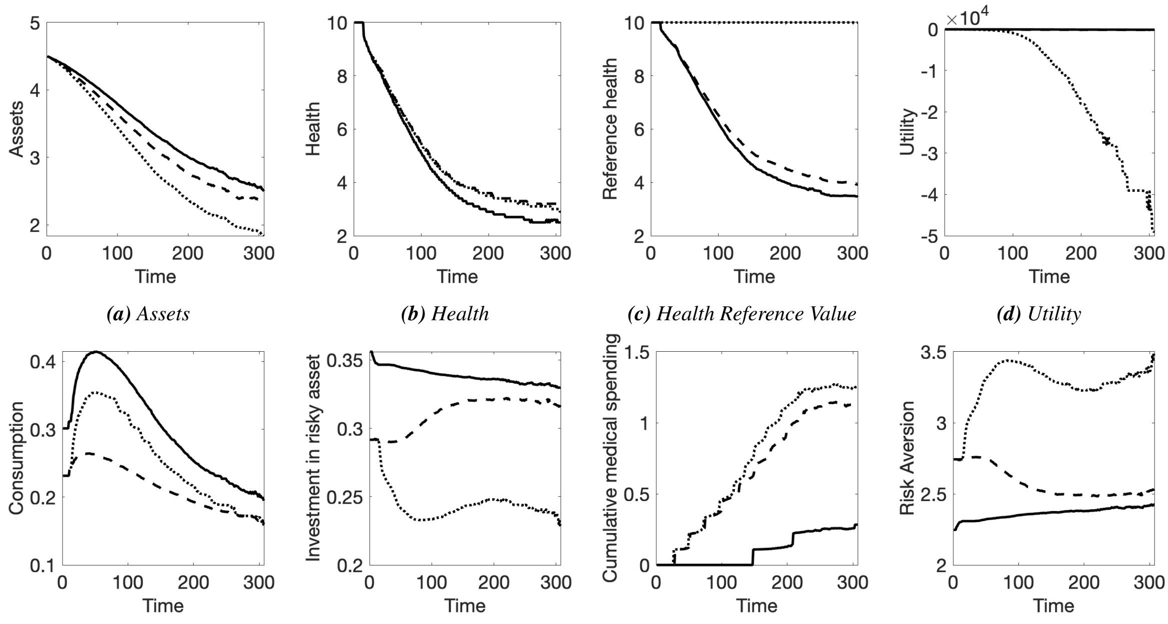
(e) Consumption *(f) Investments* *(g) Medical spending* *(h) Expected risk aversion*

Figure 7: State space and optimal choice quantiles, fully adaptive reference health without anticipation. *This figure shows the quantiles of the state variables, i.e., assets (Panel 7a) health, (Panel 7b), and reference health (Panel 7c), as well as the optimal choices in consumption (Panel 7e), investment (Panel 7f), and accumulated medical spending (Panel 7g) for a Monte Carlo sample. Each plot shows, for every time point, the median for the adaptive reference health with full anticipation (solid line) and without anticipation, i.e. the projection bias (dashed lines), and with a static reference health (dotted line). If the reference point adapts, it fully adapts to health changes, that is* $\omega = 1$. *All other parameters are as in Table 1.*

If the health reference point is fully adaptive, the optimal control values approach those of Merton (1969). Medical spending does not spike as much in response to the health shock at the price of a slightly lower life span, risk aversion increases less because the loss in utility is smaller, the fall in consumption and the restructuring of investment are mitigated.

However, in the present context projection bias is likely to affect decision making. It amounts to a failure to anticipate the impact of a health shock on utility. Accordingly, optimal choices lie between those associated with a static and a fully adaptive reference point.

This paper is subject to a number of limitations that should be addressed in future work. First, investment in prevention is neglected; it would reduce the likelihood (and hence frequency) of health shocks and/or their severity (and hence impact on the utility function). This would call for a state-dependent health production function; if healthy, the agent develops his or her health



(a) Assets *(b) Health* *(c) Health Reference Value* *(d) Utility*

(e) Consumption *(f) Investments* *(g) Medical spending* *(h) Risk Aversion*

Figure 8: State space and optimal choice quantiles, fully adaptive reference health without anticipation. *This figure shows the quantiles of the state variables, i.e., assets (Panel 8a) health, (Panel 8b), and reference health (Panel 8c), as well as the optimal choices in consumption (Panel 8e), investment (Panel 8f), and accumulated medical spending (Panel 8g) for a Monte Carlo sample. Each plot shows, for every time point, the median for the adaptive reference health with full anticipation (solid line) and without anticipation, i.e. the projection bias (dashed lines), and with a static reference health (dotted line). If the reference point adapts, it fully adapts to health changes, that is* $\omega = 0.5$*. All other parameters are as in Table 1.*

effort; if ill, medical care is required. Next, a health shock may not only cause an increase in risk aversion but also in time preference since survival is less likely.

For the same reason, the planning horizon becomes shorter, which has an effect on all optimal control values. A health shock is known to make (early) retirement more likely, with the concomitant fall in the opportunity cost of time (and hence preventive effort). Finally, there is no fear of death, the disutility of which presumably falls with the number and severity of health shocks to an extend again depending on the health reference point. In spite of this limitations, this contribution demonstrates the importance of a reference point in terms of health and its adaptiveness when the objective is to model optimal responses to optimal health shocks.

References

Acland, Dan and Matthew R. Levy (2015). “Naiveté, Projection Bias, and Habit Formation in Gym Attendance”. In: *Management Science* 61.1, pp. 146 –160.

Ameriks, John, Andrew Caplin, Steven Laufer, and Stijn van Nieuwerbergh (2011). “The Joy of Giving or Assisted Living? Using Strategic Surveys to Separate Public Care Aversion from Bequest Motives”. In: *The Journal of Finance* 66.2, pp. 519–561.

Baucells, Manel, Martin Weber, and Frank Welfens (2011). “Reference-Point Formation and Updating”. In: *Management Science* 57.3, pp. 506–519.

Bilsen, Servaas van, Roger J. A. Laeven, and Theo E. Nijman (2020). “Consumption and Portfolio Choice Under Loss Aversion and Endogenous Updating of the Reference Level”. In: *Management Science* 66.9, pp. 3927–3955.

Budhiraja, Amarjit and Kevin Ross (2007). “Convergent Numerical Scheme for Singular Stochastic Control with State Constraints in a Portfolio Selection Problem”. In: *SIAM Journal of Control and Optimization* 45.6, pp. 2169 –2206.

Cocco, João F. (2004). “Portfolio Choice in the Presence of Housing”. In: *The Review of Financial Studies* 18.2, pp. 535–567.

Crossley, Thomas F. and Steven Kennedy (2002). “The reliability of self-assessed health status”. In: *Journal of Health Economics* 21, pp. 643 –658.

De Giorgi, Enrico G. and Thierry Post (2011). “Loss Aversion with a State-Dependent Reference Point”. In: *Management Science* 57.6, pp. 1094–1110.

Eeckhoudt, Louis, Beatrice Rey, and Harris Schlesinger (2007). “A Good Sign for Multivariate Risk Taking”. In: *Management Science* 53.1, pp. 117–124.

Gottlieb, Daniel (2012). "Prospect Theory, Life Insurance, and Annuities". Working Paper, Wharton School, University of Pennsylvania.

Groot, Wim (2000). "Adaptation and scale of reference bias in self-assessments of quality of life". In: *Journal of Health Economics* 19, pp. 403 –420.

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

Harris, M. C. and J. L. Kohn (2018). "Reference health and the demand for medical care". In: *The Economic Journal* 128.615, pp. 2812–2842.

He, Xue Dong and Moris S. Strub (2022). "How Endogenization of the Referene Point Affects Loss Aversion: A Study of Portfolio Selection". In: *Operations Research* 70.6, pp. 3035–3053.

Hugonnier, Julien, Florian Pelgrin, and Pascal St-Amour (2013). "Health and (Other) Asset Holdings". In: *Review of Economic Studies* 80, pp. 553–710.

Jung, Jürgen (2022). "Estimating transition probabilities between health states using US longitudinal survey data". In: *Empirical Economics* 63, pp. 901–943.

Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk". In: *Econometrica* 47.2, pp. 263–292.

Kliger, Doron and Ori Levy (2008). "Projection bias by investors: A market approach". In: *Journal of Economic Behavior & Organization* 66, pp. 657 –668.

Köszegi, Botond and Matthew Rabin (2006). "A Model of Reference-Dependent Preferences". In: *The Quarterly Journal of Economics* 121.4, pp. 1133–1165.

— (2007). "Reference-Dependent Risk Attitudes". In: *American Economic Review* 97.4, pp. 1047–1073.

Kushner, Harold J. (1999). “Consistency Issues for Numerical Methods for Variance Control, with Applications to Optimization in Finance”. In: *IEEE Transaction on Automatic Control* 44.12, pp. 2283 –2296.

Kushner, Harold J. and Paul Dupuis (2001). *Numerical Methods for Stochastic Control Problems in Continuous Time*. Second edition. Vol. 24. Applications of Mathematics. New York, NY: Springer.

Loewenstein, G., T. O’Donoghue, and M. Rabin (2003). “Projection bias in predicting future utility”. In: *Quarterly Journal of Economics* 118.4, pp. 1209–1248.

Merton, Robert C. (1969). “Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case”. In: *Review of Economics and Statistics* 51.3, pp. 247–257.

Rosen, Harvey S. and Stephen Wu (2004). “Portfolio choice and health status”. In: *Journal of Financial Economics* 72, pp. 457 –484.

Sprangers, Mirjam A.G. et al. (2000). “Which chronic conditions are associated with better or poorer quality of life?” In: *Journal of Clinical Epidemiology* 52, pp. 895 –907.

Strub, Moirs. S. and Duan Li (2020). “Failing to Foresee the Updating of the Reference PointLeads to Time-Inconsistent Investment”. In: *Operations Research* 68.1, pp. 199–213.

Thakral, Neil and T. To Linh (2021). “Daily Labor Suppy and Adaptive Reference Points”. In: *American Economic Review* 111.8, pp. 2417 –2443.

Tversky, Amos and Daniel Kahneman (1992). “Advances in Prospect Theory: Cumulative Representation of Uncertainty”. In: *Journal of Risk and Uncertainty* 5.4, pp. 297–323.

Yogo, Motohiro (2016). “Portfolio Choice in retirement: Health risk and the demand for annuities, housing, and risky assets”. In: *Journal of Monetary Economics* 80, pp. 17–34.

Zweifel, Peter, Stefan Felder, and Markus Meier (1999). "Ageing of Population and Health Care Expenditure: A Red Herring?" In: *Health Economics* 8, pp. 485–496.

Zweifel, Peter, Stefan Felder, and Andreas Werblow (2004). "Population Ageing and Health Care Expenditure: New Evidence on the "Red Herring"". In: *The Geneva Papers on Risk and Insurance* 29.4, pp. 652–666.

Reference Health and Investment Decisions
Appendix

Chunli Cheng Christian Hilpert Alexander Szimayer Peter Zweifel

List of Appendices

Appendix I Special Cases

I.1 Derivation of the solution for the critically ill agent

Given the agent is critically ill, the value function $U(h,b,x)$ is of the form

$$U(h,b,x) = G(h,b)\,x^{1-\gamma}, \tag{I.1}$$

with $G(h,b) < 0$ for $\gamma > 1$. To see this, for the partials of Equation I.4

$$U_x(h,b,x) = (1-\gamma)\,G(h,b)\,x^{-\gamma} \text{ and } U_{xx}(h,b,x) = -\gamma(1-\gamma)\,G(h,b)\,x^{-\gamma-1}. \tag{I.2}$$

Substitution in Equation (26) of the text yields

$$\begin{aligned} 0 = & \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{(1-\gamma)^2\,G(h,b)^2\,x^{-2\gamma}}{\gamma(1-\gamma)\,G(h,b)\,x^{-\gamma-1}} + (r+\beta h)\,(1-\gamma)\,G(h,b)\,x^{1-\gamma} \\ & + \frac{\gamma}{1-\gamma}K(h,b)^{\frac{1}{\gamma}}\left((1-\gamma)\,G(h,b)\,x^{-\gamma}\right)^{\frac{\gamma-1}{\gamma}} \\ & - [\rho+\lambda(h)]\left[G(h,b)\,x^{1-\gamma}\right] - \lambda(h)\frac{r}{\rho r^{\gamma}(\gamma-1)}G(h,b)^{-1}e^{(h-b)(1-\gamma)}. \end{aligned}$$

This can be divided by $G(h,b)\,x^{1-\gamma}$ to obtain

$$0 = \frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1-\gamma}{\gamma} + (r+\beta h)\,(1-\gamma) + \gamma\left(\frac{K(h,b)}{(1-\gamma)\,G(h,b)}\right)^{\frac{1}{\gamma}} - [\rho+\lambda] - \frac{\lambda r e^{(h-b)(1-\gamma)}}{\rho r^{\gamma}(\gamma-1)}G(h,b)^{-1}.$$

Rearranging gives

$$\frac{\rho+\lambda}{\gamma}-\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1-\gamma}{\gamma^2}-(r+\beta h)\frac{1-\gamma}{\gamma}-\left(\frac{K(h,b)}{(1-\gamma)\,G(h,b)}\right)^{\frac{1}{\gamma}}=\frac{\lambda r e^{(h-b)(1-\gamma)}}{\rho r^{\gamma}(\gamma-1)}G(h,b)^{-1}$$

Set $L(h,b)=((1-\gamma)G(h,b))^{-\frac{1}{\gamma}}$ (noting that $G(h,b)<0$), then $L(h,b)>0$ and

$$\frac{\rho+\lambda}{\gamma}-\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1-\gamma}{\gamma^2}-(r+\beta h)\frac{1-\gamma}{\gamma}-(K(h,b))^{\frac{1}{\gamma}}L(h,b)=\frac{\lambda r e^{(h-b)(1-\gamma)}}{\rho r^{\gamma}}L(h,b)^{\gamma}, \quad \text{(I.3)}$$

which has exactly one positive solution.

I.2 Derivation for no health shocks

Given the agent faces no health shocks, that is, $\lambda(H)=0$, the value function $\overline{U}(h,b,x)$ is of the form

$$\overline{U}(h,b,x)=\overline{G}(h,b)\,x^{1-\gamma}, \quad \text{(I.4)}$$

with $\overline{G}(h,b)<0$ for $\gamma>1$. To see this, for the partials of Equation I.4

$$\overline{U}_x(h,b,x)=(1-\gamma)\,\overline{G}(h,b)\,x^{-\gamma} \text{ and } \overline{U}_{xx}(h,b,x)=-\gamma(1-\gamma)\,\overline{G}(h,b)\,x^{-\gamma-1}. \quad \text{(I.5)}$$

Substitution in Equation (26) of the text yields

$$0=\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{(1-\gamma)^2\,\overline{G}(h,b)^2\,x^{-2\gamma}}{\gamma(1-\gamma)\,\overline{G}(h,b)\,x^{-\gamma-1}}+(r+\beta h)\,(1-\gamma)\,\overline{G}(h,b)\,x^{1-\gamma}$$
$$+\frac{\gamma}{1-\gamma}K(h,b)^{\frac{1}{\gamma}}\left((1-\gamma)\,\overline{G}(h,b)\,x^{-\gamma}\right)^{\frac{\gamma-1}{\gamma}}-\rho\overline{G}(h,b)\,x^{1-\gamma}.$$

Table II.1: Computational parameters for the base case.

Parameter	Value	Parameter	Value
Health steps	41	Asset steps	101
Health step size	0.25	Minimal asset value	0.5
Reference grid steps	41	Reference step size	0.25
Health shock grid steps	7	Health shock step size	0.25
Minimal asset value	0.5	Maximal asset value	5.0

This can be divided by $\overline{G}(h,b)\,x^{1-\gamma}$ to obtain

$$0 = \frac{1}{2}\,\frac{(\mu-r)^2}{\sigma^2}\,\frac{1-\gamma}{\gamma} + (r+\beta h)\,(1-\gamma) - \gamma\left(\frac{K(h,b)}{(1-\gamma)\,\overline{G}(h,b)}\right)^{\frac{1}{\gamma}} - \rho.$$

Rearranging gives

$$\frac{\rho}{\gamma} - \frac{1}{2}\,\frac{(\mu-r)^2}{\sigma^2}\,\frac{1-\gamma}{\gamma^2} - (r+\beta h)\,\frac{1-\gamma}{\gamma} - \left(\frac{K(h,b)}{(1-\gamma)\,\overline{G}(h,b)}\right)^{\frac{1}{\gamma}} = 0.$$

Solving for $\overline{G}$ results in

$$\overline{G}(h,b) = \frac{K(h,b,\gamma)}{(1-\gamma)\,\left(\frac{\rho}{\gamma} - \frac{1}{2}\,\frac{(\mu-r)^2}{\sigma^2}\,\frac{1-\gamma}{\gamma^2} - (r+\beta h)\,\frac{1-\gamma}{\gamma}\right)^{\gamma}}.$$

Plugging this constant into the the consumption and investment choices delivers

$$c^\star(h,b,x) = \left(\frac{\rho}{\gamma} - \frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2 \frac{1-\gamma}{\gamma^2} - (r+\beta h)\frac{1-\gamma}{\gamma}\right)x, \qquad a^\star(h,b,x) = \frac{\mu-r}{\sigma^2}\,\frac{1}{\gamma}.$$

which concludes the proof.

Appendix II Numerical Method

This appendix addresses the question how to determine the value function $U(h,b,x)$ and based on this the optimal strategy $(a^\star, c^\star, k^\star)$ numerically. Proposition 2 characterizes the value function via an ordinary differential equation, and the optimal consumption $c^\star$ and investment $a^\star$ choices whereas Proposition determines the optimal medical spending choice $k^\star$; however, neither choices nor value function are available in closed form.

The key numerical method to determine the value function as well as consumption and investment choices is the approximating Markov chain method following Kushner (1999) and Kushner and Dupuis (2001). A direct optimization delivers the optimal medical spending choice that is numerically distinct from the remaining optimization because medication only appears at health shocks, whereas the approximating Markov chain describes choices between health shocks.

A key observation from a numerical point of view is that Proposition 2 characterizes the value function U via an ordinary differential equation in the asset space x by holding both health h and reference health value b constant. Furthermore, for critical health, the value function U is available in semi-closed form. Finally, the value function U depends on the expected after-shock utility $\mathbb{U}$. Because health shocks imply a permanent health reduction, $\mathbb{U} < U$, pointing to an iterative scheme to determine U for which health iterates upwards from critical to perfect health.

Thus, the overall structure to compute the value function U is as follows: For $h \leq H_D$, that is, health below the death threshold, the value function equals the death utility, that is, $U_D(h,b,x)$, for all $x \geq 0$. For the poor health states, $H_D < h \leq H_D + \varepsilon_H$, the value function U features a semi-analytical form given by (I.4) with $G(h,b)$ determined in equation (I.3). For better-than-critical health $h > H_D + \varepsilon_H$, $\mathbb{U}$ is computed using the already known $U(h',b,x,)$ for $h' \leq h - \varepsilon_H$ where simultaneously the optimal health expenditure $k^\star(h,b,x,\theta)$ is solved for using (18). The value function $U(h,b,x)$ is computed using an approximating Markov chain based on the ODE (26) for a given health h and reference b, which is then repeated for increasing h, and for each h, the admissible range of reference values b, that is, $b \geq h$. The structure of the value function in (I.4) is in general not preserved when $\mathbb{U}(h,b,x)$ is computed, thus no closed-form solution exists.

This appendix proceeds as follows: First, we describe the discretization of the state space and the health shocks. Then, the determination of the medical spending choice precedes the computation of the value function and the investment and consumption choices.

II.1 Discretization and boundary conditions

Consider the discretization of the state space and the health shock distribution. Let us denote the domain the distribution F_θ describing the health shocks by $[\underline{\theta}, \overline{\theta}]$, with $\varepsilon_H \leq \underline{\theta} < \overline{\theta}$. Thus, if a health shock occurs, its smallest possible impact $\underline{\theta}$ exceeds the value ε_H.

Define the grid of the state space (h,b,x) by the respective mesh Δh and Δx. For the number of asset steps N_x let

$$x_i = i\Delta x\,, \qquad \text{for } i = 0,\ldots,N_x\,. \tag{II.1}$$

For health, numerical analysis requires two numbers of health steps $\underline{N}_h$ and $\overline{N}_h$. The parameter $\overline{N}_h$ represents the number of health steps above the death threshold. The second parameter $\underline{N}_h$ adds additional health steps below the death threshold. We define

$$h_i = H_D + i\Delta h\,, \qquad \text{for } i = -\underline{N}_h,\ldots,0,\ldots,\overline{N}_h\,. \tag{II.2}$$

Similarly,

$$b_m = H_D + m\Delta h\,, \qquad \text{for } m = 0,\ldots,\overline{N}_h\,. \tag{II.3}$$

Health states below the death threshold H_D are required for a formal reason. In case a health shock for $h \approx H_D$ occurs, the utility is recursively calculated by jump conditions from utility at states with a lower health level. That is, the value function U depends on $\mathbb{U}$ which needs to be defined at a specified range of lower health levels. Thus, $\underline{N}_h = \overline{\theta}/\Delta h$ such that $H_D - \underline{N}_h \Delta h = H_D - \overline{\theta}$ covers the biggest possible health shock $\overline{\theta}$ from the lowest alive state H_D.

Next, the support $[\underline{\theta}, \overline{\theta}]$ of the health shock distribution F_θ is discretized by

$$\theta_i = i\Delta h\,, \qquad \text{for } i = \underline{N}_\theta, \ldots, \overline{N}_\theta\,, \tag{II.4}$$

with $(\underline{N}_\theta - 1)\,\Delta h < \underline{\theta} \leq \underline{N}_\theta\,\Delta h$ and $\overline{N}_\theta\,\Delta h < \overline{\theta} \leq (\overline{N}_\theta + 1)\,\Delta h$. The corresponding probabilities are determined by

$$p_i = \mathbb{P}(\theta_i - 0.5\,\Delta h < \theta \leq \theta_i + 0.5\,\Delta h)\,,\ \text{for } i = \underline{N}_\theta + 1, \ldots, \overline{N}_\theta - 1\,, \tag{II.5}$$

and $p_{\underline{N}_\theta} = \mathbb{P}(\theta \leq \theta_{\underline{N}_\theta} - 0.5\,\Delta h)$ as well as $p_{\overline{N}_\theta} = \mathbb{P}(\theta_{\overline{N}_\theta} - 0.5\,\Delta h < \theta)$.

We initialize the numerical approach by observing that in death states $h_i \leq H_D$, or, $i \leq 0$

$$\hat{U}(h_i, b_m, x_j) = U_D(h_i, b_m, x_j)\,, \qquad \text{for } i = -\underline{N}_h, \ldots, 0,\ m = 0, \ldots \overline{N}_h, \text{ and } j = 0, \ldots, N_x\,. \tag{II.6}$$

Also, for the zero wealth state $x_0 = 0$, the agent has zero utility as no further consumption can be financed. In this case, the presence of health shocks has negligible impact on utility, as neither consumption nor medical spending can be financed and we assume that the structure of the actual utility approximately equals

$$\hat{U}(h_i, b_m, x_0) \approx G(h, b, \gamma) x_0^{1-\gamma}, \qquad \text{for } i = 1, \ldots, \overline{N}_h \text{ and } m = i, \ldots, \overline{N}_h. \tag{II.7}$$

The second step of the initialization is to look at poor health levels h_i, with $H_D < h_i \leq H_D + \varepsilon_H$ according to (27). The approximation of the value function $\hat{U}$ follows from (I.4) and is given by

$$\hat{U}(h_i, b_m, x_j) = G(h_i, b_m)\, x_j^{1-\gamma}, \qquad \text{for } i = 1, \ldots, N_h^\star, m \geq i, \ldots, N_h^\star, \text{ and } j = 1, \ldots, N_x\,, \tag{II.8}$$

where G follows from the discretization

$$\left(\frac{\rho+\lambda}{\gamma}-\frac{1}{2}\frac{(\mu-r)^2}{\sigma^2}\frac{1-\gamma}{\gamma^2}-(r+\beta h)\frac{1-\gamma}{\gamma}\right)+K(h_i,b_m)^{1/\gamma}L(h_i,b_m)=\frac{\lambda r e^{(h_i-b_m)(1-\gamma)}}{\rho r^{\gamma}}L(h_i,b_m)^{\gamma}, \tag{II.9}$$

which determines L and subsequently G numerically. $N_h^\star$ is characterized by $\Delta h\,(N_h^\star-1)<\overline{\theta}\leq \Delta h\,N_h^\star$.

II.2 Optimal medical spending

We now need to solve the optimal medical spending policy before we can turn to solving the differential equation (26) in x for each given h and b numerically. This reason is that solving for U requires the computation of the health jump effect given by $\mathbb{U}$ along the lines of (21) beforehand, which in turn is based on the optimal investment in medical treatment $k^\star$ given a health shock as outline in (18).

To address the optimal medical treatment $k^\star$, we define for a given health shock $\theta=i\Delta h$, with $i=\underline{N}_\theta,\ldots,\overline{N}_\theta$ the investment $k_j(\theta_i)$ to generate a beneficial effect on the health level of magnitude $j\Delta h$ by

$$k_l(\theta_i)=\theta_i\ln\frac{\theta_i-\varepsilon_H}{\theta_i-\varepsilon_H-l\,\Delta h},\qquad \text{for } l=0,\ldots,\overline{N}_k(\theta_i), \tag{II.10}$$

with $\theta_i-\varepsilon_H-\Delta h<\overline{N}_k(\theta_i)\Delta h\leq\theta_i-\varepsilon_H$, such that $V(k_l(\theta_i),\theta_i)=l\,\Delta h$, or, equivalently, $\theta_i-V(k_j(\theta_i),\theta_i)=(i-j)\,\Delta h$. Discretizing the medical spending $k_l(\theta_i)$ in this way ensures that the medicated after-shock health remains on the discretized grid of h. To prepare the numerical equivalent to (18), define

$$l^\star(h_i,b_m,x,\theta_j,\omega)=\underset{l=0,\ldots,\overline{N}_k(\theta_j)}{\arg\max}\ \hat{U}(h_i-\theta_j+V(k_l(\theta_j),\theta_j),f(h_i,b_m,\omega,k_l(\theta,x-k_l(\theta_j)))\,, \tag{II.11}$$

where $\hat{U}$ is interpolated linearly in x. In case the maximum in (II.11) is not unique, the largest expense is selected resulting in the longest life with identical utility. The numerical equivalent to (18) is thus

$$k^\star(h_i, b_m, x, \omega, \theta_j) = k_{l^\star(h_i, b_m, x, \omega, \theta_j)}(\theta_i)\,. \tag{II.12}$$

Then, for $(h_i, b_m, x_j) = (i\Delta h, m\Delta h, j\Delta x)$, the numerical version of (21) becomes

$$\hat{\mathbb{U}}(h_i, b_m, x_j) = \sum_{l=\underline{N}_\theta}^{\overline{N}_\theta} \hat{U}((i-l+l^\star(h_i, x_j, \theta_l))\Delta h, (1-\omega)m\Delta h + \omega(i-l+l^\star(h_i, x_j, \theta_l))\Delta h, x - k^\star(h_i, x_j, \theta_l))\, p_l\,, \tag{II.13}$$

where $\hat{U}$ is interpolated linearly in x. Now, we are equipped to solve the second order ordinary differential equation (26).

II.3 Approximating Markov chain

To determine the solution of the general ordinary differential equation in Proposition 2, we restate the problem in the framework of Kushner (1999) and Kushner and Dupuis (2001). For doing so, we need a boundary condition for the value function for vanishing wealth. It is reasonable to assume $U(h, b, x) \approx G(h, b)x^{1-\gamma}$. Vanishing wealth in the discretized setup is given by $x = \Delta x = x_1$. For $i = N_h^\star + 1, \ldots, \overline{N}_h$, we approximate

$$\hat{U}(h_i, x_1) = G(h_i, b_m)x_1^{1-\gamma}\,. \tag{II.14}$$

Thus, (II.14) presents the needed boundary conditions that is specified in terms of $G(h_i, b_m)$.

To define an approximating Markov chain, we require an underlying stochastic differential equation (SDE), a target function, and appropriate boundary conditions. The asset process evolves as

$$dX_t = [(a(\mu - r) + r + \beta h)X_t - c_t]\, d_t + \sigma a_t X_t dW_t,$$

which poses the underlying SDE. The target function equals the utility value function which remains

$$U(h,b,x) = E\left[\int_0^\infty e^{-(\rho+\lambda)t}\left(u(c_t,h,b)+\lambda\mathbb{U}(h,b,X_t)\right)dt\right].$$

We define the adjusted discount factor $\beta = \rho+\lambda$. In terms of the notation of Kushner and Dupuis (2001), the instantaneous payout equals $u(c_t,h,b)+\lambda\mathbb{U}(h,b,X_t)$. The lower boundary arises from the case of vanishing wealth for x_1. The upper boundary appears for x_{N_x} and is handled via a reflection condition in the approximating Markov chain, which we discuss later.

Following Kushner and Dupuis (2001), a controlled approximating Markov chain for the regular asset grid under consideration features the control-dependent probabilities for up and down transitions as[12]

$$p^{\tilde{h}}(x,x\pm\tilde{h}|a,c) = \frac{\sigma^2a^2x^2/2+\tilde{h}\left[(a(\mu-r)+r+\beta h)x-c\right]^{\pm}}{\sigma^2a^2x^2+\tilde{h}\left|(a(\mu-r)+r+\beta h)x-c\right|}, \tag{II.15}$$

with $y^+ := \max(y,0)$ and $y^- := -\max(-y,0)$. This formulation implies the dynamic programming equation

$$U^{\tilde{h}}(x) = \max_{(a,c)}\left\{e^{-\beta\Delta t^{\tilde{h}}}\sum_y p^{\tilde{h}}(x,y|a,c)U^{\tilde{h}}(y)+\left(u(c,h,b)+\lambda\mathbb{U}(h,b,x)\right)\Delta t^{\tilde{h}}(x|a,c)\right\},$$

where the time step equals

$$\Delta t^{\tilde{h}}(x|a,c) = \frac{\tilde{h}^2}{\sigma^2a^2x^2+\tilde{h}\left|(a(\mu-r)+r+\beta h)x-c\right|}. \tag{II.16}$$

An issue in our numerical approach lies in the control-dependent variance. Following Kushner (1999), because the variance is controlled by the investment control a, to stabilize the

[12]Usually, the grid is denoted by h, which we already use for health, hence, the grid parameters is $\tilde{h}$.

numerical computation, we employ

$$\overline{Q}^{\tilde{h}} = \max_{x,a,c} Q^{\tilde{h}}(x,a,c) = \max_{x,a,c} \sigma^2 a^2 x^2 + \tilde{h}\,|(a(\mu - r) + r + \beta h)x - c| \tag{II.17}$$

and fix the time-step size to

$$\overline{\Delta} t^{\tilde{h}} = \frac{\tilde{h}^2}{\overline{Q}^{\tilde{h}}}. \tag{II.18}$$

This delivers the approximated value function

$$U^{\tilde{h}}(x,a_0,c_0) = e^{-\beta t^{\tilde{h}}} \sum_{y} \overline{p}^{\tilde{h}}(x,y|a_0(x),c_0(x))U^{\tilde{h}}(y,a_0,c_0) + \left(\frac{c_0^{1-\gamma}}{1-\gamma} + \lambda \mathbb{U}(x,h,b)\right)\overline{\Delta} t^{\tilde{h}} \tag{II.19}$$

for initial controls a_0 and c_0. This expression is equivalent to

$$U^{\tilde{h}}(x,a_0,c_0) = \frac{e^{-\beta t^{\tilde{h}}} \sum_{y \neq x} \overline{p}^{\tilde{h}}(x,y|a_0(x),c_0(x))U^{\tilde{h}}(y,a_0,c_0) + g(x,h,b)\overline{\Delta} t^{\tilde{h}}}{1 - e^{-\beta t^{\tilde{h}}}\overline{p}^{\tilde{h}}(x,x|a_0(x),c_0(x))} \tag{II.20}$$

with $g(x,h,b) = \frac{c_0^{1-\gamma}}{1-\gamma} + \lambda \mathbb{U}(x,h,b)$. In this formulation, we have included the after-health shock expected utility in the instantaneous payoff, adjusted with the probability of the shock occurring.

A final consideration poses the reflecting boundary for as the asset value grows. As for the highest asset level x_{N_x}, the transition probability for further rises lies outside the defined grid, we follow Budhiraja and Ross (2007) in imposing a reflection step with zero time length $\Delta t = 0$ that reflects the approximating Markov chain to the upmost state again.

The remaining algorithm satisfies the conditions for the standard Merton problem following the algorithm of Kushner (1999). For every health level exceeding critical health, that is, $h_i \geq H_D + \varepsilon_H$, and every admissible health reference value $b_m \geq h_i$, we use the approximation in the policy space for the value function (Kushner, 1999). As long as the error exceeds a set tolerance and the number of iterations does not reach a set limit, we recompute equation (II.20) for reoptimized controls (a_n, c_n) for iteration n until the value function converges.