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

Julien Hugonnier[1,4,5], Florian Pelgrin[2,6] and Pascal St-Amour[3,4,6]

[1]École Polytechnique Fédérale de Lausanne

[2]EDHEC Business School

[3]HEC Lausanne, University of Lausanne

[4]Swiss Finance Institute

[5]CEPR

[6]CIRANO

August 13, 2019

*Financial support from the Swiss Finance Institute is gratefully acknowledged. We thank the Associate Editor (Owen O'Donnell) and two anonymous referees for constructive remarks. We are also thankful to Audrey Laporte, Alex Michaelides, Kim Peijnenburg, Shang Wu and seminar participants at the 2016 European Workshop on Econometrics and Health Economics (Nyborg, DK), the 2017 NETSPAR International Pension Workshop (Leiden, NL), the 2017 Workshop on Household Finance and Retirement Savings (Torino, I) for very useful comments and discussions.

**Abstract**

Near the end of life, health declines, mortality risk increases and curative is replaced by uninsured long-term care, accelerating the fall in wealth. Whereas standard explanations emphasize inevitable aging processes, we propose a complementary *closing down the shop* justification where agents' decisions affect their health and the timing of death. Despite preferring to live, individuals optimally deplete their health and wealth towards levels associated with high death risk and gradual indifference between life and death. Reinstating exogenous aging processes reinforces the relevance of closing down. Using HRS-CAMS data for elders, a structural estimation of the closed-form decisions identifies, tests and confirms the relevance of closing down.

**JEL classification:** D15, I12

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

# 1 Introduction

Health declines steadily throughout the life cycle and deteriorates faster as we approach the last period of life.[1] Because how healthy we currently are is a significant predictor of future major health onsets,[2] exposure to death risk also rises at an increasing rate.[3] However, falling health is not uniform across individuals, but varies significantly with socioeconomic status (SES). Indeed, poor agents suffer faster deterioration in health[4] and in longevity,[5] with SES gradients being strongest around mid life and falling thereafter.[6] Moreover, health spending augments and changes in composition,[7] with long-term care out-of-pocket expenses increasing sharply towards the end of life, leading to a rapid drain in financial resources.[8]

The standard demand-for-health framework (Grossman, 1972; Ehrlich and Chuma, 1990) does allow for optimal declines in health when the net return on health investment is sufficiently low. Accelerating deterioration obtains when suitable exogenous aging dynamics are appended, such as age-increasing depreciation or sickness risks. Consequently, curative expenses are curtailed to accompany, but not revert, the age-induced decline in health and longevity. Given an exogenous remaining life horizon, the wealth draw-down objectives consist of ensuring sufficient resources to reach the end of life and, potentially, leave bequests (De Nardi et al., 2016; van Ooijen et al., 2015; St-Amour, 2018).

However, such models fail to capture part of the life cycle and cross-sectional evidence.[9] They counterfactually predict that decline in health is wealth- and health-independent, and assume it is inconsequential for mortality risk. Conditional on age, technology, risk distributions and preferences, all agents face the same rates of deterioration and mortality risk, independent of financial resources and current health status. In addition, these predictions are at odds with the observation that the health decline is strongest among the poor and that the SES health gradients weaken in the last phase of life.

This paper addresses these life cycle and cross-sectional shortcomings in a model emphasizing endogenous longevity obtained by appending health-dependent mortality to a demand-for-health setup. We ensure preference for life through recursive preferences and dynamically-consistent decisions. End-of-life dynamics can then be rationalized as *closing down the shop* whereby individuals who prefer life over death nonetheless optimally relinquish their financial, health and life capital stocks.

First, agents optimally select a dynamic path towards death involving depletion of both health and financial capital, as well as an increasing exposure to mortality risk. This path is enacted by curtailing health investment in order to initiate (and eventually hasten) the health decline. While the latter induces a corresponding increase in death risk, dynamic consistency ensures that agents gradually become indifferent between life and death in the last phase of life. Second, accelerating dynamics emerge independently of aging when health is an input to producing better health and longer life expectancy and when the decision to invest or not depends on the (adjustable) planning horizon. Finally, rich individuals delay entering this path for longer than poorer agents. However, optimal declines in wealth entail that *all* agents eventually choose to let health and longevity deteriorate. This guarantees that SES health and mortality gradients are initially strong, but gradually weaken in the later part of life.

Our main theoretical contribution is to prove the optimality of closing down dynamics near the end of life. Agents that prefer life over death nonetheless simultaneously act in a manner that *results* in a short terminal horizon, *and* they select a depletion strategy that is consistent with this horizon. This simultaneous feedback between decisions and horizon 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.[10] Remarkably, this optimal depletion reinforces a biological aging explanation, and reintroducing aging makes closing down even more relevant.

Our second contribution is to assess whether closing down dynamics are empirically meaningful. Taking a structural econometric characterization of the health and wealth loci where these strategies are to be expected, we test conditions and identify thresholds under which closing down does, or does not take place. Using HRS-CAMS data for relatively old (mean 75) agents, our results provide evidence that the bulk of agents optimally select to close down the shop.

# 2 Theoretical background

## 2.1 Life cycle model

Our analysis of the life cycle dynamics of health and wealth builds upon on the theoretical framework developed in Hugonnier et al. (2013), briefly reproduced here for completeness. First, the dynamics for the agent's health are:

$$\mathrm{d}H_t = \left( (I_t/H_{t-})^{\alpha} - \delta \right) H_{t-}\mathrm{d}t - \phi H_{t-}\mathrm{d}Q_{st}, \quad H_0 > 0, \tag{1}$$

where $\alpha, \delta, \phi \in (0,1)$ and $I_t > 0$ is investment (health expenses).[11] We denote by $H_{t-} = \lim_{s \uparrow t} H_s$ the agent's health prior to the realization of the morbidity shock $Q_{st}$, where the latter follows a Poisson distribution with exogenous intensity $\lambda_{s0}$.[12] Second, the age at death $T_m$ is also distributed as Poisson, although with endogenous death intensity:

$$\lim_{h \to 0} (1/h) P_t \left[ t < T_m \leq t + h \right] = \lambda_{m0} + \lambda_{m1} H_{t-}^{-\xi_m} \equiv \lambda_m(H_{t-}), \tag{2}$$

whereby healthier agents face a lower likelihood of dying.

Third, the dynamics for financial[13] wealth $W_t$ are given by:

$$\mathrm{d}W_t = (rW_{t-} + Y_t - C_t - I_t)\mathrm{d}t + \Pi_t \sigma_S \left( \mathrm{d}Z_t + \theta \mathrm{d}t \right) + X_t \left( \mathrm{d}Q_{st} - \lambda_{s0}\mathrm{d}t \right) \tag{3a}$$

with constant (e.g. annuity) income:

$$Y_t = y_0. \tag{3b}$$

We denote $C$ as consumption of nondurables and services, including health-related expenses that do not readily classify as investment $I$,[14] $\Pi$ as the portfolio invested in the risky asset with Brownian motion $Z$ and market price of financial risk $\theta = (\mu - r)/\sigma_S$, and $X$ as the actuarially-fair insurance against morbidity shock.

Fourth, the agent's objective is to solve:

$$V(W_t, H_t) = \sup_{(C,\Pi,X,I)} 1_{\{T_m > t\}} U_t \tag{4}$$

subject to (1)–(3). Preferences are defined from the continuous-time analog of the Epstein-Zin recursive preferences (Stochastic Differential Utility, Duffie and Epstein, 1992) where the continuation utility solves the recursive integral:

$$U_t = E_t \int_t^{\infty} e^{-\int_t^{\tau}\left(\frac{\lambda_m(H_v)}{1-\gamma_m}\right)\mathrm{d}v}\left(f(C_\tau, H_{\tau-}, U_{\tau-}) - \frac{\gamma|\sigma_\tau(U)|^2}{2U_{\tau-}} - F_s(U_{\tau-}, \Delta_s U_\tau)\right)\mathrm{d}\tau. \quad (5a)$$

The exposure to Brownian financial risk is $\sigma_t = 1/\mathrm{d}t\,\mathrm{d}\langle U, Z\rangle_t$, while the Kreps-Porteus aggregator $f(C, H, U)$ and the penalty $F_s(U, \Delta_s U)$ for exposure to Poisson morbidity shocks in utility $\Delta_s U_t = 1_{\{\mathrm{d}Q_{st}\neq 0\}}(U_t - U_{t-})$ are given by:

$$f(C, H, U) = \frac{\rho U}{1-1/\varepsilon}\left(((C - a + \beta H)/U)^{1-1/\varepsilon} - 1\right), \quad (5b)$$

$$F_s(U, \Delta U) = \lambda_{s0}\left[\frac{\Delta U}{U} + \frac{1-(1+\Delta U/U)^{1-\gamma_s}}{1-\gamma_s}\right] U. \quad (5c)$$

In this formulation, $a > 0$ denotes minimal (subsistence) consumption, $\beta$ captures utilitarian service flows from health[15] and $\rho$ is a discount rate. The elasticity of intertemporal substitution $\varepsilon \geq 0$ is disentangled from the source-dependent risk aversion parameters $\gamma \geq 0$ (financial risk), $\gamma_m \in (0, 1)$ (death risk) and $\gamma_s \geq 0$ (sickness risk).

The homogeneity of preferences ensures weak preference for life over death: $V(W_t, H_t) \geq 0$. Moreover, this model nests the Grossman (1972); Ehrlich and Chuma (1990) as a special case where morbidity and insurance are both abstracted from, mortality is exogenous, and preferences are VNM (see Section B.3 for implications):

$$\lambda_{s0}, \phi, X_t, \lambda_{m1}, \gamma_m, \gamma_s = 0, \text{ and } \gamma = 1/\varepsilon. \quad (6)$$

It is also straightforward to show that the functional forms we rely upon for health accumulation technology in (1) and costs in (3a) encompass other ones found in the literature.[16]

## 2.2 Optimal allocation

The optimal rules for this model are obtained as follows.

**Theorem 1 (Optimal rules)** *Assume that conditions (17) of Appendix A hold true and define net total wealth as:*

$$N_0(W_{t-}, H_{t-}) = W_{t-} + BH_{t-} + (y_0 - a)/r. \tag{7}$$

*Up to a first order approximation, the indirect utility and optimal policy functions are given by:*

$$V(W_{t-}, H_{t-}) = \left(\Theta - \mathcal{V}_1 H_{t-}^{-\xi_m}\right) N_0(W_{t-}, H_{t-}) \tag{8a}$$

$$I^*(W_{t-}, H_{t-}) = KBH_{t-} + \mathcal{I}_1 H_{t-}^{-\xi_m} N_0(W_{t-}, H_{t-}) \tag{8b}$$

$$X^*(W_{t-}, H_{t-}) = \phi BH_{t-} + \mathcal{X}_1 H_{t-}^{-\xi_m} N_0(W_{t-}, H_{t-}) \tag{8c}$$

$$C^*(W_{t-}, H_{t-}) = a + \left(A + \mathcal{C}_1 H_{t-}^{-\xi_m}\right) N_0(W_{t-}, H_{t-}) \tag{8d}$$

$$\Pi^*(W_{t-}, H_{t-}) = (\theta/(\gamma\sigma_S))\, N_0(W_{t-}, H_{t-}) \tag{8e}$$

*where $B \geq 0$ solves (18), $K = \alpha^{\frac{1}{1-\alpha}} B^{\frac{\alpha}{1-\alpha}}$, the parameters $(A, \Theta) \geq 0$ are defined in (19), and where the constants $(\mathcal{V}_1, \mathcal{I}_1) \geq 0$ and $(\mathcal{X}_1, \mathcal{C}_1) \gtreqless 0$ are defined in (20) in Appendix A.*

**Proof**. See Hugonnier et al. (2013, Thms. 1, 2, and Remark 3) for the general case and evaluate the optimal policies at the restricted exogenous morbidity case $\lambda_{s1} = 0$. ■

The net total wealth $N_0(W_{t-}, H_{t-})$ in (7) is the sum of financial assets and capitalized future income, net of subsistence consumption expenditures $a$, where $B \geq 0$ in (18) represents the shadow price (i.e. Tobin's-$Q$) of health. The expressions involving $\mathcal{V}_1, \mathcal{I}_1, \mathcal{X}_1$ and $\mathcal{C}_1$ capture the effects of endogenous mortality $\lambda_{m1} H_{t-}^{-\xi_m}$ on welfare and on the optimal rules, and are all zero when mortality is exogenous. Indeed, the death intensity in (2) mechanically increases from base risk $\lambda_0^m$ when $\lambda_1^m > 0$. Because life is valuable, higher death risk is unconditionally welfare reducing in (8a) (since $\mathcal{V}_1 > 0$), and because mortality can be adjusted, investment consequently increases in (8b) (since $\mathcal{I}_1 > 0$). The effects on consumption and insurance critically depend on preferences. Highly morbidity-risk averse agents ($\gamma_s > 1$) demand more insurance against health shocks in (8c) when exposed to higher death risks. Moreover, agents with high elasticity of inter-temporal substitution ($\varepsilon > 1$) compensate against a shorter life horizon by increasing current consumption – including non-investment health-related expenses – in (8d), thus substituting better quality for less quantity of life. Observe that for all cases, the effects of endogenous mortality are compounded for agents with high net total wealth $N_0(W, H)$.

Importantly, the optimal rules in (8) are defined only over an *admissible* state space, i.e. the set of wealth and health levels that provides minimal resources requirements to ensure survival, as well as strict preference for life over death:

$$C^*(W_{t-}, H_{t-}) - a > 0 \iff V(W_{t-}, H_{t-}) > 0 \iff N_0(W_{t-}, H_{t-}) > 0 \tag{9}$$

This admissible region $\mathcal{A}$ thus requires positive net total wealth in (7):

$$\mathcal{A} = \{(W, H) \in \mathbb{R} \times \mathbb{R}_+ : W \geq x(H) = -(y_0 - a)/r - BH\}, \tag{10}$$

where we assume that base income $y_0$ is insufficient to cover subsistence consumption $a$:

$$(y_0 - a)/r < 0. \tag{11}$$

to ensure consistency with observed financial choices.[17]

## 2.3 Optimal health and wealth dynamics

The agent's health and wealth evolve on the optimal path given by (1) and (3) evaluated at the optimal rules (8). This stochastic differential system cannot be studied with standard phase portraits, and we instead analyze the instantaneous expected changes in health (1) and wealth (3).[18]

We focus on admissible depletion regions $\mathcal{D}_H, \mathcal{D}_W \subset \mathcal{A}$ of the $(W, H)$ space where health and wealth are expected to fall. Moreover, we also study the region $\mathcal{AC} \subset \mathcal{D}_H$ where the health depletion is accelerating, i.e. where a fall in health induces a larger cut in investment, leading to further depletion. Propositions 1, 2 in Appendix B solve for the necessary and sufficient conditions for relevant depletion and acceleration regions, i.e. $(\mathcal{D}_H, \mathcal{D}_W, \mathcal{AC}) \neq \emptyset$. Proposition 3 gives more stringent sufficient conditions for relevance:

$$\tilde{\delta}^{1/\alpha} > \beta, \quad \tilde{\delta} = \delta + \phi\lambda_{s0}, \tag{12a}$$

$$\varepsilon > 1, \tag{12b}$$

$$A > \frac{\theta^2}{\gamma} + r \iff \varepsilon(\rho - r) + (\varepsilon - 1)\frac{\lambda_{m0}}{1-\gamma_m} > (1+\varepsilon)\frac{\theta^2}{2\gamma}. \tag{12c}$$

Condition (12a) states that expected health depreciation is high relative to the health gradient in utility $\beta$. A high depreciation in the absence of investment ($\delta$), or conditional upon sickness ($\phi$), as well as a high likelihood of morbidity shocks ($\lambda_{s0}$) are all to be expected in the last years of life. Other conditions require high elasticity of inter-temporal substitution (12b) and high marginal propensity to consume (12c), obtained through high impatience $\rho$, and/or high aversion to death risk $\gamma_m \in [0, 1)$, and/or high unconditional risk of dying $\lambda_{m0}$, all of which are relevant for end of life.

Under sufficient conditions (12), Propositions 1, 2 of Appendix B identify the depletion and accelerating regions as:

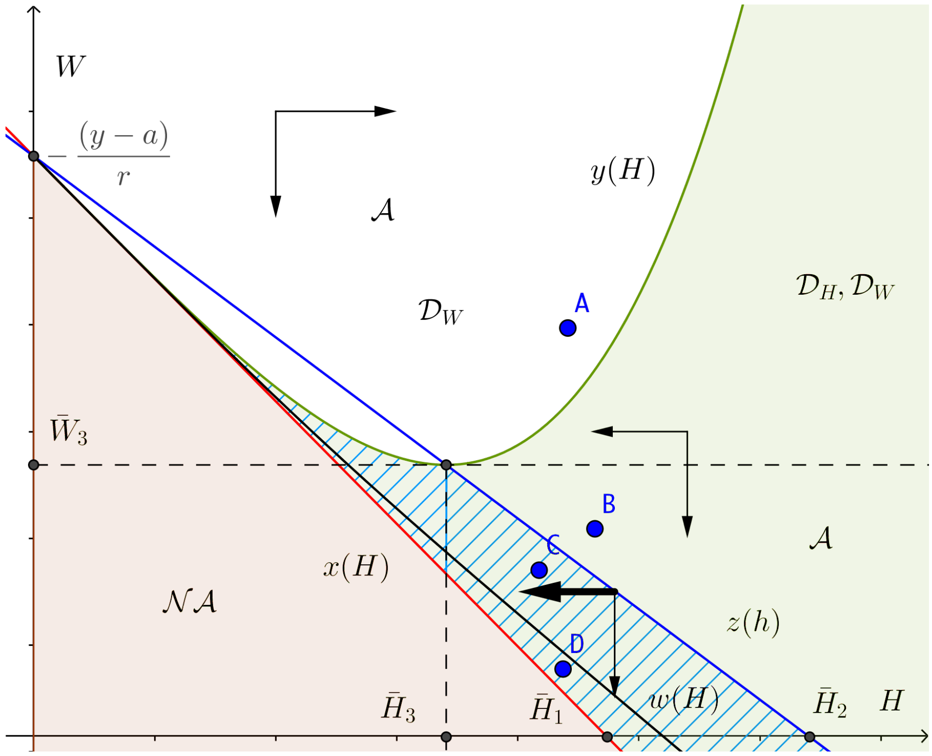
$$\begin{aligned}
\mathcal{D}_H &= \left\{(W,H) \in \mathcal{A} : \frac{1}{\mathrm{d}t}E_{t-}[\mathrm{d}H_t \mid W_{t-} = W, H_{t-} = H] < 0\right\} \\
&= \left\{(W,H) \in \mathcal{A} : W < y(H) = x(H) + DH^{1+\xi_m}\right\}, \qquad (13a)\\
\mathcal{AC} &= \left\{(W,H) \in \mathcal{D}_H : I_H^h(W,H) > 0\right\}, \\
&= \left\{(W,H) \in \mathcal{D}_H : W < \min\left[y(H), z(H) = x(H) + \frac{BH}{1+\xi_m}\right]\right\}, \qquad (13b)\\
\mathcal{D}_W &= \left\{(W,H) \in \mathcal{A} : \frac{1}{\mathrm{d}t}E_{t-}[\mathrm{d}W_t \mid W_{t-} = W, H_{t-} = H] < 0\right\} \\
&= \left\{(W,H) \in \mathcal{A} : W > w(H) = \frac{x(H)[l(H)+r]}{l(H)} + \frac{k(H)}{l(H)}\right\}, \qquad (13c)
\end{aligned}$$

where

$$\begin{aligned}
D &= \mathcal{I}_1^{-1}\left[\tilde{\delta}^{1/\alpha} - BK\right] > 0 \\
l(H) &= A - \frac{\theta^2}{\gamma} - r + (\mathcal{I}_1 + \mathcal{C}_1)\, H^{-\xi_m} > 0 \\
k(H) &= y_0 - a + H(\beta - KB).
\end{aligned}$$

The depletion and accelerating dynamics (13) can be analyzed through the local phase diagram in Figure 1. First, the admissible region $\mathcal{A}$ is bounded below by the red $x(H)$ locus (10), with complementary non-admissible area $\mathcal{NA}$ in shaded red region. The horizontal intercept of $x(H)$ is the capitalized base income deficit $-(y_0 - a)/r > 0$ under (11), whereas its vertical intercept is $\bar{H}_1 = -(y_0 - a)/(rB) > 0$. From (9), the red $x(H)$ locus is characterized by zero net total wealth, consumption at subsistence level and indifference between life and death, with the negative slope suggesting a corresponding tradeoff between health and wealth.[19]

**Figure 1:** Joint health and wealth dynamics




*Notes:* Non-admissible set $\mathcal{NA}$: shaded red area under red $x(H)$ line, admissible $\mathcal{A}$ is area above $x(H)$. 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.

Second, the health depletion region $\mathcal{D}_H$ is the shaded green area located below the green $y(H)$ locus in (13a), a U-shaped function bounded below by $\bar{W}_3 = y(\bar{H}_3)$ where:

$$\bar{H}_3 = \left(\frac{B}{D(1+\xi_m)}\right)^{\frac{1}{\xi_m}} > 0. \tag{14}$$

The reasons for the non-monotonicity stem from the effects of health on $I^h(W, H) = I(W, H)/H$. At low health $H < \bar{H}_3$, better health raises the value of the health capital $BH$ and therefore net total wealth $N_0(W, H)$, thereby increasing the investment to capital ratio $I^h$. Constant (and zero) growth thus requires an offsetting reduction in $W$. At high health $H > \bar{H}_3$, being healthier lowers the incentives for investing to control for mortality risk and therefore reduces $I^h$. Constant growth requires increasing $W$.

Third, the accelerating locus $z(H)$ in (13b) is plotted as the blue line in Figure 1; the accelerating region is the dashed blue subset of $\mathcal{D}_H$. Appendix B shows that this locus intersects the $x(H), y(H)$ loci at the same $-(y_0 - a)/r$ 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 for all health levels. Moreover, it also intersects the health depletion locus $y(H)$ at lower bound $\bar{H}_3$ in (14). Consequently, there exists a threshold wealth level $\bar{W}_3 = y(\bar{H}_3)$ below which all agents expect a health decline, and a threshold health level $\bar{H}_3$ below which all agents in the depletion region are also in the accelerating subset.

Fourth, the wealth depletion locus $w(H)$ in (13c) is represented as the black curve in Figure 1 whereby the wealth depletion region $\mathcal{D}_W$ is the area above this locus. Appendix B establishes that this locus has the same $H-$ intercept $-(y_0 - a)/r$ and it must lie above the admissibility locus $x(H)$. 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 non-empty for every $H$ under sufficient conditions (12), i.e. there exists an admissible range of $W$ for which agents optimally expect both their health and their wealth to fall.

The local expected dynamics of health are represented by the horizontal (health) and vertical (wealth) arrows in Figure 1 with agents $j$ = A, B, C and D described by their $(H_j, W_j)$ statuses. First, agent A is sufficiently rich (i.e. $W > y(H)$) and can expect a growth in health towards the steady-state locus $y(H)$, e.g. following a morbidity shock. Agent B is poorer and is located in the $\mathcal{D}_H$ region in which the health stock is expected to fall, yet is nonetheless sufficiently rich and healthy ($W > z(H)$) to optimally slow down – but not reverse – the depreciation of his health capital (i.e. $I_H^h < 0$). However, for agents C and D, wealth is below the $z(H)$ locus such that the health depletion accelerates (i.e. $I_H^h > 0$, illustrated by the thick vector) as falling health is accompanied by further cuts in the investment-to-health ratio. All three agents A, B and C expect their wealth to fall, whereas agent D, is located at very low wealth levels in the $\mathcal{AC}$ region where rapidly receding health expenses $I(W, H)$ allow for expected increases in wealth.

## 2.4 Discussion

### 2.4.1 Closing Down the Shop

These joint end-of-life dynamics of health and wealth are consistent with a deliberate closing down the shop strategy when the conditions in Propositions 1, 2, and 3 are satisfied. Sufficiently rich ($W > y(H)$) and healthy agents reinvest in their health, with the latter returning to the steady-state locus $y(H)$ following a sickness shock. However, falling wealth is also optimally chosen, leading agents to eventually enter the $\mathcal{D}_H$ region where health depletion and increasing mortality risks are optimally selected.[20] Such dynamics are consistent with strong positive SES gradients for health outcomes, as well as for longevity.[21] Furthermore, they are also consistent with the inverted U shape in the life cycle of the SES gradients that peak after middle age and fall in the last period of life.[22] Indeed, the model predicts that rich agents ($W > \bar{W}_3$) initially prevent health declines but that all ultimately enter the health depletion region after which wealth is less relevant with respect to health outcomes.

Moreover, the depreciation of the health stock accelerates once falling health and wealth draws agents into the $\mathcal{AC}$ region. Our model thus supports threshold effects whereby falling health is initially slowed down and then accelerated for $H < \bar{H}_3$ and is thus pro-factual with the accelerating deterioration in both health and longevity that is observed after age 70.[23] From the endogenous death intensity (2), 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, subsistence consumption 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 (5) ensure strictly positive continuation utility under life in (9). The agents we are considering therefore have no proclivity in favor of premature death when they deliberately initiate closing down strategies. Finally, as discussed earlier, endogenous mortality, increases investment ($\mathcal{I}_1 > 0$), as well as consumption of non-durables and services – including comfort care – if the elasticity of inter-temporal substitution is high ($\varepsilon > 1 \implies \mathcal{C}_1 > 0$). The empirical results below are consistent with $0 < \mathcal{I}_1 < \mathcal{C}_1$, i.e. higher mortality caused by falling health induces agents to shift in favor of more $C_t$ than $I_t$, consistent with an end-of-life change in the composition of health expenses towards more comfort care than curative care.[24]

### 2.4.2 Comparison with the standard model

Our model nests the seminal Grossman (1972); Ehrlich and Chuma (1990) framework under restriction (6). In particular, these restrictions abstract from endogenous mortality $\lambda_{m1} = 0$, leading to $\mathcal{V}_1, \mathcal{I}_1, \mathcal{X}_1, \mathcal{C}_1 = 0$ in the optimal rules. Appendix B.3 derives the corresponding dynamics reproduced in Figure 2. Under sufficient condition (12a), the health depletion region is the entire admissible set, whereas no accelerating region exists.

The standard model predicts common and constant depletion rates for health, that are independent of wealth or health statuses, that do not accelerate near the end of life and that have no incidence on death risk which remains counter-factually independent of age and of financial and health levels. This contradicts the evidence of strong positive SES gradients and on the life cycle of these gradients that peak at mid-life and fall thereafter.

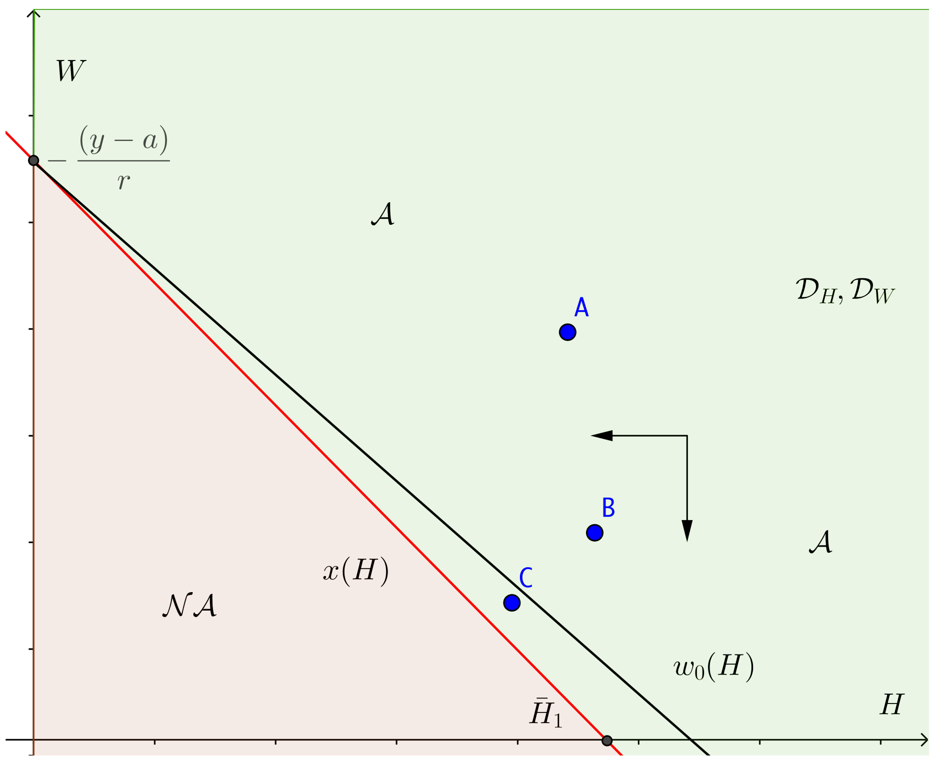
### 2.4.3 Aging

The model can be modified to account for realistic aging processes involving age-increasing depreciation, sickness and death risks exposure:

$$\dot{\delta}_t, \dot{\phi}_t, \dot{\lambda}_{s0t}, \dot{\lambda}_{m0t} \geq 0. \tag{15}$$

In that perspective, Hugonnier et al. (2013) show that the optimal rules in Theorem 1 remain valid, although with age-dependent parameters that can be solved in closed form. The predicted loci remain valid and inherit age-dependency: $x_t(H), y_t(H), z_t(H), w_t(H)$. In addition to making it more likely that sufficient conditions (12) are met, it can be shown

**Figure 2:** Joint health and wealth dynamics: Exogenous mortality



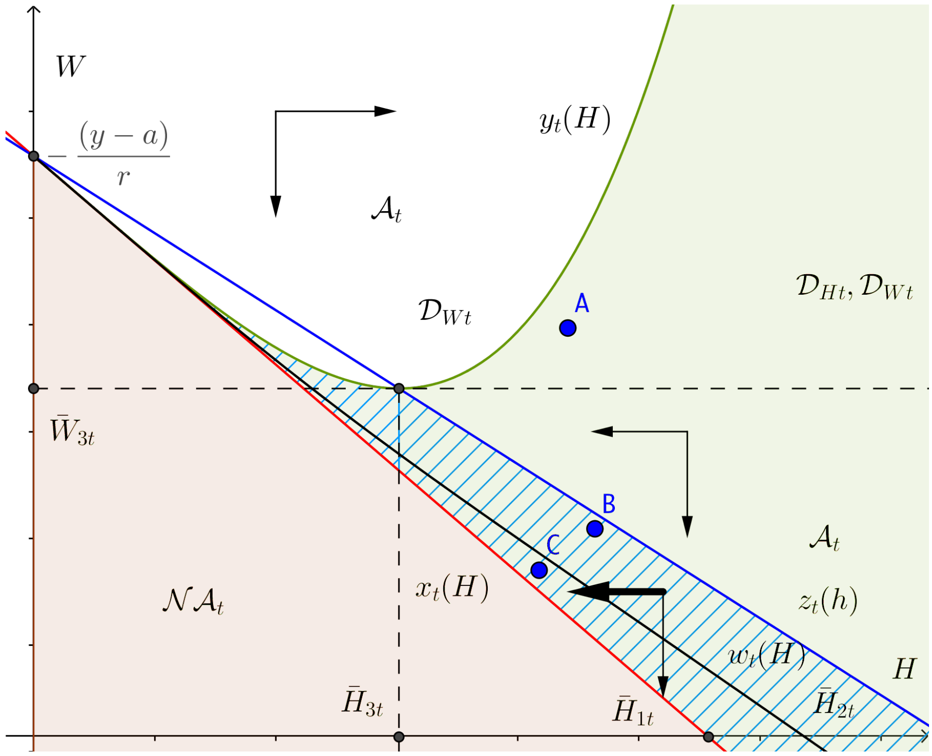

*Notes:* Effects of exogenous mortality $\lambda_{m1} = 0$. Non-admissible set $\mathcal{NA}$: shaded red area under red $x(H)$ line, admissible $\mathcal{A}$ is area above $x(H)$. Health depletion set $\mathcal{D}_H$: shaded green area under green $y(H)$ green curve. Wealth depletion set $\mathcal{D}_W$: area above $w(H)$ black curve.

that the aging process (15) generates counter-clockwise rotations in both $x_t(H), y_t(H)$ loci entails that all agents are now closer to $\mathcal{D}_H$ (see Figure 3).[25] We conclude that aging is complementary to and reinforces the closing-down process.

# 3 Empirical evaluation

The optimal health and wealth depletion strategy is arguably more appropriate for agents nearing death, than for younger ones. Indeed, a high health depreciation ($\delta$), sickness likelihood ($\lambda_{s0}$) and/or consequence ($\phi$) all seem legitimate for old agents in the last period of life, yet less so for younger ones. Moreover, a high marginal propensity to consume in (12c), potentially stemming from non-curative LTC expenses, is suitable for elders nearing the end of life. Using a database of relatively old individuals (HRS-CAMS),

**Figure 3:** Joint health and wealth dynamics: Effects of aging




*Notes:* Effects of aging process (15). Non-admissible set $\mathcal{NA}_t$: shaded red area under red $x_t(H)$ line, admissible $\mathcal{A}_t$ is area above $x_t(H)$. Health depletion set $\mathcal{D}_{Ht}$: shaded green area under green $y_t(H)$ green curve. Acceleration set $\mathcal{AC}_t$: hatched green area under blue $z_t(H)$ curve. Wealth depletion set $\mathcal{D}_{Wt}$: area above $w_t(H)$ black curve.

we next verify empirically whether or not these conditions are valid and whether the admissible, depletion and acceleration subsets have economic relevance.

## 3.1 Econometric model

Assuming that agents $j = 1, 2, \cdots, N$ follow the optimal rules in Theorem 1, we consider a tri-variate nonlinear structural econometric model defined by the optimal investment (8b), the consumption equation (8d), and the risky asset holdings (8e):

$$I_j = KBH_j + \mathcal{I}_1 H_j^{-\xi_m} N_0(W_j, H_j) + u_j^I, \tag{16a}$$

$$C_j = a + \left(A + \mathcal{C}_1 H_j^{-\xi_m}\right) N_0(W_j, H_j) + u_j^C \tag{16b}$$

$$\Pi_j = (\theta/(\gamma\sigma_S))\, N_0(W_j, H_j) + u_j^\Pi, \tag{16c}$$

where $N_0(W, H)$ denotes net total wealth in (7), the parameters $(K, B, \mathcal{I}_1, A, \mathcal{C}_1)$ are outlined in Appendix A and where $(u_j^I, u_j^C, u_j^\Pi)$ are correlated error terms. Optimal insurance (8c) is omitted from our specification under near-universal Medicare coverage for elders.

A subset of the technological, distributional and preference parameters are estimated using the joint system (16), imposing the regularity conditions (17) (see Appendix A). Due to significant non-linearities, not all the parameters can be identified. We calibrate certain parameters (i.e. $\mu, r, \sigma_S, \rho$) with standard values from the literature. For others however (i.e. $\phi, \gamma_m, \gamma_s$), scant information is available and we rely on a thorough robustness analysis.

The estimation approach is an iterative two-step ML 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 latter are fixed and the likelihood function is maximized with respect to $\xi_m$. The procedure is iterated until a fixed point is reached for all the estimated structural parameters. The likelihood function is written by assuming that there exist some cross-correlations between the three equations, i.e. $\text{Cov}(u_j^I, u_j^C, u_j^\Pi) \neq 0$. 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).

The database is the 2002 wave of the Health and Retirement Study (HRS, Rand data files) corresponding to the last HRS wave with detailed information on health spending. The HRS data set is merged with the 2001 Consumption and Activities Mail Survey (CAMS) for observable data corresponding to consumption equation (8d) (see Appendix C for details). A main advantage of CAMS data is that its detailed categories allow us to distinguish between curative and comfort care expenses (e.g. home health care or dental visits, ...) that can reasonably be considered as consumption rather than investment. For consistency with our model of end-of-life health and wealth dynamics, with bequest motives abstracted from, we restrict our analysis to elders (i.e. agents aged 65 and more), who are singles, and with positive financial wealth (1,124 remaining observations).

We report the sample statistics in Table 1, while Table 2 reports the median values stratified by wealth quintiles and self-reported health. Consistent with empirical evidence, financial wealth seems to be relatively insensitive to health,[26] health investment increases slowly in wealth, but falls sharply in health,[27] whereas risky asset holdings are higher for healthier and wealthier agents.[28] Consumption is highest for rich, less healthy agents.

**Table 1:** HRS-CAMS data statistics

| | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|
| Consumption ($C$) | 16 507 | 17 765 | 0 | 217 510 |
| Wealth ($W$) | 79 423 | 164 837 | 1 | 1 675 001 |
| Investment ($I$) | 1 959 | 2 978 | 0 | 36 049 |
| Risky holdings ($\Pi$) | 38 631 | 116 958 | 0 | 1 500 000 |
| Health ($H$) | 1.84 | 0.82 | 0.5 | 3.5 |
| Age ($t$) | 75.92 | 6.99 | 65 | 97 |

*Notes:* Statistics for HRS-CAMS data (in 2002 $ for nominal variables) used in estimation. Scaling for self-reported health is 0.5 (Poor), 1.25 (Fair), 2.00 (Good), 2.75 (Very good) and 3.5 (Excellent).

# 4 Results

## 4.1 Structural parameters

Table 3 reports the calibrated and estimated deep parameters (panels a–d), the induced parameters to define the various subsets (panel e), as well as the sufficient conditions that are relevant to Propositions 1, and 2 (panel f). The standard errors indicate that all the estimates are precisely estimated and are significant at the 5% level.

First, the law of motion parameters in panel a are consistent with significant diminishing returns to the health production function ($\alpha = 0.73$). Depreciation is important ($\delta = 4.6\%$) and sickness is consequential, with an additional depreciation of $\phi = 1.1\%$

**Table 2:** HRS-CAMS data statistics (cont'd)

| Variable | Wealth quintile | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| | a. Poor health ($H = 0.5$) | | | | |
| Consumption ($C$) | 11 708 | 10 563 | 20 881 | 30 056 | 35 841 |
| Wealth ($W$) | 31 | 1 459 | 10 481 | 55 492 | 308 243 |
| Investment ($I$) | 26 379 | 22 792 | 25 127 | 32 048 | 26 948 |
| Risky share ($\Pi/W$) | 0.00 | 0.00 | 0.25 | 0.63 | 0.86 |
| | b. Fair health ($H = 1.25$) | | | | |
| Consumption ($C$) | 15 672 | 13 090 | 13 319 | 19 237 | 24 722 |
| Wealth ($W$) | 29 | 2 008 | 12 469 | 51 375 | 237 487 |
| Investment ($I$) | 26 162 | 14 850 | 10 459 | 18 728 | 24 242 |
| Risky share ($\Pi/W$) | 0.00 | 0.02 | 0.26 | 0.45 | 0.62 |
| | c. Good health ($H = 2.0$) | | | | |
| Consumption ($C$) | 14 873 | 14 603 | 14 124 | 18 426 | 23 190 |
| Wealth ($W$) | 34 | 1 915 | 14 300 | 54 334 | 300 252 |
| Investment ($I$) | 10 637 | 16 420 | 11 592 | 9 749 | 14 965 |
| Risky share ($\Pi/W$) | 0.00 | 0.01 | 0.27 | 0.40 | 0.83 |
| | d. Very good health ($H = 2.75$) | | | | |
| Consumption ($C$) | 13 255 | 12 705 | 15 181 | 17 948 | 20 585 |
| Wealth ($W$) | 34 | 2 142 | 14 198 | 51 266 | 306 920 |
| Investment ($I$) | 4 768 | 21 220 | 6 876 | 8 060 | 10 285 |
| Risky share ($\Pi/W$) | 0.03 | 0.04 | 0.19 | 0.41 | 0.72 |
| | e. Excellent health ($H = 3.5$) | | | | |
| Consumption ($C$) | 12 418 | 11 749 | 15 340 | 18 947 | 21 254 |
| Wealth ($W$) | 68 | 2 114 | 12 679 | 60 140 | 358 548 |
| Investment ($I$) | 2 456 | 5 159 | 7 199 | 8 079 | 5 593 |
| Risky share ($\Pi/W$) | 0.00 | 0.00 | 0.32 | 0.50 | 0.82 |

*Notes:* Mean values (in 2002 $ for nominal variables) per health status and wealth quintiles for HRS-CAMS data used in estimation.

suffered upon realization of the health shock. The intensity parameters in panel b indicate a high and significant likelihood of health shocks ($\lambda_{s0} = 0.08$). The death intensity (2) parameters reject the null of exogenous exposure to death risk ($\lambda_{m1}, \xi_m \neq 0$), validating

**Table 3:** Estimated and calibrated parameter values

| Parameter | Value | Parameter | Value | Parameter | Value |
|---|---|---|---|---|---|
| a. Law of motion health (1) | | | | | |
| $\alpha$ | 0.7285* | $\delta$ | 0.0460* | $\phi$ | $0.011^c$ |
| | (0.2066) | | (0.0148) | | |
| b. Sickness and death intensities (2) | | | | | |
| $\lambda_{s0}$ | 0.0813* | $\lambda_{m0}$ | 0.0665* | | |
| | (0.0233) | | (0.0299) | | |
| $\lambda_{m1}$ | 0.0219* | $\xi_m$ | 2.2498* | | |
| | (0.0074) | | (1.0845) | | |
| c. Income and wealth (3) | | | | | |
| $y_0$ | $0.0085^{c\$}$ | $r$ | $0.048^c$ | | |
| $\mu$ | $0.108^c$ | $\sigma_S$ | $0.20^c$ | | |
| d. Preferences (5) | | | | | |
| $a$ | $0.0126^{*\$}$ | $\beta$ | 0.0091* | | |
| | (0.0064) | | (0.0023) | | |
| $\varepsilon$ | 1.7364* | $\gamma$ | 2.5721* | | |
| | (0.4192) | | (1.1925) | | |
| $\rho$ | $0.025^c$ | $\gamma_m$ | $0.75^c$ | $\gamma_s$ | $7.40^c$ |
| e. State space subsets (10), (24a), (25b), (30b) | | | | | |
| $(y_0 - a)/r$ | $-0.0836^{*\$}$ | $B$ | 0.0966* | $\bar{H}_1$ | 0.8661* |
| $D$ | 0.1674* | $\mathcal{I}_1$ | 0.1572* | $K$ | $5.88e-04^*$ |
| $\bar{H}_3$ | 0.4692* | $\bar{W}_3$ | $0.0523^{*\$}$ | $\bar{H}_2$ | 1.1193* |
| $\mathcal{C}_1$ | 2.5298* | $A$ | 0.1065* | | |
| f. Sufficient conditions (12) (must be negative) | | | | | |
| $\beta - \tilde{\delta}^{1/\alpha}$ | $-0.0059^*$ | $\theta^2/\gamma + r - A$ | $-0.0235^*$ | | |

*Notes:* Econometric model (16). *: Estimated structural and induced parameters (standard errors in parentheses), significant at 5% level; $c$: calibrated parameters; \$: In \$M.

the assumption that agent's health decisions are consequential for their expected life horizon.

Third, the returns parameters $(\mu, r, \sigma_S)$ are calibrated at standard values in panel c. The base income $y_0$ in equation (3b) is calibrated to a value of \$8,500 in 2002 dollars (\$11,927 in 2018). Fourth, the preference parameters in panel d suggest a significant subsistence consumption $a$ of \$12,600 (\$17,530 in 2018), which is larger than base income $y_0$. Both subsistence and base income values are realistic.[29] Our estimate of the inter-temporal elasticity $\varepsilon = 1.74$ is larger than one, as required for sufficient condition (12b) and as identified by others using micro data.[30] Aversion to financial risk is realistic ($\gamma = 2.57$), 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 and calibrated structural parameters are economically plausible.

## 4.2 Induced parameters and relevance of closing down

Table 3.e reports the induced parameters that are relevant for the admissible, depletion and accelerating subsets. Table 3.f shows that the sufficient conditions (12a) and (12c) 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 4 and to plot the corresponding subsets in Figure 4 using the same scaling as the one for the estimation. Finally, we can rely on the joint distribution in Table 2 in order to plot the quintile values of wealth as blue dots for $\mathrm{Q}_i$ 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 83,630\$ in 2002 dollars and confirms that condition (11) is verified. Second, we identify a relatively large marginal-$Q$ of health $B = 0.0966$ in panel e, suggesting that health depletion can remain optimal despite health being very valuable.[31] Third, the value for $D$ in Table 3.e is large and significant. From the definition of $y(H)$ in (13a), a large value of $D$ also entails a steep health depletion locus in Figure 4. It follows that its minimum is attained at a low $\bar{H}_3 = 0.4692$, with corresponding realistic value of $\bar{W}_3 = \$52{,}262$. Since this value is larger than most observed wealth levels (see Tables 1 and 2), it follows that the bulk of the population is located in the health depletion subset. Fourth, our estimates are consistent with a narrow accelerating region $\mathcal{AC}$. 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.

Table 4: Estimated values of loci

| Level | $H$ | % Pop. | $\mathcal{A}$ $x(H)$ | $\mathcal{D}_H$ $y(H)$ | $\mathcal{AC}$ $z(H)$ | $\mathcal{D}_W$ $w(H)$ |
|---|---|---|---|---|---|---|
| Poor | 0.50 | 12.74 | 0.04 | 0.05 | 0.05 | 0.04 |
| Fair | 1.25 | 28.16 | −0.04 | 0.31 | −0.00 | −0.03 |
| Good | 2.00 | 33.77 | −0.11 | 1.48 | −0.05 | −0.08 |
| Very good | 2.75 | 18.90 | −0.18 | 4.30 | −0.10 | −0.11 |
| Excellent | 3.50 | 6.43 | −0.25 | 9.56 | −0.15 | −0.12 |

*Notes:* Values (in M$) of admissible $\mathcal{A} : W \geq x(H)$; health depletion $\mathcal{D}_H : W < y(H)$; accelerating $\mathcal{AC} : W < \min[y(H), z(H)]$; and wealth depletion $\mathcal{D}_W : W > w(H)$ at observed health levels.

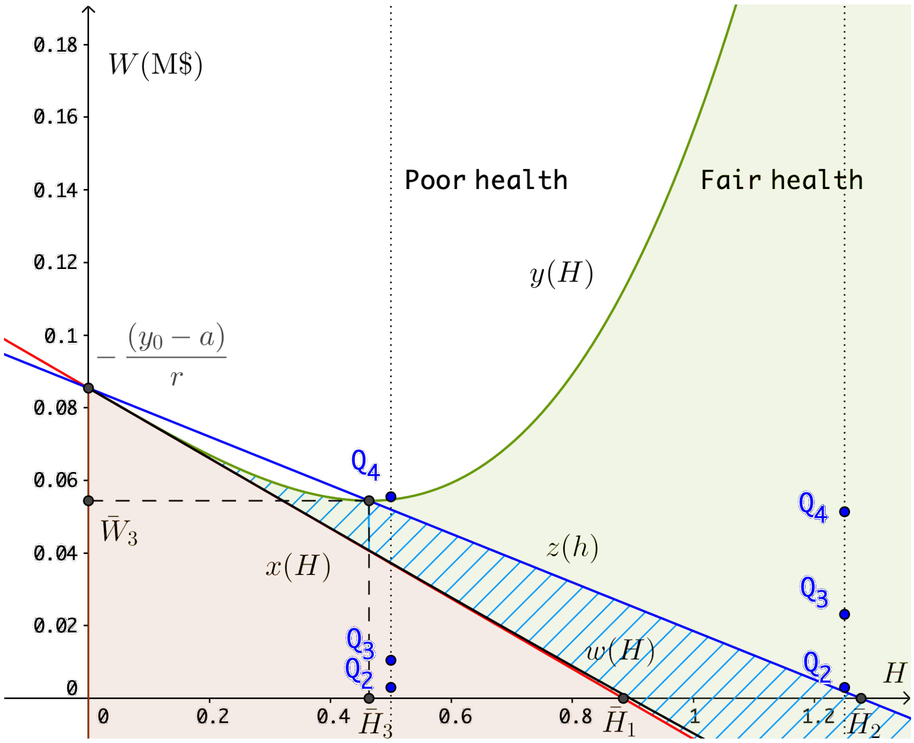
between Fair and Poor self-reported health) and close to one another (less than one discrete increment of 0.75). This feature of the model is reassuring since we would expect accelerating phases where agents are cutting down expenses in the face of falling health to coincide with the very last periods of life where health is very low. Fifth, our finding of $0 < \mathcal{I}_1 < \mathcal{C}_1$ is consistent with stronger positive effects of increasing endogenous mortality on consumption – including comfort care – than on investment, i.e. curative care (see discussion in Section 2.4.1).

Finally, the estimated wealth depletion locus $w(H)$ is lying between the $x(H)$ and $y(H)$ loci (see Proposition 2). 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 4 makes clear, 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 unambiguously confirm the empirical relevance of optimal closing-down strategies.

## 4.3 Simulation analysis

The analysis presented thus far has abstracted from exogenous depletion processes associated with aging and has focused upon optimal local expected changes for health and wealth. 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

**Figure 4:** Estimated depletion, accelerating and non-admissible regions

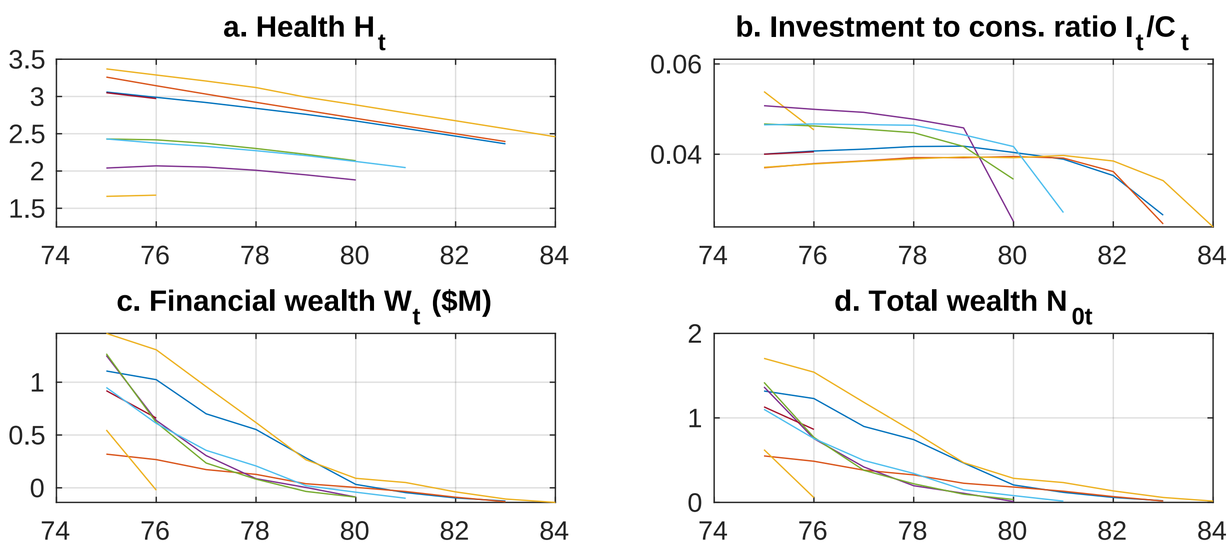



*Notes:* Non-admissible set $\mathcal{NA}$: shaded red area under red $x(H)$ line. Health depletion set $\mathcal{D}_H$: shaded green area under green $y(H)$ green curve. Acceleration set $\mathcal{AC}$: hatched green area under blue $z(H)$ curve. Wealth depletion set $\mathcal{D}_W$: area above $w(H)$ black curve. Position of loci and areas evaluated at estimated parameters in Table 3. Quintile levels for wealth quintiles $Q_2, \ldots Q_4$ are taken from Table 2 and are reported as blue points for poor and fair health levels.

described in further details in Appendix D. To summarize, the simulated life cycles draw initial health and wealth statuses from uniform distributions at age 75 and for a sample of 1,000 individuals. For each period, a common financial market shock is then drawn, whereas agent-specific sickness and death shocks are drawn using the corresponding exogenous morbidity and the endogenous mortality Poisson distributions. The health and wealth statuses are updated using (1) and (3a), evaluated at the theoretical optimal rules in Theorem 1; the process is replicated for 500 times.

Figure 5 plots a random sample of the 500,000 simulated optimal trajectories for health level $H_t$ (panel a), the investment-to-consumption ratio $I_t/C_t$ (panel b), financial wealth $W_t$ (panel c), as well as the net total wealth $N_0(W_t, H_t)$ (panel d). A common color across each panel corresponds to a common individual path. Curtailed paths indicate death and/or non-admissibility (i.e. preference for death over life). Unsurprisingly, these results confirm all our previous findings. Consistent with the data, our simulated life cycles feature 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., 2015; French et al., 2006; De Nardi et al., 2010, 2009). Indeed, the optimal strategy is to bring down net total wealth $N_0(W_t, H_t)$ in panel d to zero (i.e. reach the lower limits of admissible set $\mathcal{A}$) at terminal age at which stage agents are indifferent between life and death. This objective is attained by running down wealth very rapidly in panel c (consistent with our finding of low $w(H)$ locus) and a somewhat slower decline for (utility services-providing) health in panel a. The health decline is achieved by curtailing investment in favor of consumption (including comfort care) spending in panel b, consistent with our empirical finding of $\mathcal{C}_1 > \mathcal{I}_1$. These pro-factual life cycle profiles confirm that the Closing Down model can reproduce the data even without the self-reinforcing incidence of biological aging.

**Figure 5:** Simulated optimal trajectories




*Notes:* Sample of simulated trajectories from 1,000 individuals × 500 replications. Each color corresponding to a common admissible policies for surviving individual. Curtailed paths correspond to death and/or non-admissibility. See Appendix D for details.

Contrasting individual paths reveals that, as expected, health (panel a) and wealth (panel c) depletion are both faster for the poor and unhealthy agents. The joint health and wealth depletion means that the latter approach the non-admissible subset more rapidly. Moreover, worse health entails that exposure to death risk is higher for the poor, resulting in lower survivorship, consistent with stylized facts (Bosworth et al., 2016). 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. High initial wealth has a moderating effect on the speed of the depletion, but not on its ultimate outcome.

We conclude by emphasizing endogenous mortality as a key element in reproducing the end-of-life dynamic and cross-sectional evidence. Reinstating realistic aging processes makes our optimal dynamic strategies even more relevant. Put differently, aging is not a substitute to, but is a reinforcing complement to closing-down.

# Notes

[1]See Banks et al. (2015, Fig. 5, p. 12), Heiss (2011, Fig. 2, p. 124), Smith (2007, Fig. 1, p. 740), Case and Deaton (2005, Fig. 6.1, p. 186), or Van Kippersluis et al. (2009, Figs. 1, 2, p. 820, 823, 824) for evidence.

[2]Smith (2007, Tabs. 1–3, pp. 747–752).

[3]See Benjamins et al. (2004); Heiss (2011); Smith (2007); Hurd et al. (2001); Hurd (2002) for evidence and discussion. See also Arias (2014, Tab. B, p. 4) for Life Tables.

[4]Smith (2007, 1999); Attanasio and Emmerson (2003).

[5]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). See also Attanasio and Emmerson (2003) for mortality-SES gradient evidence.

[6]Van Kippersluis et al. (2009); Baeten et al. (2013); Case and Deaton (2005).

[7]Whereas curative expenses (e.g. doctor visits, hospital stays, drugs, ...) tend to stagnate, nursing homes and other long-term care (LTC) spending increase sharply (De Nardi et al. (2015, Fig. 3, p. 22)). LTC expenditures are more income- and wealth-elastic than curative care and can be associated with comfort care consumption (De Nardi et al. (2015); Tsai (2015); Marshall et al. (2010)). In addition, LTC expenses are not covered by Medicare and are rarely insured against through private markets.

[8]De Nardi et al. (2016, 2015); Marshall et al. (2010); Love et al. (2009); French et al. (2006); Palumbo (1999)

[9]See Grossman (2000); Galama (2015) for reviews. We also provide further details on these shortcomings for the base Grossman (1972) model with analytical solutions discussed in Section 2.4.2 and Figure 2, with formal arguments made in Theorem 4.

[10]Other exceptions with endogenous mortality include Pelgrin and St-Amour (2016); Kuhn et al. (2015); Dalgaard and Strulik (2014, 2017); Blau and Gilleskie (2008); Hall and Jones (2007). However,

none of these papers focus on end-of-life joint dynamics for health and wealth. The closest paper is Galama (2015) who also emphasizes the shortcomings of the canonical Grossman (1972) model as motivation for generalizations, while stressing the importance of combining investment with the law of motion for health to characterize the optimal (equilibrium) time paths of the latter, and to generate declining health over the life cycle. The main differences are stochastic mortality, morbidity and financial processes, more general preferences, as well as structural estimation that are abstracted from in Galama (2015).

[11]The positive investment restriction is necessary since $I_t$ is to the power $\alpha \in (0,1)$ in the Cobb-Douglas technology (1). Moreover, it is required to prevent degenerate cases where the agent could hasten the death timing by investing negative amounts, inconsitstent with both the monetary expense interpretation and the preference for life over death assumption.

[12]Hugonnier et al. (2013) also consider a more general setup with health-dependent sickness intensity which we abstract from.

[13]Our analysis focuses on financial wealth only and omits housing, a non-negligible determinant of disposable resources for elders. Unlike financial wealth, housing provides direct utilitarian service flows, involves additional life-cycle decisions to those we consider, as well as complex budget constraint considerations (e.g. leverage effects, (non) inclusion in means-tested programs, ...) that would have to be modeled for completeness. Moreover, a simpler approach of adding net housing to financial wealth yielded similar empirical results. For these two reasons, we thus maintain our current perspective on financial assets only, and we prefer to leave housing on the resarch agenda.

[14]In the empirical application of Section 3, we thus include elements such as long-term care, personal health care and dental visits in $C$. See Appendix C for details.

[15]The model also admits an alternative interpretation where preferences $f(C, H, U)$ in (5b) are health-independent, and with health-increasing income (3b) replaced by $Y_t = y_0 + \beta H_{t-}$ (Hugonnier et al., 2013, Remark 3). The theoretical results are unaffected by this change in perspective.

[16]In particular, a change of variable from $I$ to $I^g, I^h$ reveals the following equivalent formulations:

| New control in (4) | $(I/H)^\alpha$ in (1) | $I$ in (3a) |
|---|---|---|
| $I^g \equiv I^\alpha H^{1-\alpha}$ | $(I^g/H)$ | $(I^g/H)^{1/\alpha} H$ |
| $I^h \equiv (I/H)$ | $(I^h)^\alpha$ | $I^h H$ |

The optimal dynamics we recover under either $I^g$ or $I^h$ are identical to the ones we obtain in the original formulation with $I$, provided both (1) and (3a) are modified accordingly. The linear technology in $I^g$ in $\mathrm{d}H_t$ is resorted to by Grossman (1972), whereas Ehrlich and Chuma (1990) add convex costs as in $(I^g)^{1/\alpha}$. A concave technology $(I^h)^\alpha$ formulation is used notably by Galama (2015).

[17]Although not necessary for the main theoretical results, restriction (11) is also tested and confirmed empirically in Section 4 and will be relied upon in the discussion of these results. Moreover, it helps ensure that portfolio shares $\Pi^*/W$are increasing in wealth, consistent with the data (e.g. Wachter and Yogo, 2010).

[18]See also Laporte and Ferguson (2007) for an analysis of expected local changes of the Grossman (1972) model with Poisson shocks.

[19]See also Finkelstein et al. (2013, 2009) for evidence and discussion regarding health effects on marginal utility of wealth.

[20]It is worth noting that the optimal risky asset holdings in (8e) are positive when net total wealth and risk premia are both positive. Moreover, the investment in (8b) is monotone increasing in wealth, such that a sufficiently long sequence of high positive returns on financial wealth could 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. See however the simulation exercise in Section 4.3 for realistic life cycle patterns consistent with depletion.

[21]See Smith (2007, 1999) for SES gradients on health outcomes and (Bosworth et al., 2016) for longevity.

[22]See Case and Deaton (2005); Van Kippersluis et al. (2009); Baeten et al. (2013) for evidence of diminishing SES gradients over the life cycle.

[23]See for Banks et al. (2015); Van Kippersluis et al. (2009) health deterioration and Arias (2014) for longevity.

[24]See De Nardi et al. (2015); Marshall et al. (2010) for evidence and discussion.

[25]In particular, the aging process (15) yields age-decreasing $\dot{B}_t, \dot{K}_t, \dot{L}_{mt}, \dot{\mathcal{I}}_{1t} \leq 0$ and age-increasing $\dot{D}_t \geq 0$. The combination of the two entails counter-clockwise rotations in all the loci, with common intercept $(y - a)/r$ unaffected.

[26]See Hugonnier et al. (2013); Michaud and van Soest (2008); Meer et al. (2003); Adams et al. (2003) for additional evidence.

[27]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). In addition, consumption is highest for rich, unhealthy agents and health (e.g. Hugonnier et al., 2013; Smith, 1999; Gilleskie and Mroz, 2004; Yogo, 2009) have been discussed elsewhere.

[28]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 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).

[29]For example, the 2002 poverty threshold for elders above 65 was \$8,628 (source: U.S. Census Bureau).

[30]For example, Gruber (2013) finds estimates centered around 2.0, relying on CEX data.

[31]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 $w_h(0.10, W, H) = 0.10. * B * 10^6 = \$9{,}656$.

# References


Acemoglu, D., Finkelstein, A., and Notowidigdo, M. J. (2013). Income and health spending: Evidence from oil price shocks. *Review of Economics and Statistics*, 95(4):1079 – 1095.

Adams, P., Hurd, M. D., McFadden, D., Merrill, A., and Ribeiro, T. (2003). Healthy, wealthy, and wise? tests for direct causal paths between health and socioeconomic status. *Journal of Econometrics*, 112(1):3–56.

Arias, E. (2014). United States life tables, 2010. *National Vital Statistics Report*, 63(7):1–62.

Attanasio, O. P. and Emmerson, C. (2003). Mortality, health status, and wealth. *Journal of the European Economic Association*, 1(4):821–850.

Baeten, S., Ourti, T. V., and Doorslaer, E. V. (2013). The socioeconomic health gradient across the life cycle: what role for selective mortality and institutionalization? *Social Science and Medicine*, 97:1–20.

Banks, J., Blundell, R., Levell, P., and Smith, J. P. (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, London.

Benjamins, M. R., Hummer, R. A., Eberstein, I. W., and Nam, C. B. (2004). Self-reported health and adult mortality risk: An analysis of cause-specific mortality. *Social Science & Medicine*, 59(6):1297 – 1306.

Berkowitz, M. K. and Qiu, J. (2006). A further look at household portfolio choice and health status. *Journal of Banking and Finance*, 30(4):1201–1217.

Blau, D. M. and Gilleskie, D. B. (2008). The role of employee health insurance in the employment behavior of older men. *International Economic Review*, 49(2):475–514.

Bosworth, B. P., Burtless, G., and Zhang, K. (2016). Later retirement, inequality in old age, and the growing gap in longevity between rich and poor. Economic Studies at Brookings, Brookings Institute.

Carroll, C. D. (2002). Portfolios of the rich. In Guiso, L., Haliassos, M., and Jappelli, T., editors, *Household Portfolios*, chapter 10, pages 389–429. MIT Press, Cambridge MA.

Case, A. and Deaton, A. (2005). Broken down by work and sex: How our health declines. In Wise, D. A., editor, *Analyses in the Economics of Aging*, chapter 6, pages 185 – 205. University of Chicago Press, Chicago.

Coile, C. and Milligan, K. (2009). How household portfolios evolve after retirement: The effect of aging and health shocks. *Review of Income and Wealth*, 55(2):226 – 248.

Dalgaard, C.-J. and Strulik, H. (2014). Optimal aging and death: Understanding the preston curve. *Journal of the European Economic Association*, 12(3):672 – 701.

Dalgaard, C.-J. and Strulik, H. (2017). The genesis of the golden age: Accounting for the rise in health and leisure. *Review of Economic Dynamics*, 24:132 – 151.

De Nardi, M., French, E., and Jones, J. B. (2009). Life expectancy and old age savings. *American Economic Review*, 99(2):110–115.

De Nardi, M., French, E., and Jones, J. B. (2010). Why do the elderly save? The role of medical expenses. *Journal of Political Economy*, 118(1):39–75.

De Nardi, M., French, E., and Jones, J. B. (2016). Savings after retirement: A survey. *Annual Review of Economics*, 8:177–204.

De Nardi, M., French, E., Jones, J. B., and McCauley, J. (2015). Medical spending of the U.S. elderly. Working Paper 21270, National Bureau of Economic Research.

DiMatteo, L. (2003). The income elasticity of health care spending: A comparison of parametric and nonparametric approaches. *European Journal of Health Economics*, 4(1):20–29.

Duffie, D. and Epstein, L. G. (1992). Asset pricing with stochastic differential utility. *Review of Financial Studies*, 5(3):411–436.

Ehrlich, I. and Chuma, H. (1990). A model of the demand for longevity and the value of life extension. *Journal of Political Economy*, 98(4):761–782.

Fan, E. and Zhao, R. (2009). Health status and portfolio choice: Causality or heterogeneity? *Journal of Banking and Finance*, 33(6):1079 – 1088.

Finkelstein, A., Luttmer, E. F., and Notowidigdo, M. J. (2009). Approaches to estimating the health state dependence of the utility function. *American Economic Review*, 99(2):116 – 121.

Finkelstein, A., Luttmer, E. F., and Notowidigdo, M. J. (2013). What good is wealth without health? The effect of health on the marginal utility of consumption. *Journal of the European Economic Association*, 11(s1):221–258.

French, E., De Nardi, M., Jones, J. B., Baker, O., and Doctor, P. (2006). Right before the end: Asset decumulation at the end of life. *Federal Reserve Bank of Chicago Economic Perspectives*, 30(3):2 – 13.

Galama, T. J. (2015). A contribution to health-capital theory. CESR-Schaeffer Working Paper Series 2015-004, USC Dornsife, USC Schaeffer.

Gilleskie, D. B. and Mroz, T. A. (2004). A flexible approach for estimating the effects of covariates on health expenditures. *Journal of Health Economics*, 23(2):391 – 418.

Goldman, D. and Maestas, N. (2013). Medical expenditure risk and household portfolio choice. *Journal of Applied Econometrics*, 28(4):527 – 550.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–255.

Grossman, M. (2000). The human capital model. In Culyer, A. J. and Newhouse, J. P., editors, *Handbook of Health Economics*, volume 1, chapter 7, pages 347–408. North-Holland, Amsterdam.

Gruber, J. (2013). A tax-based estimate of the elasticity of intertemporal substitution. *Quarterly Journal of Finance*, 3(1):1 – 20.

Guiso, L., Jappelli, T., and Terlizzese, D. (1996). Income risks, borrowing constraints, and portfolio choice. *American Economic Review*, 86(1):158–172.

Hall, R. E. and Jones, C. I. (2007). The value of life and the rise in health spending. *Quarterly Journal of Economics*, 122(1):39–72.

Heiss, F. (2011). Dynamics of self-rated health and selective mortality. *Empirical Economics*, 40(1):119 – 140.

Hugonnier, J., Pelgrin, F., and St-Amour, P. (2013). Health and (other) asset holdings. *The Review of Economic Studies*, 80(2):663–710.

Hurd, M. D. (2002). Portfolio holdings of the elderly. In Guiso, L., Haliassos, M., and Jappelli, T., editors, *Household Portfolios*, pages 431–472. MIT Press, Cambridge MA.

Hurd, M. D., McFadden, D., and Merrill, A. (2001). Predictors of mortality among the elderly. In Wise, D. A., editor, *Themes in the economics of aging*, NBER Conference Report series, pages 171–97. University of Chicago Press, Chicago and London.

Kuhn, M., Wrzaczek, S., Prskawetz, A., and Feichtinger, G. (2015). Optimal choice of health and retirement in a life-cycle model. *Journal of Economic Theory*, 158(Part A):186 – 212.

Laporte, A. and Ferguson, B. S. (2007). Investment in health when health is stochastic. *Journal of Population Economics*, 20(2):423 – 444.

Love, D. A., Palumbo, M. G., and Smith, P. A. (2009). The trajectory of wealth in retirement. *Journal of Public Economics*, 93(1-2):191–208.

Marshall, S., McGarry, K. M., and Skinner, J. S. (2010). The risk of out-of-pocket health care expenditure at end of life. Working paper 16170, National Bureau of Economic Research.

Meer, J., Miller, D. L., and Rosen, H. S. (2003). Exploring the health-wealth nexus. *Journal of Health Economics*, 22(5):713–730.

Michaud, P.-C. and van Soest, A. (2008). Health and wealth of elderly couples: Causality tests using dynamic panel data models. *Journal of Health Economics*, 27(5):1312–1325.

Palumbo, M. G. (1999). Uncertain medical expenses and precautionary saving near the end of the life cycle. *Review of Economic Studies*, 66(2):395–421.

Pelgrin, F. and St-Amour, P. (2016). Life cycle responses to health insurance status. *Journal of Health Economics*, 49:79–96.

Rosen, H. S. and Wu, S. (2004). Portfolio choice and health status. *Journal of Financial Economics*, 72(3):457–484.

Smith, J. P. (1999). Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspective*, 13(2):145–166.

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

St-Amour, P. (2018). The lifetime dynamics of health and wealth. In *Oxford Research Encyclopedia of Economics and Finance*. Oxford University Press.

Tsai, Y. (2015). Social security income and the utilization of home care: Evidence from the social security notch. *Journal of Health Economics*, 43:45 – 55.

Van Kippersluis, H., Van Ourti, T., O'Donnell, O., and van Doorslaer, E. (2009). Health and income across the life cycle and generations in Europe. *Journal of Health Economics*, 28:818–830.

van Ooijen, R., Alessie, R., and Kalwij, A. (2015). Saving behavior and portfolio choice after retirement. *De Economist*, 163(3):353 – 404.

Wachter, J. A. and Yogo, M. (2010). Why do household portfolio shares rise in wealth? *Review of Financial Studies*, 23(11):3929 – 3965.

Yogo, M. (2009). Portfolio choice in retirement: Health risk and the demand for annuities, housing, and risky assets. NBER Working Paper 15307, National Bureau Of Economic Research.

# A Parametric restrictions

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

$$\beta < \left(r + \tilde{\delta}\right)^{\frac{1}{\alpha}}, \tag{17a}$$

$$\max\left(0; r - \frac{\lambda_{m0}}{1-\gamma_m} + \theta^2/\gamma\right) < A, \tag{17b}$$

$$0 < A - \max\left(0, r - \frac{\lambda_{m0}}{1-\gamma_m} + \theta^2/\gamma\right) - F(-\xi_m), \tag{17c}$$

where $\tilde{\delta} = \delta + \phi\lambda_{s0}$ and we define

$$\begin{aligned} \chi(x) &= 1 - (1-\phi)^{-x} < 0, \\ F(x) &= x(\alpha B)^{\frac{\alpha}{1-\alpha}} - x\delta - \lambda_{s0}\chi(-x), \\ L_m &= [(1-\gamma_m)(A - F(-\xi_m))]^{-1} > 0. \end{aligned}$$

The shadow price of health $B$ is defined by:

$$\begin{aligned} g(B) &= \beta - (r + \delta + \phi\lambda_{s0})B - (1 - 1/\alpha)(\alpha B)^{\frac{1}{1-\alpha}} \\ &= \beta - (r + \tilde{\delta})B + \left(\frac{1-\alpha}{\alpha}\right) BK = 0 \end{aligned} \tag{18a}$$

subject to:

$$\begin{aligned} g'(B) &= -(r + \tilde{\delta}) + (\alpha B)^{\frac{\alpha}{1-\alpha}} \\ &= -(r + \tilde{\delta}) + \frac{BK}{\alpha B} < 0, \end{aligned} \tag{18b}$$

where we denote by $\tilde{\delta} = \delta + \phi\lambda_{s0}$ the sickness-adjusted expected depreciation rate of health and where $K = \alpha^{\frac{1}{1-\alpha}} B^{\frac{\alpha}{1-\alpha}} \geq 0$. The shadow price $B$ in (18) is independent of preferences and increases in the health gradient $\beta$ in preferences, while falling in the sickness and depreciation parameters $(\lambda_{s0}, \phi, \delta)$. The marginal propensity to consume parameter $A$ is defined in (18) and in (19a).

**Closed-form solutions for optimal rules parameters** The closed-form expression for the parameters in the optimal rules are obtained as follows:

$$A = \varepsilon\rho + (1-\varepsilon)\left(r - \frac{\lambda_{m0}}{1-\gamma_m} + \frac{\theta^2}{2\gamma}\right) \geq 0. \tag{19a}$$

$$\Theta = \left(\rho^{-\varepsilon} A\right)^{\frac{1}{1-\varepsilon}} \geq 0. \tag{19b}$$

as well as:

$$\mathcal{V}_1 = \lambda_{m1}\Theta L_m \geq 0, \tag{20a}$$

$$\mathcal{I}_1 = \lambda_{m1}\left(\xi_m K/(1-\alpha)\right) L_m \geq 0, \tag{20b}$$

$$\mathcal{X}_1 = \lambda_{m1}\chi(\xi_m)\left(1/\gamma_s - 1\right) L_m \gtreqless 0, \tag{20c}$$

$$\mathcal{C}_1 = \lambda_{m1} A(\varepsilon - 1) L_m \gtreqless 0. \tag{20d}$$

# B Main theoretical results and proofs

## B.1 Health dynamics

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

$$\mathcal{D}_H = \left\{(W,H) \in \mathcal{A} : \frac{1}{\mathrm{d}t} E_{t-}[\mathrm{d}H_t \mid W_{t-} = W, H_{t-} = H] < 0\right\} \tag{21}$$

$$\mathcal{AC} = \left\{(W,H) \in \mathcal{D}_H : I_H^h(W,H) > 0\right\}, \tag{22}$$

**Proposition 1 (Health depletion and acceleration)** *Assume that the agent follows the approximate optimal rules in Theorem 1. Then, the health depletion set $\mathcal{D}_H$ in (21) is non-empty if and only if:*

$$BK < \tilde{\delta}^{1/\alpha}. \tag{23}$$

*Under condition (23):*

1. *The health depletion zone is given by:*

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

*where the health depletion locus is*

$$\begin{aligned} y(H) &= x(H) + DH^{1+\xi_m}, \\ D &= \mathcal{I}_1^{-1}\left[\tilde{\delta}^{1/\alpha} - BK\right] > 0. \end{aligned} \tag{24a}$$

2. *The accelerating region (22) is given by:*

$$\mathcal{AC} = \{(W, H) \in \mathcal{D}_H : W < \min[y(H), z(H)]\}, \tag{25a}$$

*where the acceleration locus is*

$$z(H) = x(H) + \frac{BH}{1+\xi_m}. \tag{25b}$$

**Proof**. The expected local change in health capital is given by:

$$\frac{1}{\mathrm{d}t} E_{t-}[\mathrm{d}H_t] = \left[I^h(W_{t-}, H_{t-})^\alpha - \tilde{\delta}\right] H_{t-}. \tag{26}$$

The investment-to-health ratio evaluated at the optimal investment in (8b) is given by:

$$I^h(W_{t-}, H_{t-}) = \frac{I^*(W_{t-}, H_{t-})}{H_{t-}} = BK + \mathcal{I}_1 H_{t-}^{-\xi_m - 1} N_0(W_{t-}, H_{t-}). \tag{27}$$

Substituting the investment-to-capital ratio (27) in the expected local change for health (26) and using the definition of net total wealth (7) shows that:

$$\begin{aligned} \frac{1}{\mathrm{d}t} E_{t-}[\mathrm{d}H_t \mid W_{t-} = W, H_{t-} = H] &= \left\{\left[BK + \mathcal{I}_1 H^{-\xi_m - 1} N_0(W, H)\right]^\alpha - \tilde{\delta}\right\} H, \\ &< 0 \iff W < y(H) = x(H) + DH^{1+\xi_m}, \end{aligned}$$

where $D = \mathcal{I}_1^{-1}\left[\tilde{\delta}^{1/\alpha} - BK\right]$.

Assume that necessary and sufficient condition (23) is violated. Because $\mathcal{I}_1 > 0$ in (20b), we have that $D < 0$. Consequently, we have that $y(H) \leq x(H), \forall H$ and it follows that

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

Hence a non-empty health depletion set obtains if and only if restriction (23) is verified, under which $D > 0$.

Second, observe that the health depletion locus is characterized by:

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

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

The locus $y(H)$ is therefore convex and U-shaped under condition (23) and attains a unique minimum at $\bar{H}_3$ in the $(H, W)$ space, where $\bar{H}_3$ is given in (14), with corresponding wealth level $\bar{W}_3 = y(\bar{H}_3)$.

Next, taking the derivative of the investment-to-health ratio (27) with respect to $H$ and rearranging shows that the accelerating region can be characterized by:

$$\begin{aligned} I^h_H(W, H) &= -(1+\xi_m)H^{-\xi_m-2}\mathcal{I}_1 N_0(W, H) + H^{-\xi_m-1}\mathcal{I}_1 B \\ &> 0 \iff W < z(H) = x(H) + \frac{BH}{1+\xi_m}. \end{aligned}$$

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. $\mathcal{AC} \subset \mathcal{A}$. 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 (14). 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 $I_H^h > 0$) is the entire $\mathcal{D}_H$ for $H \in [0, \bar{H}_3]$ and otherwise the area between $y(H), z(H)$, as given in (25).

■

## B.2 Wealth dynamics

The following proposition characterizes the wealth depletion zone:

$$\mathcal{D}_W = \left\{ (W, H) \in \mathcal{A} : \frac{1}{\mathrm{d}t} E_{t-}[\mathrm{d}W_t \mid W_{t-} = W, H_{t-} = H] < 0 \right\}, \tag{28}$$

**Proposition 2 (Wealth depletion)** *Assume that the agent follows the approximate optimal rules in Theorem 1 and that condition (23) in Proposition 1 is verified. Then, the wealth depletion set $\mathcal{D}_W$ in (28) is non-empty if and only if there exists health levels $H$ such that:*

$$l(H) = A - \frac{\theta^2}{\gamma} - r + (\mathcal{I}_1 + \mathcal{C}_1) H^{-\xi_m} > 0. \tag{29}$$

*Under condition (29), the wealth depletion zone is given by:*

$$\mathcal{D}_W = \{(W, H) \in \mathcal{A} : W > w(H)\}, \tag{30a}$$

*where the wealth depletion locus is*

$$w(H) = \frac{x(H)[l(H) + r]}{l(H)} + \frac{k(H)}{l(H)}, \tag{30b}$$

$$k(H) = y_0 - a + H(\beta - KB). \tag{30c}$$

**Proof**. Observing that the expected net return on actuarially fair insurance contracts (3) is zero, the expected local change in wealth is:

$$\begin{aligned} \frac{1}{\mathrm{d}t} E_{t-}[\mathrm{d}W_t] = & [rW_{t-} + Y(H_{t-}) - C^*(W_{t-}, H_{t-}) - I^*(W_{t-}, H_{t-}) \\ & + \Pi^*(W_{t-}, H_{t-}) \sigma_S \theta] \,. \end{aligned} \tag{31}$$

We can use the definition of net total wealth (7) and substitute the optimal investment (8b), as well as the optimal consumption (8d) and risky portfolio (8e) in the expected local change for wealth (31) to obtain:

$$\frac{1}{\mathrm{d}t} E_{t-}[\mathrm{d}W_t \mid W_{t-} = W, H_{t-} = H] = \{rW + k(H) - N_0(W, H)[l(H) + r]\}$$
$$< 0 \iff W l(H) > x(h)[l(H) + r] + k(H),$$

where,

$$l(H) = \left[ A - \frac{\theta^2}{\gamma} - r + (\mathcal{I}_1 + \mathcal{C}_1) H^{-\xi_m} \right],$$
$$k(H) = (y_0 - a) + H(\beta - BK),$$

as given in (29), (30c).

Assume that necessary and sufficient restriction (29) is violated such that $l(H) < 0$, then $E_{t-}[\mathrm{d}W_t \mid W_{t-} = W, H_{t-} = H]/\mathrm{d}t < 0$ obtains if:

$$W < w(H) = \frac{x(H)[l(H) + r]}{l(H)} + \frac{k(H)}{l(H)}.$$

Since $l(H) < 0$, it follows that

$$w(H) \leq x(H) \iff x(H)r + k(H) \geq 0.$$

Relying on the definition of $g(B)$ in (18a) and from necessary and sufficient condition (23) shows that

$$\begin{aligned} x(H)r + k(H) &= H[\beta - B(r + K)] \\ &= HB[\tilde{\delta} - K/\alpha] \\ &= HB[\tilde{\delta} - (BK)^\alpha] > 0. \end{aligned}$$

When (29) is violated and $l(H) < 0$ the wealth depletion zone thus simplifies to:

$$\mathcal{D}_W = \{(W, H) \in \mathcal{A} : W < w(H)\} = \emptyset$$

since $w(H) \leq x(H)$. Consequently, a non-empty wealth depletion set obtains if and only if restriction (29) is verified and is delimited by:

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

where $w(H)$ is given by (30b), as stated. It is straightforward to show that:

$$\lim_{H \to 0} \frac{l(H) + r}{l(H)} = 1, \quad \lim_{H \to 0} \frac{k(H)}{l(H)} = 0, \implies \lim_{H \to 0} w(H) = x(0) = -(y_0 - a)/r$$

such that the $w(H)$ shares the same intercept with $x(H), y(H), z(H)$ and which is non-negative under condition (11). ■

**Proposition 3 (sufficient conditions)** *Assume that the agent follows the approximate optimal rules in Theorem 1. The following conditions:*

$$\tilde{\delta}^{1/\alpha} > \beta, \tag{32a}$$

$$\varepsilon > 1, \tag{32b}$$

$$A > \frac{\theta^2}{\gamma} + r \tag{32c}$$

*are sufficient for non-empty individual $\mathcal{D}_H, \mathcal{D}_W$ and joint depletion sets ($\mathcal{D}_W \cap \mathcal{D}_H$).*

**Proof**. This simplifies to showing:

$$\begin{aligned} w(H) \leq y(H) &\iff rx(H) + k(H) \leq l(H) D H^{1+\xi_m} \\ &\iff \beta - Br - \tilde{\delta}^{1/\alpha} \leq D H^{\xi_m} \left[ A - \frac{\theta^2}{\gamma} - r \right] + \mathcal{C}_1 D \end{aligned}$$

Since $\beta < \tilde{\delta}^{1/\alpha}$ under (32a), the left-hand side is negative, whereas $D > 0$. Moreover (32b) implies that $\mathcal{C}_1 \geq 0$, whereas the right-hand term in square bracket is also positive under condition (32c). It follows that the right-hand side is positive, and consequently sufficient for $w(H) \leq y(H)$, as required. ■

## B.3 Standard model

The well-known demand-for-health framework of Grossman (1972); Ehrlich and Chuma (1990) can be analyzed as a restricted case relying on (6).

**Proposition 4 (Exogenous mortality)** *Assume that the exposure to mortality risk cannot be adjusted, i.e.* $\lambda_{m1} = 0$*. Then, the health depletion set is non-empty if and only if condition* (23) *is verified, under which:*

1. *Health depletion is expected everywhere in the admissible set:*

$$\mathcal{D}_H = \mathcal{A}, \tag{33}$$

2. *The accelerating subset is empty:*

$$\mathcal{AC} = \emptyset, \tag{34}$$

3. *The wealth depletion set is non-empty if and only if condition* (32c) *is verified, under which* $\mathcal{D}_w$ *remains delimited by* (30a)*, where the wealth depletion locus is modified as:*

$$w(H) = \frac{x(H)[l + r]}{l} + \frac{k(H)}{l} > x(H), \tag{35a}$$

$$l = A - \frac{\theta^2}{\gamma} - r \tag{35b}$$

*and* $k(H)$ *remains as in* (30c)*.*

**Proof**. First, setting $\lambda_{m1} = 0$ results in the first-order adjustment $\mathcal{I}_1 = 0$ in (20b). Consequently, the investment-to-capital ratio in (27) is constant and given by $I^h = BK$. Substituting in (26) reveals that so is the expected growth rate:

$$E_{t-}[\mathrm{d}H_t] = \left[(BK)^\alpha - \tilde{\delta}\right] H_{t-}\mathrm{d}t,$$

and that the latter is negative under condition (23) for all admissible health and wealth levels. Consequently, the health depletion subset corresponds to the entire admissible set, as stated in (33). Moreover, a constant $I^h$ implies that it is orthogonal to the health status; consequently no accelerating region exists as stated in (34).

Finally, setting $\lambda_{m1} = 0$ also sets $\mathcal{I}_1, \mathcal{C}_1 = 0$ in equation (29) for $l(H)$. Condition (32c) implies that $l > 0$ in (35b) and as showed in Appendix B.2, is necessary and sufficient for $\mathcal{D}_W \neq \emptyset$. The wealth depletion locus $w(H)$ is modified accordingly by using $l$ in (35a).

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 Data

The database is the 2002 wave of the Health and Retirement Study (HRS, Rand data files), merged with the 2001 Consumption and Activities Mail Survey (CAMS), with 1,124 observations used for estimation.

**Consumption** The CAMS consumption data used to construct $C_j$ includes nondurables expenses on utilities (electricity, water, heating, phone, cable), house and yard supplies, food, recreation (dining out, vacations, tickets, hobbies) and others (gifts, contributions). Health expenses not considered as investment (drugs, personal health services and medical supplies, insurance, nursing homes, home health care and special facilities, dental visits) are appended to consumption.

**Wealth and risky assets** 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 investment** Health investments $I_j$ are obtained as the sum of OOP expenses on hospital, doctor visits, medication, and outpatient surgery.

**Health** Health status $H_j$ is evaluated using the self-reported general health status with the following scaling: 0.5 (poor), 1.25 (fair), 2.00 (good), 2.75 (very good) and 3.50 (excellent).[32]

# D 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 levels at base age $t = 75$ by drawing uniformly $W_{75} \sim [W_{\min}, W_{\max}]$, and $H_{75} \sim [H_{\min}, H_{\max}]$ taken from Table 1.

2. We simulate individual-specific Poisson health shocks $\mathrm{d}Q_s \sim P(\lambda_{s0})$, as well as a population-specific sequence of Brownian financial shocks $\mathrm{d}Z \sim N(0, \sigma_s^2)$ 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) 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 realized sickness and financial shocks $\mathrm{d}Q_{st}, \mathrm{d}Z_t$ in the stochastic laws of motion $\mathrm{d}H_t, \mathrm{d}W_t$.

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

$$
\begin{aligned}
H_{t+1} &= H_t + \mathrm{d}H_t(H_t, I_t, \mathrm{d}Q_{st}) \\
W_{t+1} &= W_t + \mathrm{d}W_t \left[W_t, C(W_t, H_t), I(W_t, H_t), \Pi(W_t, H_t), X(W_t, H_t), \mathrm{d}Q_{s,t}, \mathrm{d}Z_t\right]
\end{aligned}
$$

4. For each agent with health $H_t$, we generate the Poisson death shocks with endogenous intensities $\mathrm{d}Q_m \sim P[\lambda_m(H_t)]$ and keep only the surviving agents at each date.

5. We verify admissibility, for each agent with health and wealth $(H_t, W_t)$ and keep only surviving agents in the admissible region $\mathcal{A} : W_t \geq x(H_t)$.

6. We replicate the simulation 1–5 for 500 times.