Closing Down the Shop: Optimal Health and Wealth Dynamics near the End of Life*

Julien Hugonnier\textsuperscript{1,4,5}, Florian Pelgrin\textsuperscript{2,6} and Pascal St-Amour\textsuperscript{3,4,6}

\textsuperscript{1}École Polytechnique Fédérale de Lausanne
\textsuperscript{2}EDHEC Business School
\textsuperscript{3}University of Lausanne, Faculty of Business and Economics (HEC)
\textsuperscript{4}Swiss Finance Institute
\textsuperscript{5}CEPR
\textsuperscript{6}CIRANO

December 14, 2016

*Financial support from the Swiss Finance Institute is gratefully acknowledged. We are thankful to Audrey Laporte, and seminar participants at the European Workshop on Econometrics, and Health Economics for very useful comments and discussions.
Abstract

Health declines, and mortality risk increases rapidly near the end of life. Curative care stagnates, while comfort care increases, accelerating the fall in wealth. Standard explanations emphasize inevitable health declines associated with aging. We propose a closing down the shop alternative where agents’ decisions affect their health, and the timing of death. Despite strictly preferring to live, agents optimally deplete their health and wealth statuses towards levels associated with high death risk and indifference between life and death. A structural estimation of the closed-form decisions identifies and tests conditions for these strategies to be optimal, and confirm their economic relevance.

JEL classification: D91, D14, I12

Keywords: End of life; Life cycle; Dis-savings; Endogenous mortality risk; Unmet medical needs; Right to refuse treatment.
1 Introduction

Health falls rapidly as we approach the last period of life. Because how healthy we are is a significant predictor of future major health onsets, exposure to death risk also increases. Moreover, health spending augments, and changes in composition. Whereas curative expenses (e.g. doctor visits, hospital stays, drugs, ...) tend to stagnate, nursing homes, and other long-term care (LTC) spending increase sharply. LTC expenditures are more income- and wealth-elastic than curative care, and can be associated with comfort care consumption. Furthermore, LTC expenses are not covered by Medicare; out-of-pocket expenses thus increase sharply towards the end of life, leading to a rapid drain in financial resources.

The standard explanations of these joint end-of-life health and wealth dynamics emphasize ineluctable exogenous health declines that are driven by the aging process. Mortality risks mechanically increase as a result, whereas comfort care expenses are mainly accompanying, but not reverting the biological decline in status. Given an expected remaining life horizon, the wealth management objectives simplify to insuring sufficient resources to reach the end (and eventually leave bequests).

We propose a different perspective that abstracts from the inevitability in end-of-life health and wealth processes. This alternative relies on four modeling hypotheses regarding individual decisions. More precisely, we assume that agents’ choices can affect their health status, through which they can also alter their exposure to mortality. We further assume that individuals prefer life over death, and that they make dynamically-consistent choices. Put differently, the agents’ decisions are coherent with a remaining life horizon, and this horizon is endogenously determined through their decisions. Under these hypotheses, we ask whether the observed dynamics can be rationalized as an optimal relinquishment strategy whereby agents near the end of life choose to close down the shop. Under reasonable, and empirically verifiable conditions, closing down strategies involve

---

1See Banks et al. (2015, Fig. 5, p. 12), Heiss (2011, Fig. 2, p. 124), Smith (2007, Fig. 1, p. 740), or Case and Deaton (2005, Fig. 6.1, p. 186) for evidence.
3See Benjamins et al. (2004); Heiss (2011); Smith (2007); Hurd et al. (2001); Hurd (2002b) for evidence and discussion. See also Arias (2014, Tab. B, p. 4) for Life Tables.
4De Nardi et al. (2015b, Fig. 3, p. 22).
5See De Nardi et al. (2015b); Tsai (2015); Marshall et al. (2010) for evidence and discussion.
6De Nardi et al. (2015a,b); Marshall et al. (2010); French et al. (2006); Palumbo (1999)
7See De Nardi et al. (2015a, 2009), or French and Jones (2011) for examples.
selecting a depletion of the health stock, which can eventually be accelerated towards the end, and which will invariably increase the risk of dying. Moreover, wealth is also selected to fall in response to the shorter life horizon, thereby reducing disposable resources for health spending. As they approach the end of life, dynamic consistency entails that agents gradually become indifferent between life and death.

The contributions of this paper are twofold. We first build upon a rich theoretical framework (Hugonnier et al., 2013) to define the conditions under which closing down the shop could take place. This life cycle model features demand for health in the spirit of Grossman (1972), combined with diminishing returns to health spending, exogenous morbidity, and endogenous mortality exposure. Importantly, its recourse to non-expected utility guarantees preference for life over death. The main theoretical novelty of our approach is to prove the optimality of the joint health and wealth depletion processes near the end of life without recourse to exogenous aging processes. Despite preferring to live, our agents optimally close down the shop; they simultaneously act in a manner that results in a short terminal horizon, and they select a depletion strategy that is consistent with this horizon. It is this simultaneous feedback between decisions and horizon that makes the solution of this model particularly challenging. To our knowledge, this is the first attempt to rationalize end-of-life health and wealth dynamics, rather than model them as ex-post responses to an irreversible sequence of exogenous health and/or wealth declines. Note that this remarkable result does not preclude biological aging explanations; our main findings would be reinforced by exogenous deterioration of health conditions driven through the aging process.

Our second contribution is to assess whether closing down strategies are empirically relevant. Towards that purpose, we innovate by providing a structural econometric characterization of the health and wealth loci where these strategies are to be expected. This allows us to test conditions, and precisely pinpoint thresholds under which closing down does, or does not take place. Using the observed joint distribution of health and wealth for elders (75.3 years old on average), and for self-reported health levels ranging between Poor, Fair, Good, Very Good, and Excellent, our results indicate that, regardless of wealth, all agents with at least Fair health can optimally select to close down the shop.

Exceptions with endogenous mortality include Pelgrin and St-Amour (2016); Hugonnier et al. (2013); Blau and Gilleskie (2008); Hall and Jones (2007). However, none of these papers focus on end-of-life joint dynamics for health and wealth.
This paper also indirectly contributes to policy debates regarding exploding end-of-life health expenses. It reinforces arguments against aggressive therapy, and in favor of the right to refuse treatment. From this perspective, at least part of the spending in curative care towards the end of life may not accord with what agents actually want. Moreover, we show that although it is feasible to reduce the incidence of closing-down, the normative rationale for doing so is unclear. Indeed closing down strategy obtains optimally in a complete markets setup, with horizon, and poverty levels endogenously determined. From a pure normative perspective, standard arguments for intervening such as myopic decisions, market failure, or redistribution cannot be made. Finally, we show that the main conclusions can also be extended to cases of terminal incurable illnesses such as certain forms of cancer.

The rest of this paper proceeds as follows. We summarize the theoretical model in Section 2. The depletion and accelerating regions are defined, and formally characterized in Section 3. The empirical evaluation is performed in Section 4, with main results outlined in Section 5, and concluding remarks in Section 6.

2 Theoretical framework

Our analysis builds upon on the theoretical framework developed in Hugonnier et al. (2013) which we extend to analyze the joint dynamics of health and wealth. The main features of this model and its approximate solution are briefly reproduced here for completeness.

2.1 Economic environment

2.1.1 Health dynamics

The agent’s health level is denoted by $H_t$ and evolves according to a stochastic version of the Grossman (1972) demand-for-health model:

$$
\mathrm{d}H_t = (\frac{I_t}{H_t} - \delta) H_t \mathrm{d}t - \phi H_t \mathrm{d}Q_{st}, \quad H_0 > 0.
$$

The patient’s right to refuse treatment is protected under both common law, and the US constitution (Legal Advisors Committee of Concerns for Dying, 1983), and recognized by the AMA (American Medical Association, 2016). Unmet medical needs is found to be prevalent in 23.4% of the population below poverty line (National Center for Health Statistics, 2012, Tab. 79, pp. 272-75.). See also Ayanian et al. (2000); Park et al. (2016) for evidence.
In this equation $I_t \geq 0$ represents the agent’s flow rate of health spending,\textsuperscript{10} $dQ_{st}$ is the increment of a Poisson process with intensity $\lambda_{s0}$ that captures the arrival of health shocks,\textsuperscript{11} and $(\delta, \alpha, \phi) \in (0, 1)^3$ are constants that represent the decay rate of health in the absence of shocks, the severity of health adjustment costs and the fraction of the agent’s health that is lost upon the occurrence of a shock.

The agent’s health level endogenously determines both the instantaneous likelihood of the agent’s death and his income. Specifically, we assume that the agent’s flow rate of income is given by

$$Y(H_t) = y_0 + \beta H_t$$

(2)

where $(y_0, \beta) \in \mathbb{R}_+^2$ are constants that capture health-independent elements—such as Social Security revenue—and the sensitivity of the agent’s income to his health status.\textsuperscript{12}

We also assume that the agent’s time of death $T_m$ is the first jump time of a Poisson process $Q_{mt}$ whose arrival intensity is given by

$$\lim_{h \to 0} \left( \frac{1}{h} \right) P_t \left[ t < T_m \leq t + h \right] = \lambda_{m0} + \lambda_{m1} H^{-\xi_m}_t \equiv \lambda_m(H_t^-)$$

(3)

for some constants $(\xi_m, \lambda_{m0}, \lambda_{m1}) \in \mathbb{R}_+^3$ where $H_t^- = \lim_{s \uparrow t} H_s$. In this expression the first term represents the agent’s endowed exposure to mortality risk while the second term captures the fact that the agent can influence the distribution of his lifetime by investing in his health. The fact that both $\xi_m$ and $\lambda_{m1}$ are nonnegative implies that a healthier agent can naturally expect to live longer.

### 2.1.2 Investment opportunity set

The agent can continuously invest in three assets: A risk-free asset with constant rate of return $r$, and two risky assets. The first risky asset proxies for the stock market. Its

\textsuperscript{10}The restriction to positive health investment rates is standard and simply reflects the fact that the agent cannot monetize his own health.

\textsuperscript{11}Hugonnier \textit{et al.} (2013) consider a more general model in which the arrival intensity of health shocks is a decreasing function of the agent’s health level. We focus on the case of a constant arrival intensity to facilitate the presentation but our results can be extended to cover this more general case.

\textsuperscript{12}Old-age male participation in the labor market has increased from 26% in 1995, to 35% in 2014, 60% of which involves full time work (Bosworth \textit{et al.}, 2016, Figs. II.1, and 2, pp. 7, and 9). See also Bureau of Labor Statistics (2008); Toossi (2015) for further evidence of increased old age participation in the labor force.
prices is denoted by $S_t$ and evolves according to

$$\frac{dS_t}{S_t} = rt + \sigma_S (dZ_t + \theta_t dt)$$

where $dZ_t$ is the increment of a Brownian motion that captures market risk, and $(\sigma_S, \theta) \in \mathbb{R}^2_+$ are constants which represent, respectively, the volatility of market returns and the instantaneous remuneration that investors earn for exposure to market risk.

The second risky asset that the agent can invest in is an actuarially fair health insurance contract that pays one unit of the numéraire upon the occurrence of a health shock. The instantaneous return that the agent earns by investing the amount $X_t \geq 0$ in this asset over the time interval $(t, t + dt]$ is given by

$$X_t (dQ_{st} - \lambda_s dt)$$

(4)

where the first term captures the payment that the agent receives from the insurer upon the occurrence of a health shock and the second term captures the instantaneous insurance premium that he pays to the insurer.

Denote by $\Pi_t$, $X_t$ and $C_t$ the predictable processes that track the amount that the agent invests in the stock market, the amount he invests in the insurance and the amount he consumes per unit of time. With this notation we have that the dynamic budget constraint that governs the evolution of the agent’s wealth is

$$dW_t = (rW_t + Y(H_t) - C_t - I_t)dt + \Pi_t \sigma_S (dZ_t + \theta dt) + X_t (dQ_{st} - \lambda_s dt).$$

(5)

Investment in the riskless asset and the stock market is unconstrained, but we naturally assume that the agent can neither consume negative amounts nor sell insurance by imposing a nonnegativity constraint on both $C_t$ and $X_t$.

2.1.3 Preferences

To close the model it remains to specify the agent’s preferences. Following Hugonnier et al. (2013) we assume that the continuation utility $U_t = U_t(C)$ to an alive agent of a
lifetime consumption schedule $C$ solves a recursive integral equation of the form

$$
U_t = E_t \int_t^{T_m} \left( f(C_s, U_s) - \frac{\gamma \sigma_s^2}{2U_s} - \sum_{k=m}^{s} F_k(U_s, H_s, \Delta_k U_s) \right) ds
$$

(6)

where $\gamma$ is a strictly positive constant that measures the agent’s local risk aversion to financial market shocks, $\sigma_t = \sigma_t(U) = d(U, Z)/dt$ measures the sensitivity of the continuation utility process to these shocks, and

$$
\Delta_k U_t = 1_{\{dQ_{kt} \neq 0\}} (U_t - U_{t-})
$$

represents the predictable jump in continuation utility triggered by the occurrence of a health shock ($k = s$) or the agent’s death ($k = m$). In the above equation

$$
f(C, U) = \frac{\rho U}{1 - 1/\varepsilon} \left( ((C - a)/U)^{1-1/\varepsilon} - 1 \right)
$$

(7)

is the Kreps-Porteus aggregator with elasticity of intertemporal substitution $\varepsilon > 0$, time preference rate $\rho > 0$, and subsistence consumption level $a \geq 0$; and the penalty terms below the sum are given by

$$
F_k(U, H, \Delta U) = \lambda_k(H) \left[ \frac{\Delta U}{U} + \frac{1 - (1 + \Delta U/U)^{1-\gamma_k}}{1 - \gamma_k} \right] U
$$

(8)

for some constants $\gamma_s > 0$ and $\gamma_m \in [0, 1)$. This recursive preference specification allows to disentangle the agent’s behavior toward intertemporal substitution from his attitude towards risk but, as explained in Hugonnier et al. (2013), it goes one step further than Duffie and Epstein (1992) by allowing to discriminate between the various sources of risk present in the model. Our specification thus implies that the agent has constant relative risk aversion $\gamma > 0$ towards financial market risk, $\gamma_s \geq 0$ towards health risk, and $\gamma_m \in [0, 1)$ toward mortality risk. Importantly, the restriction that $\gamma_m < 1$ guarantees that, irrespective of his attitude towards the other sources of risk, the agent prefers life over death.
2.1.4 The decision problem

The agent’s decision problem consists in choosing a portfolio, consumption, health insurance and health investment strategy to maximize his lifetime utility. The indirect utility associated with this problem is defined by

\[ V(W_t, H_t) = \sup_{(C,H,X,I)} U_t(C) \]

subject to the dynamics of the health process (1) and the budget constraint (5) where \( U_t(C) \) is the continuation utility process associated with the lifetime consumption and health investment plan \((C,I)\) through (6).

In the absence of bequests, the continuation utility process defined by (6) becomes zero at death.\(^{13}\) As a result, we have \( \Delta_m U_t = -U_t \) and it follows that the penalty associated with mortality risk satisfies

\[ \frac{F_m(U_s, H_s, \Delta_m U_s)}{\lambda_m(H_s) U_s} = \frac{\gamma_m}{1 - \gamma_m} \equiv \Phi \in [1, \infty). \]

Using this observation and integrating over the conditional distribution of the agent’s time of death, Hugonnier et al. (2013) show that the agent’s decision problem, which features incomplete markets and an endogenous random horizon, can be conveniently recast into an equivalent infinite horizon problem with endogenous discounting and complete markets. Specifically, they show that

\[ V(W_t, H_t) = \sup_{(C,H,X,I)} \bar{U}_t(C) \]

where the modified continuation utility process \( \bar{U}_t = \bar{U}_t(C) \) solves the infinite horizon recursive integral equation given by

\[ \bar{U}_t = E_t \int_t^\infty e^{-\int_t^s \lambda_m(H_k)(1+\Phi)dk} \left( f(C_s, U_s) - \frac{\gamma \sigma_s(U)}{2U_s} - F_s(U_s, H_s, \Delta_s U_s) \right) ds. \quad (9) \]

\(^{13}\)This assumption is imposed for tractability and can be justified by noting that while bequest motives are potentially relevant in an endogenous mortality setting, panel data evidence suggests that their role in explaining the behavior of retired agents is debatable. In particular, Hurd (2002a) finds no clear evidence of a bequest motive behind savings decisions and Hurd (1987) finds no differences in the saving behavior of the elderly who have children compared to those who don’t.
This formulation brings to light the channels through which the agent’s health status enters his decision problem. First, health can be interpreted as a durable good that generates service flows through the income rate $Y(H_t)$. Second, health determines the instantaneous probability of morbidity shocks and the rate $\lambda_m(H_t)(1 + \Phi)$ at which the agent discounts future consumption and continuation utilities.

**Remark 1 (Health dependent preferences)** One might also reasonably object that old agents are likely to be retired and thus do earn labor income. However, this objection is inconsequential for our purposes since the agent’s decision problem is iso-morphic to one with health-dependent utility, and constant base income. This results follows by effecting the change of variable $C_t = C_t + \beta H_t$ throughout the above equations, see (Hugonnier et al., 2013, Remark 3) for details.

### 2.2 Optimal dynamic policies

The fact that the discount rate in (9) is endogenous implies that the agent’s decision problem does not admit a closed-form solution. To circumvent this difficulty Hugonnier et al. (2013) rely on a two step procedure. First, they show that in the exogenous mortality case where $\lambda_m = 0$ the agent’s decision problem admits a closed form solution. Second, they use an asymptotic expansion of the solution to the dynamic programming equation around the point $\lambda_m = 0$ to compute the first order effect of endogenous mortality on the optimal policy. Adapting their results to our setting allows to derive an following approximation to the optimal policy.

**Theorem 1 (Optimal policy functions)** Assume that conditions (39), (40), and (41) of Appendix A.2 hold true, and let

\[
N_0(W, H) = W + BH + (y_0 - a) / r
\]  

(10)
where \( B > 0 \) is the smallest solution to (45). Up to a first order approximation the optimal policy functions are given by:

\[
X^*(W, H) = \phi BH + X_1 H^{-\xi_m} N_0(W, H) \quad (11)
\]

\[
C^*(W, H) = a + A N_0(W, H) + C_1 H^{-\xi_m} N_0(W, H) \quad (12)
\]

\[
\Pi^*(W, H) = \left( \frac{\theta}{\gamma \sigma_S} \right) N_0(W, H) \quad (13)
\]

and

\[
I^*(W, H) = KBH + I_1 H^{-\xi_m} N_0(W, H) \quad (14)
\]

where the nonnegative constants \( B, X_1, A, C_1, K, \) and \( I_1 \) are defined in Appendix A.2.

The optimal rules in Theorem 1 are all increasing in net total wealth \( N_0(W, H) \). The latter in (10) includes financial wealth \( W \) plus the present value of the human capital \( BH \), for which \( B \) is the marginal-\( Q \) of health solving equation (45). This value can be interpreted as the capitalized value of the health-dependent capacity to generate income \( \beta H \) in (2), where \( B \) is an increasing function of \( \beta \). In parallel, \((y_0 - a)/r\) is the net present value (NPV) of base income \( y_0 \) minus subsistence consumption \( a \).

The optimal rules are defined only over an admissible state space, i.e. the set of wealth and health levels such that net total wealth \( N_0(W, H) \) is nonnegative in (10). Indeed, observe from optimal consumption (12) that admissibility is required to ensure that consumption \( C_t \) is above subsistence \( a \). Moreover, as shown in Hugonnier et al. (2013), the homogeneity of the Kreps-Porteus aggregator \( f \) in (7), and of the penalty function \( F_k \) in (8) guarantees that the continuation utility (6) is also homogeneous. It follows that excess consumption and welfare are measured in the same units. Positive excess consumption \( C_t - a > 0 \) entails positive continuation utility \( V(W_t, H_t) > 0 \), and therefore strict preference of life over death, i.e. \( V(W_t, H_t) \equiv 0 \). In order to ensure that resources are sufficient to cover subsistence consumption, and that life is preferable, positive net total wealth in equation (10) can be relied upon to define the admissible region \( \mathcal{A} \):

\[
\mathcal{A} = \{ (W, H) \in \mathbb{R} \times \mathbb{R}_+ : W \geq x(H) \equiv -(y_0 - a)/r - BH \}.
\]
Note finally that we might expect base income $y_0$ to be insufficient to cover subsistence consumption $a$, therefore requiring strictly positive financial resources. Moreover, observe from optimal risky holdings (13) that the risky portfolio shares out of total wealth can be written as:

$$\Pi'(W, H)/W = (\theta/(\gamma \sigma_S)) (1 - x(H)/W).$$

As is well-known, portfolio shares are increasing in the financial wealth level $W$ (e.g. Wachter and Yogo, 2010), which therefore requires that $x(H)$ be nonnegative, and consequently, that:

$$(y_0 - a)/r < 0.$$ (16)

Restriction (16) is also tested and confirmed empirically in Section 5, and will henceforth be imposed.

The first term in optimal rule (14) is the order-0 investment, and is proportional to health capital’s economic value $BH$. The second term captures the incremental demand for health that arises from its death risk hedging capacity; that demand is increasing in the endogenous component $\lambda_{mH} H^{-\xi_m}$ of the death intensity (3). Hence the demand for health is larger when better health reduces the exposure to death risk. As will be seen next, the non-monotonic effects of $H$ on $I(W, H)$ induced by the demand for death risk hedging will play a key role in the complex nonlinear dynamics for health and wealth.

3 Optimal health and wealth dynamics

We assume from now on that the agent follows the approximate optimal rules prescribed by Theorem 1. Consequently, his health and wealth evolve according to the dynamical system formed by (1) and (5) evaluated at (11)–(14). Due to the presence of Brownian financial shocks and Poisson health shock, this dynamical system is stochastic and thus cannot be directly analyzed using standard tools such as phase portraits. To circumvent
this difficulty we focus on the instantaneous expected changes in health and wealth that are implied by the optimal rules.\textsuperscript{14}

We start by defining, and characterizing health depletion, as well as speed of depreciation in Section 3.1. We then analyze wealth depletion, as well as implications for closing-down strategies in Section 3.2. Policy instruments aimed at reducing the prevalence of these depletion dynamics are presented in Section 3.3. Finally, we analyze the implications in the case of terminal incurable illness in Section 3.4.

### 3.1 Optimal health depletion

The expected local change in health capital is given by:

\[
E_{t-}[dH] = \left[ I^h(W, H)^{\alpha} - \tilde{\delta} \right] H dt,
\]

where we denote the conditional expectation, given filtration \( F_{t-} \), as \( E_{t-}[\cdot] = E[\cdot \mid F_{t-}] \), and let \( \tilde{\delta} = \delta + \phi\lambda_{s0} \) be the sickness-adjusted expected depreciation rate. The investment-to-health ratio evaluated at the optimal investment in (14) is given by:

\[
I^{h}(W, H) = \frac{I_{1}^{*}(W, H)}{H} = BK + I_{1}H^{-\xi=-1}N_{0}(W, H).
\]

Since our main focus concerns end-of-life decumulation, we can define the admissible region of the state space \( D_{H} \) where health depletion is expected:

\[
D_{H} = \{(W, H) \in A : E_{t-}[dH] < 0\}.
\]

Interestingly, it is also possible to analyze how fast the health capital is allowed to deplete. To do so, we can define an acceleration subset in the health depletion region \( \mathcal{AC} \subseteq D_{H} \) whereby the investment-to-health ratio is an increasing function of health:

\[
\mathcal{AC} = \{(W, H) \in D_{H} : I^{h}_{H}(W, H) > 0\}.
\]

Hence for agents with \((W, H) \in \mathcal{AC}\), a positive health gradient of the investment-to-capital ratio (18) entails that health depletion is followed by more important cuts in

\textsuperscript{14}See also Laporte and Ferguson (2007) for an analysis of expected local changes of the Grossman (1972) model with Poisson shocks.
health investment, leading to declines in $I^h(W, H)$, and further depletion of the health capital in (17).

The following result characterizes both the health depletion, and acceleration regions of the state space.

**Theorem 2 (health depletion)** Assume that the agent follows the optimal rules in Theorem 1, and that:

$$\beta < \tilde{\delta}^{1/\alpha}. \quad (19)$$

Then,

1. the health depletion zone is given by:

$$\mathcal{D}_H = \{(W, H) \in A : W < y(H)\}, \quad (20)$$

where the health depletion locus is

$$y(H) = x(H) + DH^{1+\xi_m},$$

$$D = \mathcal{T}^{-1}_1 \left[ \tilde{\delta}^{1/\alpha} - BK \right] > 0,$$

2. there exists a threshold $\bar{H}_3$ given by:

$$\bar{H}_3 = \left( \frac{B}{D(1 + \xi_m)} \right)^{1/\xi_m} > 0,$$

such that the accelerating region is given by:

$$\mathcal{AC} = \begin{cases} 
\mathcal{D}_H, & \text{if } H < \bar{H}_3; \\
\{(W, H) \in \mathcal{D}_H : W < z(H)\}, & \text{otherwise},
\end{cases} \quad (23)$$

where the acceleration locus is

$$z(H) = x(H) + \frac{BH}{1 + \xi_m}.$$
Condition (19) states that expected health depreciation $\tilde{\delta}$ is high relative to the health-dependent income contribution $\beta$, and is appropriate for end-of-life characterization. Indeed, a high depreciation in the absence of investment ($\delta$), or conditional upon sickness ($\phi$), a high likelihood of morbidity shocks ($\lambda_{s0}$), as well as a low variable component in income ($\beta$) are all to be expected for elders. It is also more stringent than, and therefore induces the required transversality condition (39) in Appendix A.1.

Importantly, combining condition (19) with transversality restriction (46) in Appendix A.2 reveals that

$$BK < \tilde{\delta}^{1/\alpha}.$$ \hfill (25)

As shown in Appendix B.2.1, the expression $(\tilde{\delta}^{1/\alpha} - BK) > 0$ captures the order-0 expected depletion, i.e. in the absence of endogenous mortality. When the latter is reintroduced, it was shown earlier that optimal investment in (14) is larger, reflecting the positive incremental demand for death risk hedging provided by health capital. If condition (25) is violated, then health grows in expectation absent mortality control value; positive growth is even larger when endogenous mortality is re-introduced and no relevant health depletion region exists. Moreover, this restriction ensures that the constant $D > 0$ in equation (21), such that $y(H) > x(H)$, and the admissible health depletion subset is therefore non-empty.

The health dynamics characterized by Theorem 2 can be analyzed through the phase diagram in $(W, H)$ space in Figure 1. First, the admissible region $A$ is bounded below by the $x(H)$ locus (15) in red, with complementary non-admissible area $N_A$ in shaded red region. The $W-$intercept of $x(H)$ is given by the NPV of base income deficit $-(y_0 - a)/r$ which is positive under restriction (16), whereas the $H-$intercept is given by $\bar{H}_1 = -(y_0 - a)/(rB) > 0$.

Second, equation (20) states that the health depletion region $D_H$ is the shaded green area located below the green $y(H)$ locus (21). Both $x(H), y(H)$ loci intersect at the same $-(y_0 - a)/r$ intercept. Under condition (25), we show in Appendix B.2.1 that the $y(H)$ locus is U-shaped, and attains a unique minimum at $\bar{H}_3$ given by (22). The reasons for the non-monotonicity stem from the non-monotone effects of $H$ on $I^k(W, H)$. Indeed, the investment-to-health ratio in (18) is monotone increasing in wealth, but not in health due to the opposing forces of net total wealth, and mortality effects. On the one hand,
an increase in $H$ raises net total wealth $N_0(W, H)$, and therefore raises $I^h$. Consequently, constant (and zero) expected growth is obtained by reducing $W$. On the other hand, an increase in $H$ also reduces endogenous mortality risk $I_1^{-\xi_m-1}$, and therefore also reduces $I^h$. Therefore, constant zero growth requires increasing $W$. The analysis of the $y(H)$ locus in (21) thus reveals that the net total wealth effect is dominant at low health ($H < \bar{H}_3$), whereas the mortality risk effect dominates for healthier agents ($H > \bar{H}_3$).

Third, the accelerating locus $z(H)$ in (24) is plotted as the blue line in Figure 1; the accelerating region is the dashed blue and green subset of $D_H$. Appendix B.2.2 shows that this locus intersects the $x(H), y(H)$ loci at the same $-(y_0 - \alpha)/\rho$ intercept, and that it intersects the $H-$axis at $\bar{H}_2 = \bar{H}_1(1 + \xi_m)/\xi_m > \bar{H}_1$; consequently, the admissible accelerating region $x(H) < W < z(H)$ is non-empty. Moreover, it also intersects the health depletion locus $y(H)$ at its unique minimal value $\bar{H}_3$ in (22). Consequently, all agents with $H < \bar{H}_3$ in the depletion region are also in the accelerating subset.

The local expected dynamics of health are represented by the horizontal directional arrows in Figure 1. First, only agents who are sufficiently rich (i.e. $W > y(H)$) can expect a growth in health; all others are located in the $D_H$ region in which the health stock is expected to fall. In particular, there exists a threshold wealth level $\bar{W}_3 = y(\bar{H}_3)$ below which all agents, regardless of their health status, expect a health decline. Second, the sufficiently rich and healthy ($W > z(H)$) agents in the health depletion region $D_H$ optimally slow down (but do not reverse) the depreciation of their health capital (i.e. $I^h_H < 0$). However, for $W < z(H)$, the health depletion accelerates (i.e. $I^h_H > 0$, illustrated by the thick directional vector) as falling health is accompanied by further cuts in the investment-to-health ratio. The health dynamics thus crucially depend on the wealth levels and dynamics, an issue we now address.

3.2 Optimal wealth depletion

Since the expected net return is zero on actuarially fair insurance contracts (4), the expected changes in wealth is:

$$E_{t-}[dW] = [rW + Y(H) - C^*(W, H) - I^*(W, H) + \Pi^*(W, H)\sigma_s\theta] dt.$$  

(26)
Notes: Non-admissible set $NA$: shaded red area under red $x(H)$ line, admissible $A$ is area above $x(H)$. Health depletion set $D_H$: shaded green area under green $y(H)$ green curve. Acceleration set $AC$: hatched green area under blue $z(h)$ curve. Wealth depletion set $D_W$: area above $w(H)$ black curve.

In parallel with the earlier analysis, the wealth depletion $D_W$ region of the admissible state space can then be written as:

$$D_W = \{(W, H) \in A : E_t[-dW] < 0\}.$$

The following result identifies realistic sufficient conditions for characterizing wealth depletion.

**Theorem 3 (wealth depletion)** Assume that the conditions in Theorem 2 are verified, and that preferences are sufficiently elastic with respect to inter-temporal substitution, i.e. $\varepsilon \geq 1$, and that the following condition holds:

$$\frac{\theta^2}{\gamma} + r < A.$$  

(27)
Then,

1. the wealth depletion zone is given by:

\[ \mathcal{D}_W = \{(W,H) \in \mathcal{A} : W > w(H)\} \tag{28} \]

where the wealth depletion locus is

\[ w(H) = \frac{x(H)[l(H) + r]}{l(H)} + \frac{k(H)}{l(H)}, \tag{29} \]

\[ l(H) = A - \frac{\theta^2}{\gamma} - r + (I_1 + C_1) H^{-\zeta_m} \tag{30} \]

\[ k(H) = y_0 - a + H(\beta - KB), \tag{31} \]

2. the joint health and wealth depletion region \( \mathcal{D} = (\mathcal{D}_W \cap \mathcal{D}_H) \) is everywhere non-empty, i.e.:

\[ x(H) \leq w(H) \leq y(H), \quad \forall H. \]

As will be seen shortly, the assumption of a sufficiently high elasticity of inter-temporal substitution is reasonable for our data, and also verified empirically. The condition (27) refers to high consumption patterns and is economically plausible and relevant for end-of-life analysis. To see this, observe from (42) that the former can be rewritten as:

\[ (1 + \varepsilon) \frac{\theta^2}{2\gamma} < \varepsilon(\rho - r) + (\varepsilon - 1) \frac{\lambda_{m0}}{1 - \gamma_m}. \]

Since \( \gamma_m \in [0, 1) \), and if the elasticity of intertemporal substitution \( \varepsilon \geq 1 \), the high marginal propensity to consume \( A \) in (27) obtains whenever the agent is impatient, i.e. \( \rho \) is high, and/or the unconditional risk of dying \( \lambda_{m0} \) is high, and/or the aversion to death risk \( \gamma_m \) is high.

The wealth depletion locus \( w(H) \) in (29) is represented as the black curve in Figure 1. Appendix B.3 establishes that this locus has the same \( H \)– intercept \( -(y_0 - a)/r \), and must lie everywhere between the \( x(H) \), and the \( y(H) \) loci. Equation (28) states that the wealth depletion \( \mathcal{D}_W \) is the area above the black curve \( w(H) \), with corresponding wealth dynamics captured by the vertical directional arrows. Most agents thus expect
their wealth to fall, except at very low wealth levels in the $\mathcal{AC}$ region where rapidly receding health expenses $I(W, H)$ in (26) allow for expected increases in wealth. Since $w(H)$ is located between the admissible, and the health depletion loci, the joint depletion region ($\mathcal{D}_W \cap \mathcal{D}_H$) is everywhere non-empty. Put differently, for every $H$, there exists an admissible range of $W$ for which agents jointly expect health and wealth to fall.

These joint end-of-life health and wealth dynamics are thus consistent with a deliberate closing down the shop strategy. Sufficiently rich and healthy agents ($W > y(H)$) postpone health declines, e.g. by restoring levels following sickness. However, falling wealth is optimally chosen and agents eventually enter the $\mathcal{D}_H$ region where health depletion is also selected.$^{15}$ Depreciation of the health stock accelerates once falling health and wealth push agents into the $\mathcal{AC}$ region. Our model thus supports threshold effects whereby falling health is initially fought back, before eventually accelerated. Although we do not distinguish between various inputs in health care, such behavior would be consistent with an end-of-life change in composition in health expenses towards more comfort, and less curative care (De Nardi et al., 2015b; Marshall et al., 2010). From the endogenous death intensity (3), falling health is invariably accompanied by an increase in mortality, and a decline towards the admissible locus $x(H)$ characterized by zero net total wealth, and indifference between life and death. Importantly, this optimal relinquishment occurs even when life is strictly preferred. Indeed, as discussed earlier, the non-separable preferences (6) ensure strictly positive continuation utility under life (versus zero under death), under admissible health and wealth statuses. The agents we are considering therefore have no proclivity in favor of premature death.

Remarkably, our closing-down strategies rationalize the fall in health and wealth without recourse to an exogenous aging process. Indeed, allowing for an exogenous aging process would only reinforce our results. As shown in Hugonnier et al. (2013), the model can also be solved in the presence of time variation in the deep parameters. Incorporating aging would involve age-increasing depreciation $\tilde{\delta}_t$, and mortality risk $\lambda_{m0,t}$, as well as age-decreasing ability to work $\beta_t$. Although a complete derivation is beyond the scope of

$^{15}$It is worth noting that the optimal risky asset holdings in (13) are positive when net total wealth, and risk premia are both positive. Moreover, the investment in (14) is monotone increasing in wealth, such that a sufficiently long sequence of high positive returns on financial wealth could be sufficient to pull the agents away from the depletion region $\mathcal{D}_H$. Put differently, falling health, and higher mortality is locally expected, yet is not absolute for agents in the depletion region. We will return to this issue in the simulation exercise discussed below.
this paper, we note that these three time variations all concur to make the closing-down conditions (19), and (27) more easily met.

### 3.3 Reducing the prevalence of closing-down strategies

Assuming that such an objective is warranted (e.g. for public health purposes), the prevalence of closing-down strategies can be reduced through income policies. One instrument that can be used towards that aim is the base income $y_0$ which can be altered through Social Security, or minimal revenue policies. Figure 2 shows that an increase in base income lowers the intercept $-(y_0 - a)/r$. It follows that the four loci are shifted downwards without affecting $\bar{H}_3$, leading to a lower $\bar{W}_3^1 < \bar{W}_3^0$. If we take as given the current health and wealth distribution, admissibility is increased, and the prevalence of the health depletion $D_H$ is reduced.\footnote{Note however that this lower incidence remains theoretical, since the optimal rules in Theorem 1 would adjust to the change in $y_0$ resulting in a different joint distribution of health and wealth in the $(W,H)$ space.}

Whereas the tools for reducing the incidence of closing down are readily identified, the normative arguments in favor of intervention are less clear. Indeed, the traditional rationale of market incompleteness can hardly be invoked since closing down is obtained as an optimal dynamic strategy under a complete markets setting. Moreover, poor agents are subject to faster depreciation of their health capital and higher mortality risk. However, redistribution arguments cannot be invoked to the extent that poverty, and life expectancy are both endogenously determined as an optimal strategy. Consequently, redistributive arguments for intervening can neither be invoked.

### 3.4 Incurable terminal illness

While closing down strategies are optimal when exposure to mortality can be altered by health-related choices, it’s unclear whether they remain so when health improvements no longer mitigate the risk of dying, e.g. in the case of incurable terminal illness, such as certain forms of cancer. Because exposure to death risk no longer responds to treatment, such conditions may be approximated in our model by (i) removing the endogeneity of the mortality intensity, i.e. setting $\lambda_{m1} = 0$, (ii) increasing the fixed intensity of death, i.e. $\lambda_{m0}$ to capture the shorter expected survival, and (iii) increasing the decay rate...
Figure 2: Reducing the prevalence of closing-down.

Notes: Effects of increase in base income $y_0$. Dotted lines: Initial position, solid lines: resulting position. Non-admissible set $\mathcal{N}A$: shaded red area under red $x(H)$ line. Health depletion set $\mathcal{D}_H$: shaded green area under green $y(H)$ green curve. Acceleration set $\mathcal{A}C$: hatched green area under blue $z(H)$ curve. Wealth depletion set $\mathcal{D}_W$: area above $w(H)$ black curve.

$\dot{\bar{\delta}} \equiv \bar{\delta} + \lambda_{a_0}\phi$ to capture the faster expected depreciation of the health stock. As the following result makes clear, terminal illnesses alter the state space segments, but not the ultimate conclusion that closing-down is optimal.

**Theorem 4 (Incurable terminal illness)** Assume that the conditions in Theorem 2 are verified, but that the exposure to mortality risk can no longer be adjusted, i.e. $\lambda_{m1} = 0$. Then:

1. health depletion is expected everywhere in the admissible set:

$$\mathcal{D}_H = \mathcal{A},$$

(32)
2. the accelerating subset is empty:

\[ \mathcal{AC} = \emptyset. \]  \hfill (33)

If, in addition, condition (27) in Theorem 3 is verified, then:

3. the wealth depletion subset is given by

\[ \mathcal{D}_W = \{(W, H) \in \mathcal{A} : W > w(H)\}, \]

where the wealth depletion locus is modified as

\[ w(H) = \frac{x(H)[l + r]}{l} + \frac{k(H)}{l} > x(H), \]  \hfill (34)

\[ l = A - \frac{\theta^2}{\gamma} - r \]

and where \( k(H) \) is given as (31),

4. the joint health and wealth depletion \( \mathcal{D} = (\mathcal{D}_W \cap \mathcal{D}_H) \) is everywhere non-empty.

For reasons discussed earlier, setting the endogenous death intensity parameter \( \lambda_{m1} = 0 \) reverts the model to its order-0 solutions. In particular, it implies that \( \mathcal{I}_1 = 0 \), such that the investment to health ratio (18) is now constant, and expected depletion obtains throughout the state space under condition (19), and induced condition (25). It follows that all agents in the admissible subset expect health to deplete as stated in (32). Second, a constant investment-to-health capital ratio \( I^h \) obviously does not respond to health levels; consequently no accelerating region exists, as stated in (33). Third, the order-zero solution for consumption also implies that \( \mathcal{C}_1, \mathcal{I}_1 = 0 \) in (30). Substituting in (29) reveals that the \( w(H) \) locus (34) is now a straight line which lies above the admissible locus \( x(H) \) under condition (27). As health depletion is the entire admissible set, the joint health and wealth depletion is everywhere non-empty.

Hence, closing down strategies remain optimal in the presence of incurable terminal illness. The health depletion now extends to all health and wealth levels in the admissible region. Accelerating depletion is no longer a feasible strategy as the health decay rate is optimally set at a constant value, whereas wealth depletion remains marginally effected. Note finally that increasing \( \lambda_{m0} \), and \( \tilde{\delta} \) under incurable terminal diseases reinforces the
incidence of closing down since conditions (19), and (27) are both more likely to be met, compared to cases where exposure to death risk can be altered.

4 Empirical evaluation

The closing-down strategy of optimal health and wealth depletion we have identified is arguably more appropriate for agents nearing death, than for younger ones. Indeed, a high sickness-augmented depreciation rate for the health capital, and a low ability to generate labor revenues (condition (19)) both seem legitimate for old agents in the last period of life, yet less so for younger ones. Moreover, a high marginal propensity to consume (condition (27)), as well as a base income deficit relative to subsistence consumption (condition (16)) are suitable for elders nearing end of life. Using a database of relatively old individuals, we next verify empirically whether or not these conditions are valid, and whether the admissible, depletion, and acceleration subsets have economic relevance.

Towards that objective, the structural econometric model that we rely upon to estimate the deep parameters and evaluate the induced parameters that are relevant for the various regions of the state space is based on a subset of the optimal rules in Section 2.2.

4.1 Econometric model

The tri-variate nonlinear structural econometric model that we estimate over a cross-section of agents \( j = 1, 2, \ldots, n \) is the optimal investment (14), and the risky asset holdings (13), to which we append the income equation (2):

\[
I_j = KBH_j + I_1 H_j^{-\delta_m} N_0(W_j, H_j) + u^I_j, \quad (35)
\]
\[
\Pi_j = (\theta/(\gamma s)) N_0(W_j, H_j) + u^\Pi_j, \quad (36)
\]
\[
Y_j = y_0 + \beta H_j + u^Y_j, \quad (37)
\]

where \((u^I_j, u^\Pi_j, u^Y_j)\) are (potentially correlated) error terms. Data limitations discussed below explain why optimal insurance (11), and consumption (12) are omitted from the econometric model. The latter thus assumes that agents are heterogeneous only with respect to their health, and wealth statuses; the deep parameters are considered to be
the same across individuals. This assumption is justifiable since we are considering a relatively homogeneous subset of old individuals, thereby ruling out potent cohort effects.

A subset of the technological, distributional, and preference parameters are estimated using the joint system (35), (36) and (37), imposing the regularity conditions outlined in Appendix A.1, and under the theoretical restrictions governing \((K, \mathcal{L}_1, B)\) that are outlined in Appendix A.2. The identification of the deep parameters is complicated by the significant non-linearities that are involved. Consequently, not all the parameters can be estimated, and we calibrate a subset. Certain calibrated parameters are set at standard values from the literature. For others however, scant information is available, and we rely on thorough robustness analysis, especially with respect to \(\gamma_m\), and \(\phi\).

The estimation approach is an iterative two-step procedure. In a first step, the convexity parameter \(\xi_m\) is fixed and a maximum likelihood approach is conducted on the remaining structural parameters. In a second step, the other structural parameters are fixed and the likelihood function is maximized with respect to \(\xi_m\). The procedure is iterated until a fixed point is reached for both the structural parameters and the convexity parameters.

The likelihood function is written by assuming that there exist some cross-correlation between the three equations, i.e. \(\text{Cov}(u_I^j, u_{\Pi}^j, u_Y^j) \neq 0\). For the first two equations, the cross-correlation can be justified by the fact that we use an approximation of the exact solution (see Hugonnier et al., 2013, for details). Moreover, our benchmark case assumes that the three dependent variables are continuous. However, the observed risky holdings \(\Pi_j\) contain a significant share of zero observations. For that reason, we also experiment a mixture model specification in which the asset holdings variable is censored (Tobit) and the other two dependent variables (investment and income) are continuous, resulting in qualitatively similar results.

4.2 Data

The data base used for estimation is the 2002 wave of the Health and Retirement Study (HRS, Rand data files). This data set is the last HRS wave with detailed information on

\[\text{17}\] These alternative estimates, which are available upon request, are reasonably robust, with main interpretations qualitatively unaffected.

\[\text{18}\] Note however that our structural model neither rules out zero holdings, nor does predict a Tobit-based specification for the portfolio equation.
total health spending; subsequent waves only report out-of-pocket expenses. Under OOP ceilings, total health expenses $I$ are not uniquely identified for insured agents. Also, even though the HRS contains individuals aged 51 and over, we restrict our analysis to elders (i.e. agents aged 65 and more), with positive financial wealth (9,817 observations, mean age 75.3). In doing so, we avoid endogenizing the insurance choice $X_t$ in (4) which can be considered as exogenous under near-universal Medicare coverage. Unfortunately, this data set does not include a consumption variable, so that we omit equation (12) from the econometric model.

We construct financial wealth $W_j$ as the sum of safe assets (checking and saving accounts, money market funds, CD’s, government savings bonds and T-bills), bonds (corporate, municipal and foreign bonds and bond funds), retirement accounts (IRAs and Keoghs), and risky assets (stock and equity mutual funds) $\Pi_j$. Health status $H_j$ is evaluated using the self-reported general health status, where we express the polytomous self-reported health variable in real values with increments of 0.75 corresponding to: 0.5 (poor), 1.25 (fair), 2.00 (good), 2.75 (very good), and 3.50 (excellent).\(^{19}\)

Health investments $I_j$ are obtained as the sum of medical expenditures (doctor visits, outpatient surgery, hospital and nursing home, home health care, prescription drugs and special facilities), and out-of-pocket medical expenses (uninsured cost over the two previous years). Finally, we resort to wage/salary income $Y_j$, to which we add any Social Security revenues. The estimates presented below are obtained for a scaling of $\$M$ applied to all nominal variables ($I_j, W_j, \Pi_j, Y_j$).

Table 1 reports the median values for wealth, investment and risky asset holdings, for wealth quintiles, and self-reported health. Overall, these statistics confirm earlier findings. A first observation concerns the relative insensitivity of financial wealth to the health status.\(^{20}\) Second, we find that health investment increases moderately in wealth, and falls sharply in health.\(^{21}\) Conversely, risky holdings increase sharply in wealth, and are also higher for healthier agents.\(^{22}\)

\(^{19}\)Self-reported health has been shown to be a valid predictor of the objective health status (Benítez-Silva and Ni, 2008; Crossley and Kennedy, 2002; Hurd and McGarry, 1995).

\(^{20}\)See Hugonnier et al. (2013); Michaud and van Soest (2008); Meer et al. (2003); Adams et al. (2003) for additional evidence.

\(^{21}\)Similar findings with respect to wealth (e.g. Hugonnier et al., 2013; Meer et al., 2003; DiMatteo, 2003; Gilleskie and Mroz, 2004; Acemoglu et al., 2013) and health (e.g. Hugonnier et al., 2013; Smith, 1999; Gilleskie and Mroz, 2004; Yogo, 2009) have been discussed elsewhere.

\(^{22}\)Similar positive effects of wealth on risky holdings have been identified in the literature (e.g. Hugonnier et al., 2013; Wachter and Yogo, 2010; Guiso et al., 1996; Carroll, 2002) whereas positive
5 Results

Table 2 reports the calibrated, and estimated deep parameters (panels a–d), the induced parameters that are relevant for the various subsets (panel e), as well as the hypothesis testing for the assumptions relevant to Theorems 2, and 3 (panel f). The standard errors indicate that all the estimates are precisely estimated, and are significant at the 5% level.

5.1 Deep parameters

First, the law of motion parameters in panel a are indicative of significant diminishing returns to the health production function ($\alpha = 0.69$). Moreover, depreciation is important ($\delta = 7.2\%$), and sickness is rather consequential, with additional depreciation ($\phi = 1.1\%$) suffered upon realization of the health shock.

Second, in panel b the intensity parameters indicate a high, and significant incidence of health shocks ($\lambda_{s0} = 0.29$). The death intensity (3) parameters reject the null of exogenous exposure to death risk ($\lambda_{m1}, \xi_m \neq 0$), indicating that agent’s health decisions are consequential for their expected life horizon. These parameters are also realistic with respect to expected longevity. In particular, Hugonnier et al. (2013) show that an age-$t$ person’s remaining life expectancy can be computed using:

$$\ell(H_t) = (1/\lambda_{m0})(1 - \lambda_{m1}\kappa_0H_t^{-\xi_m}), \quad \text{where } \kappa_0 = [\lambda_{m0} - F(-\xi_m)]^{-1} > 0,$$

The average age in our HRS sample is 75.3 years which can be added to $\ell(H)$ to obtain the expected longevity reported in column 3 of Table 3 as an out-of-sample validity check. These numbers are remarkably close to the expected lifetime in 2002 ($77.3$ all, $74.5$ male, $79.9$ female, Arias, 2004).

Third, the returns parameters ($\mu, r, \sigma_S$) are calibrated at standard values in panel c. The income parameters of equation (2) are both significant, and indicative of a positive health effects on income ($\beta \neq 0$), while the the base income $y_0$ is estimated to a value of $8,200$ (representing $10,824$ in 2016). Fourth, the preference parameters in panel d suggest a significant subsistence consumption $a$ of $12,700$ ($16,760$ in 2016), which is

effects of health have also been highlighted (e.g. Hugonnier et al., 2013; Guiso et al., 1996; Rosen and Wu, 2004; Coile and Milligan, 2009; Berkowitz and Qiu, 2006; Goldman and Maestas, 2013; Fan and Zhao, 2009; Yogo, 2009).
larger than base income $y_0$. Both subsistence, and base income values are realistic. Our estimate of the inter-temporal elasticity $\varepsilon$ is larger than one, as required for Theorem 3, and as identified by others using micro data. Aversion to financial risk is realistic ($\gamma = 2.78$), whereas aversion to mortality and morbidity risks are calibrated in the admissible range ($0 < \gamma_m < 1$), and similar to the values set by Hugonnier et al. (2013). Finally, the subjective discount rate is set at usual values ($\rho = 2.5\%$). Overall, we conclude that the estimated structural parameters are economically plausible.

5.2 Induced parameters

Table 2.e reports the induced parameters that are relevant for the admissible, depletion and accelerating subsets. Table 2.f shows that the two main conditions for our theoretical results are verified at these induced parameters. These composite parameters allow us to evaluate the values of the four loci $x(H)$, $y(H)$, $z(H)$, and $w(H)$ at the various self-reported health levels in Table 3, and to plot the corresponding subsets in Figure 3 using the same scaling as the one for the estimation. Finally, and rely on the joint distribution in Table 1 in order to plot the median values of wealth as blue points for $Q_i$ for the poor ($H = 0.5$), and fair ($H = 1.25$) health statuses.

First, the large negative value for $(y_0 - a)/r$ corresponds to a capitalised base income deficit of 92,900$ (122,628$ in 2016), and confirms that condition (16) in Theorem 3 is verified. Moreover, we identify a relatively large marginal-$Q$ of health $B = 0.1148$ in panel e, suggesting that health depletion can remain optimal despite health being very valuable. The zero-wealth admissible health level $\bar{H}_1$ is estimated between the poor, and fair health values.

Second, the value for $D$ in Table 2.e is large, and significant. From the definition of $y(H)$ in (20), a large value of $D$ also entails a very steep health depletion locus in Figure 3. It follows that its minimum is attained at a low $\bar{H}_3 = 0.1743$, with corresponding realistic value of $\bar{W}_3 = 78,100$ ($103,092$ in 2016). Since this value is larger than most observed wealth levels, it follows that the bulk of the population is located in the health depletion subset. Moreover, our estimates are consistent with a narrow accelerating region $\mathcal{AC}$.

\footnote{For example, the 2002 poverty threshold for elders above 65 was $8,628 (source: U.S. Census Bureau).}
\footnote{For example, Gruber (2013) finds estimates centered around 2.0, relying on CEX data.}
\footnote{Adapting the theoretical valuation of health in Hugonnier et al. (2013, Prop. 3) reveals that an agent at the admissible locus (i.e. with $N_0(W,H) = 0$) would value a 0.10 increment in health as $w_h(0.10,W,H) = 0.10 \times B \times 10^6 = 11,480$ ($15,150$ in 2016).}
Figure 3: Estimated depletion, accelerating, and non-admissible regions

Notes: Non-admissible set $\mathcal{N}\mathcal{A}$: shaded red area under red $x(H)$ line. Health depletion set $\mathcal{D}_H$: shaded green area under green $y(H)$ green curve. Acceleration set $\mathcal{AC}$: hatched green area under blue $z(H)$ curve. Wealth depletion set $\mathcal{D}_W$: area above $w(H)$ black curve. Position of loci, and areas evaluated at estimated parameters in Table 2. Quintile levels for wealth quintiles $Q_2,\ldots,Q_5$ are taken from Table 1, and are reported as blue points for health levels poor, and fair.

Indeed, the values for $B, (y_0-a)/r, \xi_m$ are such that intercepts $\bar{H}_1, \bar{H}_2$ are relatively low (i.e. between Fair, and Poor self-reported health), and close to one another (less than one discrete increment of 0.75). This feature of the model is reassuring since we would expect accelerating phases where agents are cutting down expenses in the face of falling health to coincide with the very last periods of the end of life where health is very low.

Finally, as expected from Theorem 3, the estimated wealth depletion locus $w(H)$ is lying between the $x(H), y(H)$ loci. It is also very low, confirming that most of the agents are also in the wealth depletion region. It follows that unless very wealthy, and very unhealthy, the bulk of the population would be located in the $(\mathcal{D}_H \cap \mathcal{D}_W)$ regions. Indeed, as Table 3 makes clear, all the population with at least a Fair level of health, and non-negative financial wealth is located in the joint health and wealth depletion.
Put differently, our estimates do confirm the empirical relevance of optimal closing-down strategies.

5.3 Simulation analysis

The dynamic analysis presented thus far has focused upon local expected changes for health and wealth $E_t[-dH], E_t[-dW]$. In order to assess whether such small anticipated depletion translate into realistic life cycle paths for health and wealth, we conduct a Monte-Carlo simulation exercise described in further details in Appendix C. Figure 4 plots the resulting mean values for the optimal life cycles for financial wealth $W_t$ (panel a), health level $H_t$ (panel b), the remaining years of survival $\ell(H_t)$ (panel c), as well as the value function $V^*(W_t, H_t)$ (panel d). To highlight the effects of initial financial wealth on optimal dynamics, we separate the dynamic paths based on poor, and high initial wealth.

Overall, these results confirm all our previous findings. Consistent with the data, our simulated life cycles feature a rapid end-of-life depletion of both health (Banks et al., 2015; Case and Deaton, 2005; Smith, 2007; Heiss, 2011), and wealth (De Nardi et al., 2015b; French et al., 2006). Indeed, recalling that expected longevity at good health is 79.0 years, the optimal strategy is to bring down net total wealth $N_0(W_t, H_t)$ to zero (i.e. reach the lower limits of admissible set $A$) at terminal age at which stage agents are indifferent between life and death (panel d). This objective is attained by running down wealth (panel a) very rapidly (consistent with our finding of low $w(H)$ locus), and a somewhat slower decline for health (panel b).

Contrasting rich versus poor cohorts reveals that, as expected, wealth (panel a), and health (panel b) depletion is faster for poor agents, except towards age 80 where attrition implies that only the relatively healthy poor agents survive.\footnote{The short-lived increase in wealth for poor and unhealthy agents after age 80 occurs as they exit the $D_W$ region below the $w(H)$ locus.} The health differences with the rich are therefore attenuated around that age. The joint health and wealth depletion means that low-wealth individuals approach the non-admissible subset more rapidly. Moreover, worse health entails that exposure to death risk is higher for the poor, resulting in shorter residual horizon (panel c), consistent with stylized facts.\footnote{For example, longevity for males from a 1940 cohort in HRS based on deciles of career earnings are 73.3 years (1st decile), 77.9 (3rd decile), 81.8 (6th decile), and 84.6 (10th decile) (Bosworth et al., 2016, Tab. IV-4, p. 87).} These results again accord with the model predictions: poor and rich agents exhaust net total
wealth and therefore become indifferent between life and death (panel d) by the time they approach the zero expected remaining lifetime (panel c). Put differently, our simulations indicate that agents entering the last period of life optimally select an expected lifespan given current health and wealth, and choose allocations that are consistent with optimal closing down, i.e. depletion of the health and wealth capitals during their remaining lifetime. High initial wealth thus has a moderating effect on the speed of the depletion, but not on its ultimate outcome.

Figure 4: Simulated optimal paths

Notes: Mean values for simulated optimal life cycles taken over surviving admissible agents from an initial population of 1,000 agents with 1,000 replications. Rich ($W_{t=76} \geq 2^{nd}$ tercile, blue lines, left-hand scale in panels a, d), and poor ($W_{t=76} \leq 1^{st}$ tercile, red lines, right-hand scale in panels a, d).

The negative expected residual survival in panel c is explained by the stochastic properties of the joint health and wealth processes. Since death remains a non-degenerate random process throughout, some agents with very low health remain alive beyond their expected lifetime.
6 Conclusion

This paper identifies conditions under which agents approaching the end of life optimally select to close down the shop, i.e. run down their health, and wealth capitals, bringing them to a state where the probability of death is high, and where they are indifferent between life and death. We rely on analytical solutions to a life cycle model of optimal health spending and insurance, portfolio, and consumption to characterize the end of life dynamics for health, and wealth. Our findings can be summarized as follows. First, under plausible, and empirically verified conditions, agents optimally choose an expected depletion of their health capital, unless they are sufficiently healthy and wealthy. We also identify a threshold wealth level below which health decline is independent on how healthy the agent is. Moreover, this depletion is accelerated below certain levels of health and wealth. Importantly, wealth is also expected to fall for most, such that all agents eventually close down the shop.

The incidence of depletion strategies can be reduced by increasing base income (e.g. through enhanced Social Security, Medicaid, or minimal revenue programs). However, whereas the positive arguments are readily obtained, the normative reasons for intervening are much less clear. Indeed, continuous depletion of the health stock leading to very high death risks, and indifference between life and death is optimally selected, even in the case of agents with no predisposition for early death. Moreover, this downward spiral is obtained in a complete markets setting, such that no market failure argument for intervention can be invoked. Finally, terminal poverty is endogenously determined as an optimal state such that redistributive rationales for intervening cannot be made.

References


American Medical Association (2016) ‘AMA policy on provision of life-sustaining medical treatment.’ AMA website


Banks, James, Richard Blundell, Peter Levell, and James P. Smith (2015) ‘Life-cycle consumption patterns at older ages in the US and the UK: can medical expenditures explain the difference?’ IFS Working Papers W15/12, Institute for Fiscal Studies


Park, Sojung, BoRin Kim, and Soojung Kim (2016) ‘Poverty and working status in changes of unmet health care need in old age.’ *Health Policy* pp. –. forthcoming


A Parametric restrictions

A.1 Regularity and transversality restrictions

Define the following elements:

\[ \chi(x) = 1 - (1 - \phi)^{-x}, \]
\[ F(x) = x(\alpha B)^{\frac{x}{\gamma}} - x\delta - \lambda_{0}\chi(-x), \]
\[ L_{m} = [(1 - \gamma_{m})(A - F(-\xi_{m}))]^{-1} > 0. \]

The theoretical model is solved under three regularity and transversality conditions that are reproduced for completeness:

\[ \beta < \left( r + \delta + \phi\lambda_{0}\right)^{\frac{1}{\gamma}}, \]  
\[ \max \left( 0; r - \frac{\lambda_{0}}{1 - \gamma_{m}} + \frac{\theta^{2}}{\gamma} \right) < A, \]  
\[ 0 < A - \max \left( 0, r - \frac{\lambda_{0}}{1 - \gamma_{m}} + \frac{\theta^{2}}{\gamma} \right) - F(-\xi_{m}), \]

where the consumption parameter \( A \), and the price of health \( B \), are defined in (42), and in (45).

A.2 Closed-form solutions for optimal rules parameters

The closed-form expression for the parameters in the optimal rules are obtained as follows. The insurance parameter in (11) is given as:

\[ X_{1} = \lambda_{m1}\chi(\xi_{m}) \left( \frac{1}{\gamma_{m}} - 1 \right) L_{m}. \]

The consumption parameters in (12) are:

\[ A = \varepsilon\rho + (1 - \varepsilon) \left( r - \frac{\lambda_{m0}}{1 - \gamma_{m}} + \frac{\theta^{2}}{2\gamma} \right), \]  
\[ C_{1} = \lambda_{m1}A(\varepsilon - 1)L_{m}. \]
The parameters of the optimal investment in (14) are:

\[ K = \alpha \frac{1}{\alpha - 1} B \frac{1}{\alpha - 1}, \]
\[ I_1 = \lambda_1 (\xi_m K/(1 - \alpha)) L_m, \] (44)

where the price of health \( B \) of health solves:

\[ g(B) = \beta - (r + \delta) B - (1 - 1/\alpha)(\alpha B)^{\frac{1}{1-\alpha}} \]
\[ = \beta - (r + \tilde{\delta}) B + \left( \frac{1 - \alpha}{\alpha} \right) BK = 0 \] (45)

subject to:

\[ g'(B) = -(r + \tilde{\delta}) + (\alpha B)^{\frac{\alpha}{1-\alpha}} \]
\[ = -(r + \tilde{\delta}) + \frac{BK}{\alpha B} < 0. \]

Combining the two thus implies that:

\[ BK < \beta. \] (46)

B Proofs

B.1 Theorem 1

See Hugonnier et al. (2013, Thm. 2) for the general case, and evaluate the optimal policies at the restricted exogenous morbidity case \( \lambda_{s1} = 0. \)

B.2 Theorem 2

B.2.1 Health depletion

First, substituting the investment-to-capital ratio (18) in the expected local change for health (17) shows that:

\[ E_t[-dH] = \left\{ [BK + I_1 H^{-\xi_m^{-1}} N_0(W, H)]^{\alpha} - \tilde{\delta} \right\} H dt. \]
By admissibility, \(N_0(W, H) > 0\), whereas \(I_1 \geq 0\) in (44). It follows that

\[ E_t - [dH]_{t=\lambda_m=0} \begin{cases} \begin{array}{ll}
[BK]^{\alpha} - \delta 
\end{array} \end{cases} Hdt < E_t-[dH]_{t=\lambda_m=0} \leq 0, \]

Since \(E_t-[dH] < 0\) by definition of health depletion, condition (25) – which is induced by conditions (19), and (46) – is thus required for order-0 depletion, and therefore a necessary condition for existence of depletion zones when mortality risk is endogenous.

Second using the definition of net total wealth (10) shows that we can write:

\[ E_t - [dH] < 0 \iff W < y(H) = x(H) + DH^{1+\xi_m}, \quad D = \mathcal{I}_1^{-1}\left[\delta^{1/\alpha} - BK\right]. \]

Conditions (19) and (46) jointly imply that \(BK < \delta^{1/\alpha} \implies D > 0\) in (20). Condition (25) is therefore sufficient for \(y(H) \geq x(H), \forall H\), i.e. the locus \(y(H)\) lies everywhere in the admissible zone, and consequently \(\mathcal{D}_H\) is a non-empty subset of \(\mathcal{A}\). Third, observe that the health depletion locus is characterized by:

\[ y_H(H) = -B + (1 + \xi_m)DH^{1+\xi_m} \begin{cases} < 0, & \text{if } H < \bar{H}_3, \\
= 0, & \text{if } H = \bar{H}_3, \quad \text{and} \\
> 0, & \text{if } H > \bar{H}_3, \end{cases} \]

\[ y_{HH}(H) = \xi_m(1 + \xi_m)DH^{1+\xi_m} > 0. \]

The locus \(y(H)\) is therefore convex, and U-shaped and attains a unique minima at \(\bar{H}_3\) in the \((H,W)\) space, where \(\bar{H}_3\) is given in (22), with corresponding wealth level \(\bar{W}_3 = y(\bar{H}_3)\).

**B.2.2 Acceleration**

Taking the derivative of the investment-to-health ratio (18) with respect to \(H\) and rearranging shows that the accelerating region can be characterized by:

\[ I^h_H(W, H) > 0 \iff W < -(y_0 - a)/r - \frac{BH\xi_m}{1 + \xi_m} = x(H) + \frac{BH}{1 + \xi_m} = z(H). \]

Since \(B, \xi_m > 0\), \(x(H) \leq z(H)\), i.e. this locus lies above the \(x(H)\) locus, and is therefore admissible, i.e. \(\mathcal{A}C \subset \mathcal{A}\). Observe furthermore that \(z(0) = x(0) = y(0) = -(y_0 - a)/r\),
and that:

\[ z(H) - y(H) = H \left[ \frac{B}{1 + \xi_m} - DH^{\xi_m} \right] \begin{cases} > 0, & \text{if } H < \bar{H}_3 \\ = 0, & \text{if } H = \bar{H}_3 \\ < 0, & \text{if } H > \bar{H}_3 \end{cases} \]

again using the definition of $\bar{H}_3$ in (22). Consequently, the $z(H)$ locus is downward-sloping, has the same intercept and intersects $y(H)$ at its unique minimal value $\bar{H}_3$, and lies above (below) the $y(H)$ locus for $H < \bar{H}_3$ ($H > \bar{H}_3$). It follows that the acceleration set (i.e. the health depletion subset where $I_h > 0$) is the entire $D_H$ for $H \in [0, \bar{H}_3]$, and otherwise the area between $y(H), z(H)$, as given in (23), and (24).

### B.3 Theorem 3

Substituting the optimal investment (14), consumption (12), risky portfolio (13), and insurance (11) in the expected local change for wealth (26), and using the definition of net total wealth (10) reveals that

\[ E_t[dW] = 0 \iff Wl(H) = x(h)[l(h) + r] + k(H), \]

where

\[ l(H) = \left[ A - \sigma_s \theta L_0 - r + (\mathcal{I}_1 + \mathcal{C}_1) H^{-\xi_m} \right], \]

\[ k(H) = (y_0 - a) + H(\beta - BK), \]

as given in (30), (31). Observe from (43), and (44) that $\epsilon \geq 1$ implies that $\mathcal{C}_1 \geq 0$, and consequently that $\mathcal{I}_1 + \mathcal{C}_1 \geq 0$. It follows that condition (27) in (30) is therefore sufficient to guarantee that $l(H) > 0, \forall H$. Consequently, the wealth depletion zone $D_W$ is delimited by:

\[ W > \frac{x(H)[l(H) + r]}{l(H)} = w(H). \]
It is straightforward to show that:

$$\lim_{H \to 0} w(H) = x(0) = -(y_0 - a)/r$$

such that the $w(H)$ shares the same intercept with $x(H), y(H), z(H)$, and which is nonnegative under condition (16).

In order to demonstrate that the joint health and wealth depletion set $D = (D_W \cap D_H) \neq \emptyset$ we must show that the wealth depletion set $D_W$ lies everywhere in the admissible set, i.e. $x(H) \leq w(H)$, and everywhere below the health depletion locus, i.e. $w(H) \leq y(H)$. Using the health price $B$ definition (45), we first show that:

$$x(H) \leq w(H) \iff 0 \leq x(H)r + k(H)$$
$$\iff 0 \leq \beta - Br + K$$
$$\iff 0 \leq B(\tilde{\delta} - K/\alpha)$$
$$\iff 0 \leq \tilde{\delta}^{1/\alpha} - BK$$

which is verified under under induced restriction (25). Consequently the locus is everywhere admissible.

Second, we show that $w(H) \leq y(H)$, which simplifies to:

$$rx(H) + k(H) \leq Dl(H)H^{1+\xi_m} \iff \beta - Br - \tilde{\delta}^{1/\alpha} \leq DH^{\xi_m} \left[ A - \frac{\theta^2}{\gamma} - r \right] + C_1D$$

Since $\beta < \tilde{\delta}^{1/\alpha}$ under (19), the left-hand side is negative. Moreover, $D > 0$ under condition (25), whereas the right-hand term in square bracket is also positive under condition (27), and $C_1 \geq 0$ for $\varepsilon \geq 1$. It follows that the right-hand side is positive, and consequently sufficient for $w(H) \leq y(H)$, as required.

**B.4 Theorem 4**

First, setting $\lambda_m = 0$ results in the first-order adjustment $I_1 = 0$ in (44). Consequently, the investment-to-capital ratio in (18) is constant, and given by $I^h = BK$. Substituting
in (17) reveals that so is the expected growth rate:

\[ E_t[-dH] = \left[(BK)^\alpha - \delta \right] H dt, \]

and that the latter is negative under condition (25) for all admissible health and wealth levels. Consequently, the health depletion subset corresponds to the admissible subset, as stated in (32). Moreover, a constant \( I^h \) implies that it is orthogonal to the health status; consequently no accelerating region exists as stated in (33).

Finally, setting \( \lambda_{m1} = 0 \) also results in \( I_1, C_1 = 0 \) in equation (30) for \( l(H) \), and modifies the locus \( w(H) \) accordingly in (34). As stated before, \( 0 \leq x(H) r + k(H) \) ensures that \( x(H) \leq w(H) \), and is induced by the restriction (25). Because the health depletion set is the entire admissible set, the conditions relating \( w(H) \), and \( y(H) \) are irrelevant, and the joint health and wealth depletion set is everywhere non-empty.

### C Monte-Carlo simulation

The Monte-Carlo framework used to simulate the dynamic model is as follows:

1. Relying on a total population of \( n = 1,000 \) individuals, we initialize the health and wealth distributions at base age \( t = 75 \) using the observed unconditional distribution for health \( P(H) \), as well as the conditional wealth distribution \( P(W \mid H) \).

2. We simulate individual-specific Poisson health shocks \( dQ_s \sim P(\lambda_{s0}) \), as well as a population-specific sequence of Brownian financial shocks \( dZ \sim N(0, \sigma_s^2) \) over a 10-year period \( t = 75, \ldots, 85 \).

3. At each time period \( t = 75, \ldots, 85 \), and using our estimated and calibrated parameters:
   
   (a) For each agent with health \( H_t \), we generate the Poisson death shocks with endogenous intensities \( dQ_m \sim P[\lambda_m(H_t)] \), and keep only the surviving agents, with positive wealth (as imposed in the estimation) for the computation of the statistics.
   
   (b) We verify admissibility, for each agent with health and wealth \( (H_t, W_t) \) and keep only surviving agents in the admissible region.
(c) We use the optimal rules $I(W_t, H_t), c(W_t, H_t), \Pi(W_t, H_t), X(H_{t-}),$ as well as income function $Y(H_t),$ and the sickness and financial shocks $dQ_{s,t}, dZ_t$ in the stochastic laws of motion $dH_t, dW_t.$

(d) We update the health and wealth variables using the Euler approximation:

$$H_{t+1} = H_t + dH_t(H_t, I_t, dQ_{s,t})$$
$$W_{t+1} = W_t + dW_t[W_t, c(W_t, H_t), I(W_t, H_t), \Pi(W_t, H_t), X(W_t, H_t), dQ_{s,t}, dZ_t]$$

4. We replicate the simulation 1–3 for 1,000 times.

5. We rely on age-76 wealth to separate sub-samples as:

- Poor: $W_{76} \leq Q(1/3),$
- Rich: $W_{76} \geq Q(2/3),$

in order to compute the sub-sample means using only agents who are alive, and within the admissible subset.

D Tables

---

29 The morbidity risk aversion parameter involved in the optimal insurance $X^*(W, H)$ is calibrated at the same value of $\gamma_s = 7.40$ in Hugonnier et al. (2013).
Table 1: HRS data statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Wealth quintile</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial wealth (W)</td>
<td>a. Poor health ((H = 0.5))</td>
<td>0.000</td>
<td>0.051</td>
<td>0.075</td>
<td>0.519</td>
<td>100.012</td>
</tr>
<tr>
<td></td>
<td>- Quintile</td>
<td>0.000</td>
<td>0.030</td>
<td>0.220</td>
<td>0.814</td>
<td>2.930</td>
</tr>
<tr>
<td></td>
<td>- Median</td>
<td>0.379</td>
<td>0.417</td>
<td>0.469</td>
<td>0.427</td>
<td>0.615</td>
</tr>
<tr>
<td></td>
<td>Investment ((I))</td>
<td>0.005</td>
<td>0.079</td>
<td>0.216</td>
<td>0.485</td>
<td>0.800</td>
</tr>
<tr>
<td>Financial wealth (W)</td>
<td>b. Fair health ((H = 1.25))</td>
<td>0.000</td>
<td>0.030</td>
<td>0.210</td>
<td>0.983</td>
<td>71.000</td>
</tr>
<tr>
<td></td>
<td>- Quintile</td>
<td>0.000</td>
<td>0.030</td>
<td>0.230</td>
<td>0.760</td>
<td>3.400</td>
</tr>
<tr>
<td></td>
<td>- Median</td>
<td>0.255</td>
<td>0.254</td>
<td>0.233</td>
<td>0.252</td>
<td>0.266</td>
</tr>
<tr>
<td></td>
<td>Investment ((I))</td>
<td>0.000</td>
<td>0.046</td>
<td>0.253</td>
<td>0.514</td>
<td>0.782</td>
</tr>
<tr>
<td>Financial wealth (W)</td>
<td>c. Good health ((H = 2.0))</td>
<td>0.010</td>
<td>0.100</td>
<td>0.402</td>
<td>1.407</td>
<td>45.000</td>
</tr>
<tr>
<td></td>
<td>- Quintile</td>
<td>0.000</td>
<td>0.040</td>
<td>0.220</td>
<td>0.770</td>
<td>3.300</td>
</tr>
<tr>
<td></td>
<td>- Median</td>
<td>0.157</td>
<td>0.149</td>
<td>0.156</td>
<td>0.129</td>
<td>0.168</td>
</tr>
<tr>
<td></td>
<td>Investment ((I))</td>
<td>0.002</td>
<td>0.082</td>
<td>0.299</td>
<td>0.510</td>
<td>0.824</td>
</tr>
<tr>
<td>Financial wealth (W)</td>
<td>d. Very good health ((H = 2.75))</td>
<td>0.040</td>
<td>0.222</td>
<td>0.720</td>
<td>2.100</td>
<td>71.000</td>
</tr>
<tr>
<td></td>
<td>- Quintile</td>
<td>0.000</td>
<td>0.040</td>
<td>0.230</td>
<td>0.840</td>
<td>3.500</td>
</tr>
<tr>
<td></td>
<td>- Median</td>
<td>0.100</td>
<td>0.112</td>
<td>0.106</td>
<td>0.105</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>Investment ((I))</td>
<td>0.011</td>
<td>0.107</td>
<td>0.368</td>
<td>0.604</td>
<td>0.854</td>
</tr>
<tr>
<td>Financial wealth (W)</td>
<td>e. Excellent health ((H = 3.5))</td>
<td>0.050</td>
<td>0.280</td>
<td>0.874</td>
<td>2.800</td>
<td>100.120</td>
</tr>
<tr>
<td></td>
<td>- Quintile</td>
<td>0.000</td>
<td>0.050</td>
<td>0.210</td>
<td>0.800</td>
<td>3.820</td>
</tr>
<tr>
<td></td>
<td>- Median</td>
<td>0.137</td>
<td>0.065</td>
<td>0.063</td>
<td>0.105</td>
<td>0.091</td>
</tr>
<tr>
<td></td>
<td>Investment ((I))</td>
<td>0.010</td>
<td>0.131</td>
<td>0.350</td>
<td>0.520</td>
<td>0.861</td>
</tr>
</tbody>
</table>

Notes: Quintile, and median values of wealth, and mean values (investment, risky holdings), measured in 100K$ (year 2002) per health status, and wealth quintiles for HRS data used in estimation.
Table 2: Estimated and calibrated parameter values

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>a. Law of motion health (1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.6940*</td>
<td>$\delta$</td>
<td>0.0723*</td>
<td>$\phi$</td>
<td>0.011c</td>
</tr>
<tr>
<td></td>
<td>(0.1873)</td>
<td></td>
<td>(0.0366)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>b. Sickness and death intensities (3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{s0}$</td>
<td>0.2876*</td>
<td>$\lambda_{m0}$</td>
<td>0.2356*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.1419)</td>
<td></td>
<td>(0.0844)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\lambda_{m1}$</td>
<td>0.0280*</td>
<td>$\xi_{m}$</td>
<td>2.8338*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0108)</td>
<td></td>
<td>(1.1257)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>c. Income and wealth (2), (5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$y_0$</td>
<td>0.0082*§</td>
<td>$\beta$</td>
<td>0.0141*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0029)</td>
<td></td>
<td>(0.0059)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.108c</td>
<td>$r$</td>
<td>0.048c</td>
<td>$\sigma_S$</td>
<td>0.20c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d. Preferences (6), (7)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a$</td>
<td>0.0127*§</td>
<td>$\varepsilon$</td>
<td>1.6738*</td>
<td>$\gamma$</td>
<td>2.7832*</td>
</tr>
<tr>
<td></td>
<td>(0.0063)</td>
<td></td>
<td>(0.6846)</td>
<td></td>
<td>(1.3796)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.025c</td>
<td>$\gamma_m$</td>
<td>0.75c</td>
<td>$\gamma_s$</td>
<td>7.40c</td>
</tr>
<tr>
<td></td>
<td></td>
<td>e. State space subsets (15), (21), (24), (29)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(y_0 - a)/r$</td>
<td>-0.0929*§</td>
<td>$B$</td>
<td>0.1148*</td>
<td>$H_1$</td>
<td>0.8093*</td>
</tr>
<tr>
<td>$D$</td>
<td>4.5088*</td>
<td>$I_1$</td>
<td>0.0053*</td>
<td>$K$</td>
<td>0.0022*</td>
</tr>
<tr>
<td>$\bar{H}_3$</td>
<td>0.1743*</td>
<td>$\bar{W}_3$</td>
<td>0.0781*§</td>
<td>$\bar{H}_2$</td>
<td>1.0460*</td>
</tr>
<tr>
<td>$C_1$</td>
<td>0.1115*</td>
<td>$A$</td>
<td>0.6336*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>f. Conditions (19), (27) (must be negative)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta - \delta^{1/\alpha}$</td>
<td>-0.0086*</td>
<td>$\theta^2/\gamma + r - A$</td>
<td>-0.5533*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: *: Estimated structural and induced parameters (standard errors in parentheses), significant at 5% level; c: calibrated parameters; §: In $\text{M}$. 
Table 3: Estimated longevity and values of loci

<table>
<thead>
<tr>
<th>Level</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>A</th>
<th>D_H</th>
<th>AC</th>
<th>D_W</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H</td>
<td>% Pop.</td>
<td>Exp. longev.</td>
<td>x(H)</td>
<td>y(H)</td>
<td>z(H)</td>
<td>w(H)</td>
</tr>
<tr>
<td>Poor</td>
<td>0.50</td>
<td>10.7</td>
<td>51.94</td>
<td>0.04</td>
<td>0.35</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td>Fair</td>
<td>1.25</td>
<td>21.1</td>
<td>77.49</td>
<td>−0.05</td>
<td>10.56</td>
<td>−0.01</td>
<td>−0.03</td>
</tr>
<tr>
<td>Good</td>
<td>2.00</td>
<td>31.5</td>
<td>79.00</td>
<td>−0.14</td>
<td>64.15</td>
<td>−0.08</td>
<td>−0.11</td>
</tr>
<tr>
<td>Very good</td>
<td>2.75</td>
<td>26.9</td>
<td>79.32</td>
<td>−0.22</td>
<td>217.74</td>
<td>−0.14</td>
<td>−0.18</td>
</tr>
<tr>
<td>Excellent</td>
<td>3.50</td>
<td>9.9</td>
<td>79.43</td>
<td>−0.31</td>
<td>549.12</td>
<td>−0.20</td>
<td>−0.26</td>
</tr>
</tbody>
</table>

Notes: Estimated longevity $75.3 + \ell(H)$ in (38); values (in MM$) of admissible $x(H)$; health depletion $y(H)$; accelerating $z(H)$; and wealth depletion $w(H)$ at observed health levels.