

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

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

¹École Polytechnique Fédérale de Lausanne

²EDHEC Business School

³University of Lausanne, Faculty of Business and Economics (HEC)

⁴Swiss Finance Institute

⁵CIRANO

July 12, 2016

*Financial support from the Swiss Finance Institute is gratefully acknowledged.

Abstract

The observed health decline near the end of life coincides with less curative (e.g. hospital stay, doctor visits), and more comfort (e.g. nursing home) care, which accelerate both the fall in wealth, and the timing of death. We investigate whether these dynamics jointly result from a *closing down the shop* decision i.e. a depletion of the health stock is optimally selected (and eventually accelerated), leading to states characterized by indifference between life, and death. Towards that aim, we expand, structurally estimate, and simulate a life cycle model of financial, and health expenses with endogenous mortality exposure (Hugonnier et al., 2013). Under economically plausible, and statistically verified conditions, we find that, unless sufficiently rich and healthy, agents will optimally select expected depletion of their health capital, and associated increase in death likelihood. Moreover, we identify a wealth and health locus below which agents *accelerate* their health depletion. Importantly, wealth is also expected to decline for all, such that all surviving agents eventually enter the closing down phase.

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 recedes rapidly in old age.¹ Towards the end of life, this decline is associated with both an increase,² and a change in composition in health expenditures. Indeed, the top panel of Figure 1, reproduced from De Nardi et al. (2015b, Fig. 3, p. 22) shows that agents approaching the last phase of old age substitute away from curative medical expenses (e.g. doctor visits and short-term hospital stays, . . .) in favor of more nursing home, and other long-term care (LTC) spending.³ Importantly, LTC expenses are uninsured under Medicare, as well as under most Medigap private plans, while remaining subject to means-testing under Medicaid. Consequently, as shown in the bottom panel of Figure 1, increasing health expenses entail large out-of-pocket (OOP) spending for most individuals,⁴ leading to a rapid drain of financial resources.⁵

Put differently, the observed health decline in old age is met with accompanying (comfort), rather than countervailing (curative) health choices, which decumulate assets, and likely hasten the timing of death.⁶ The objective of this paper is to assess the

¹Banks et al. (2015, Fig. 5, p. 12) find that the percentage of agents reporting worse health doubles between ages 40, and 70. Similar age declines in self-reported health status are reported in Smith (2007, Fig. 1, p. 740), and Case and Deaton (2005, Fig. 6.1, p. 186), who highlight faster deterioration at lower income quartiles. Controlling for attrition, Heiss (2011, Fig. 2, p. 124) finds a doubling of the share of agents in poor health between age 70, and 80. See also Van Kippersluis et al. (2009, Figs. 2, 3, pp. 822–826) for declining age profiles in health status in European countries.

²De Nardi et al. (2015b, p. 3, 23, and Tab. 12.b, p. 24) find that health spending by all payors doubles between age 70–90, reaching \$25,000 by age 90, and increasing to \$43,030 in the last year of life. End-of-life spending has also been estimated to represent a quarter of lifetime health expenditures (Philipson et al., 2010).

³De Nardi et al. (2015b, Tab. 2, p. 7, and Figs. 3, 4, p. 22, 26) show that most of the end-of-life increase in health spending is explained by LTC (e.g. nursing homes) expenses who averaged between \$77,000–\$88,000 per year in 2014. Professional services (e.g. doctor visits) are falling after age 65. To the extent that LTC and home care demand is strongly related to wealth, it shares attributes with luxury goods consumption, and it is associated more with comfort, more than curative care (De Nardi et al., 2015b; Marshall et al., 2010). For example, sharp increases in the use of formal home care by elders is observed under more generous Social Security (Tsai, 2015).

⁴De Nardi et al. (2015b, Tab. 4, p. 13) estimate that OOPs account for 19.4% of total expenditures. See also De Nardi et al. (2012, 2010); Marshall et al. (2010); Palumbo (1999) for further evidence and discussion of the importance of OOPs.

⁵De Nardi et al. (2015a, p. 9) shows that median assets for individuals aged 76 in 1996 and who survived to 2006 fell from \$84,000 to \$44,000. Wealth depletion is especially acute towards the end of life where wealth falls by 50% in the last three years (30% in the last year alone) for those agents who die, compared to only 2% for those who don't (French et al., 2006, Fig. 1, p. 7). Falling wealth is also correlated with occurrence of severe illness late in life, and to a lesser extent to chronic diseases (Lee and Kim, 2008). Wealth profiles remain comparatively flat in the absence of significant changes in health status and/or family composition (Poterba et al., 2015).

⁶The number of survivors in U.S. Life Tables drops by 88.4% between ages 70, and 95, compared to a drop of 18.5% between ages 45, and 70 (Arias, 2014, Tab. B, p. 4). Smith (2007, Tabs. 1–3, pp. 747–752) finds that current self-reported health status is a significant predictor of major health onsets. Benjamins et al. (2004, Tab. 4, p. 1303) report a doubling of mortality risk for those reporting poor compared to

optimality of these joint end-of-life health and wealth dynamics. More precisely, we look at conditions under which health, and wealth depletion is *optimally selected* to fall towards a region associated with very high mortality risk, and indifference between life and death. Put differently, agents optimally choose to let go, i.e. to *close down the shop*. The conditions under which they do are associated with lower wealth, such that richer individuals delay the health depletion. However, we show that under reasonable assumptions, wealth depletion is *also* optimally selected, such that agents eventually enter the closing-down phase. We further identify threshold effects whereby the depletion of the health capital is initially slowed down, before being *accelerated*.

This paper has two main contributions: theoretical and empirical. First, we build upon a rich life cycle model developed in Hugonnier, Pelgrin and St-Amour (2013) in order to identify optimal depletion, and acceleration. This model encompasses a health investment setup with endogenous exposure to death risk, and exogenous sickness shocks that further depreciate the health capital. In addition to investing in their health, agents can buy actuarially fair insurance against health shocks, and save in risky, and risk-less assets. Agents also earn income, part of which is fixed (e.g. social security), and part which is health-dependent, reflecting their physical ability to work. Finally, preferences are characterized by subsistence consumption, as well as by generalized attitudes towards the various sources of risk (mortality, morbidity, and financial), and towards inter-temporal substitution. Importantly, they also guarantee strict *ex-ante* preference for life, so that agents have no proclivity for death over life.

We rely on the closed-form solutions of that model in order to characterize the optimal dynamics for health, and wealth capitals. Our main theoretical results (i) define the conditions under which expected health and wealth depletion arises, and (ii) partition the health and wealth state space to identify whether or not agents are in these regions. First, the conditions necessary for depletion are economically plausible and relevant for agents approaching the end of life. Indeed, they require that consumption (including subsistence) propensities as well as, sickness-adjusted depreciation of health capital are high, whereas the health-adjusted ability to generate income is reduced.

Second, under these assumptions, we identify a U-shaped locus in the health-wealth nexus such that all agents who are below the locus, i.e. who are insufficiently rich/healthy,

excellent health. See also Heiss (2011); Hurd et al. (2001); Hurd and McGarry (2002) for additional evidence on health-dependent mortality.

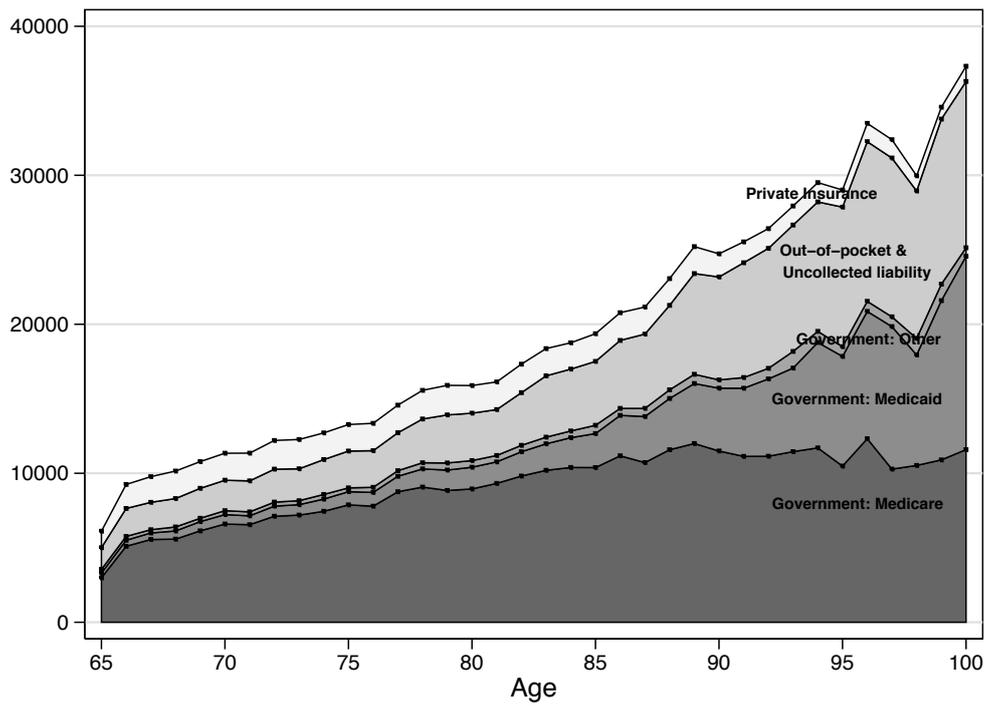
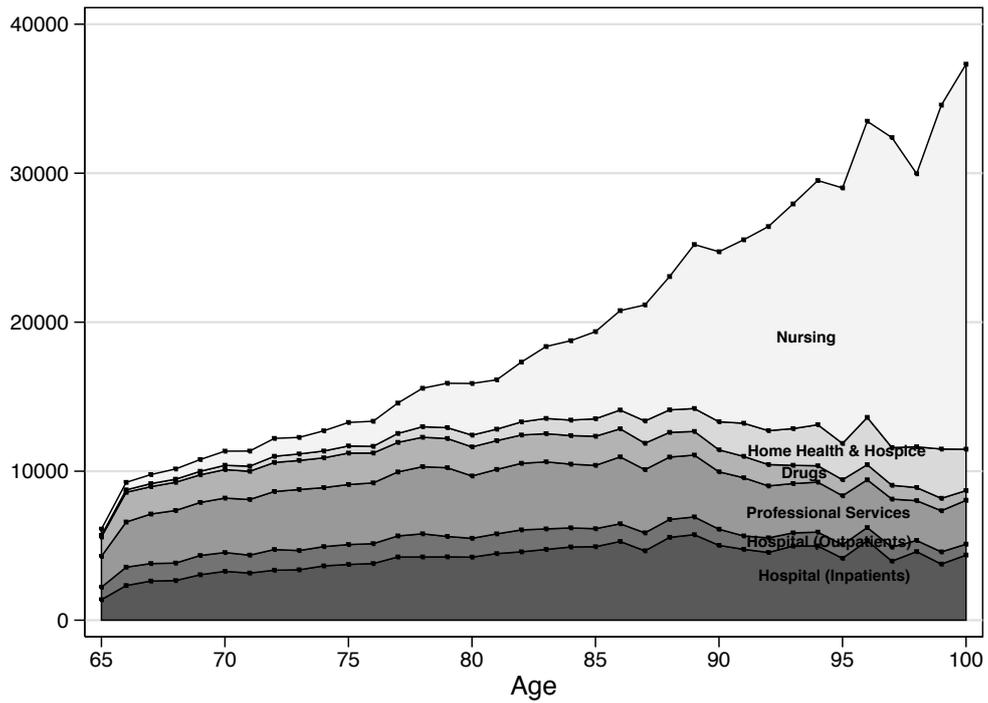


Figure 1: Health expenses composition, and payors

Notes: Source: Reproduced from De Nardi et al. (2015b, Fig. 3, p. 22). © 2015 by Mariacristina De Nardi, Eric French, John Bailey Jones, and Jeremy McCauley.

optimally select expected depletion of their health stock. Consequently, there exists a threshold wealth level below which all agents expect a health depletion, regardless of their health status. Importantly, wealth depletion is also optimally selected, irrespective of the health and wealth levels. Combining these elements entails that health is set on a downward spiral leading to drops in available resources, further cuts in health spending, and additional depletion of the health stock. We can also identify an accelerating locus below which health spending falls faster than health, such that agents initially slow down (yet do not reverse) the depletion, before choosing to *accelerate* the decline in health. Health thus eventually falls towards low levels that are associated with very high mortality risks, and indifference between life and death.

Our second contribution is empirical. Using HRS cross-sectional data, we rely on a trivariate econometric system composed of optimal health spending, risky asset holdings, and health-dependent income to structurally estimate the model over a population of relatively old agents. This exercise allows us to estimate the model's deep parameters, and evaluate the induced parameters that are used to partition the state space. The first set of results helps gauge and confirm the model's realism. The second set allows us to evaluate and also confirm the economic relevance of the depletion zones.

In particular, we show that all the required conditions are met for the existence of optimal closing-down strategies. Moreover, we show that the bulk of the population is located in the health depletion region, with a subset located in the accelerating zone. We also substitute the estimated theoretical allocations in the laws of motion for health and wealth in order to simulate the life cycles in the last period of life. The results we get are consistent with theoretical predictions, and with stylized facts for end-of-life dynamics. Indeed, the simulated optimal trajectories show rapid wealth, and health depletion. Net total human and financial capital is thus exhausted by the end of the expected lifespan, at which point indifference between life and death is predicted. These results therefore are pointing towards agents jointly selecting a short lifespan, and corresponding closing-down strategies that are consistent with remaining lifetime.

The main novelty of our approach concerns the *optimality* of the joint health and wealth depletion processes near the end of life. Despite strictly 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. 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. Indeed most life cycle models of asset decumulation in old age rely on exogenous health status, and expenses.⁷ In the other cases, endogenous expenses provide direct utility flows, but have no bearing on health status (e.g. De Nardi et al., 2010; Yogo, 2009). Regardless of whether they do, feedback effects on exposure to death risks are almost always abstracted from.⁸ Consequently, longevity is exogenously set, and cannot be altered through the agent's health decisions; in the absence of bequest motives, the optimal strategy thus fully depletes wealth reserves at death.

These results raise two important normative issues. First, from a distributional point of view, we show that reducing the incidence of depletion zones can be achieved through an increase in base income, (e.g. through social security, or minimal revenue programs). This feature is consistent with observed shorter horizons for the poor,⁹ and adds supplemental resonance to the usual finding of savings inadequacy for U.S. households.¹⁰ Yet, to the extent that health and wealth depletion stems from optimizing behavior in a complete, and frictionless market setting, whether or not the state should intervene to prevent their occurrence is open to debate. Second, and related, an unresolved ethical question is whether or not medical treatment should be imposed to agents in the closing-down phase.¹¹ This paper argues that high curative spending may not always reflect what agents actually want.¹² Again from the perspective that the downward spiral in health

⁷See De Nardi et al. (2015a, 2009), or French and Jones (2011) for examples.

⁸Exceptions 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.

⁹For example, longevity for males from a 1940 cohort in HRS are 73.3 years in the first decile of career earnings, and 84.6 if in the 10th decile (Bosworth et al., 2016, Tab. IV-4, p. 87).

¹⁰See Hubbard et al. (1994, 1995); Skinner (2007) among others on insufficient financial, and pension savings.

¹¹The legal right to refuse treatment is protected under both common law, and the American constitution (Legal Advisors Committee of Concerns for Dying, 1983), and recognized as such by the AMA (American Medical Association, 2016).

¹²Indeed, in 2010, among adults below 100% of the poverty level, 23.4% did not get or delayed medical care due to cost, 21.5% did not get prescription drugs due to cost, and 30.4% did not get dental care due to cost; these numbers fell to 6.8%, 3.9%, and 7.0% for richer households 400% above poverty (Tab. 79, National Center for Health Statistics, 2012, pp. 272–75). High prevalence of unmet medical needs due to financial reasons for uninsured Americans is also identified by Ayanian et al. (2000), especially in the case of unhealthy individuals. Park et al. (2016) find similar high incidence due to financial limitations in the case of Korean elders.

is optimally selected, a more subtle approach favoring end-of-life comfort care may be required.

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. We close the discussion with concluding remarks in Section 6.

2 Theoretical framework

Our analysis relies on the theoretical framework we developed in Hugonnier et al. (2013) which is built upon to analyze the existence of depletion regions of the state space. The main features of this model are briefly reproduced here for completeness.

2.1 Economic environment

The agent's health level H_t follows a generalized stochastic version of the Grossman (1972) demand-for-health model:

$$dH_t = ((I_t/H_t)^\alpha - \delta) H_t dt - \phi H_t dQ_{st}, \quad H_0 > 0, \quad (1)$$

where $I_t \geq 0$ is health spending. The positive restriction on investment is standard, and implies that the agent cannot sell his own health in markets. The Cobb-Douglas parameter $\alpha \in (0, 1)$ captures diminishing returns to investing in one's health, and the continuous deterministic depreciation δ is augmented by a factor ϕ upon occurrence of a stochastic sickness shock dQ_{st} . The latter follows a Poisson process with constant intensity:¹³

$$\lambda_s(H_t) = \lambda_{s0}. \quad (2)$$

¹³Hugonnier et al. (2013) consider a more general endogenous sickness intensity function given by:

$$\lambda_s(H_{t-}) = \eta + \frac{\lambda_{s0} - \eta}{1 + \lambda_{s1} H_{t-}^{-\xi_s}},$$

and where we have set $H_{t-} = \lim_{\tau \uparrow t} H_\tau$ as the health level prior to occurrence of health shocks. For the current study, we restrict our analysis to the case of $\lambda_{s1} = 0$ corresponding to exogenous morbidity $\lambda_s(H_{t-}) = \lambda_{s0}$. Importantly, it will considerably facilitate the exposition.

In a parallel vein, the age of death T_m also follows a Poisson process, however with health-dependent endogenous death intensity:

$$\lambda_m(H_t) = \lim_{\tau \rightarrow 0} \frac{1}{\tau} P_t [t < T_m \leq t + \tau] = \lambda_{m0} + \lambda_{m1} H_t^{-\xi_m}. \quad (3)$$

The component λ_{m0} captures endowed exposure to death risk, whereas the second term $\lambda_{m1} H_t^{-\xi_m}$ determines endogenous exposure in that healthier agents can expect longer time horizon. The parameter ξ_m controls diminishing returns to investing in one's health to prolong life, whereas λ_{m1} controls the degree of endogeneity, with $\lambda_{m1} = 0$ restricting the death intensity to be exogenous.

Regarding the budget constraint, agents receive an income Y_t at a rate that positively depends on their health:

$$Y(H_t) = y_0 + \beta H_t. \quad (4)$$

The base income y_0 captures health-independent elements such as Social Security revenue, whereas the health-dependent component βH_t captures the enhanced work ability for healthier agents.¹⁴ Furthermore, individuals can save and invest π_t in a risky asset whose returns follow a Brownian motion with market price of financial risk $\theta = \sigma_S^{-1}(\mu - r) \geq 0$, where μ is the drift, and σ_S the diffusion of the risky asset, and r is the risk-free asset rate. They can also purchase X_{t-} units of an actuarially fair health insurance contracts paying one unit of the numeraire upon positive occurrence of the health shock.¹⁵ The net return on insurance contracts dM_{st} is thus:

$$X_{t-} dM_{st} = X_{t-} dQ_{st} - X_{t-} \lambda_{s0} dt. \quad (5)$$

Denoting c_t the consumption, the budget constraint can be written as:

$$dW_t = (rW_{t-} + Y_t - c_t - I_t) dt + \pi_t \sigma_S (dZ_t + \theta dt) + X_{t-} dM_{st}. \quad (6)$$

¹⁴Old-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 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.

¹⁵We have set $X_{t-} = \lim_{\tau \uparrow t} X_\tau$ as the insurance purchased prior to occurrence of health shocks.

To close the model, the agent's preferences are characterized by generalized recursive utility pioneered by Duffie and Epstein (1992), which are augmented by source-dependent risk aversion:

$$U_t = 1_{\{T_m > t\}} E_t \int_t^{T_m} \left(f(c_\tau, U_{\tau-}) - \frac{\gamma \sigma_\tau^2}{2U_\tau} - \sum_{k=m}^s F_k(U_{\tau-}, H_{\tau-}, \Delta_k U_\tau) \right) d\tau, \quad (7)$$

where U_t is the continuation utility. The Kreps-Porteus aggregator function is

$$f(c, U) = \frac{\rho U}{1 - 1/\varepsilon} \left(((c - a)/U)^{1 - \frac{1}{\varepsilon}} - 1 \right) \quad (8)$$

with elasticity of intertemporal substitution $\varepsilon > 0$, time preference rate $\rho > 0$ and subsistence consumption level $a \geq 0$. For tractability, bequests are abstracted from.¹⁶

As explained in Duffie and Epstein (1992), financial risk aversion $\gamma > 0$ in (7) is disentangled from the elasticity of intertemporal substitution ε in (8), while the term $\gamma \sigma_\tau^2 / U_\tau$ is the utility cost associated with exposure to Brownian financial risks. In parallel, we can compute the expected utility jumps induced by exposure to discrete Poisson health-related risks as:

$$\Delta_k U_t = E_{t-} [U_t - U_{t-} | dQ_{kt} \neq 0].$$

Given two separate CRRA curvature indices $0 \leq \gamma_m < 1$ for death risk aversion, and $\gamma_s \geq 0$ for sickness risk aversion, the utility costs associated with these jumps are then given by the functions

$$F_k(U, H, \Delta_k U) = U \lambda_k(H) \left[\frac{\Delta_k U}{U} + u(1; \gamma_k) - u \left(1 + \frac{\Delta_k U}{U}; \gamma_k \right) \right], \quad (9)$$

where we have set

$$u(x; \gamma_k) = \frac{x^{1-\gamma_k}}{1-\gamma_k}, \quad k = m, s.$$

¹⁶The no-bequest assumption in preferences (7) can be justified by inconclusive empirical relevance. Hurd (2002) could find no evidence of a bequest motive in savings decisions, while Hurd (1987) finds no differences between the saving behavior of the elders with and without children. Both elements suggest accidental, rather than deliberate bequest decisions.

The term in square brackets in (9) is a positive penalty for exposure to jumps $\Delta_k U/U$. This penalty is U-shaped, with zero utility costs at $\Delta_k U = 0$, and is increasing in risk aversion γ_k (Hugonnier et al., 2013, Fig. 2). For reasons explained in more details in Hugonnier et al. (2013), risk aversion to morbidity risk $\gamma_s \geq 0$ is unrestricted, whereas risk aversion to mortality risk is bounded, $\gamma_m \in [0, 1)$.

The agent's problem is therefore to select optimal consumption, portfolio, insurance and investment so as to maximize utility (7):

$$V(W_t, H_t) = \sup_{(c, \pi, X, I)} U_t(c, I, H)$$

subject to the distributional assumptions, and laws of motion for health (1), and wealth (6). Observe that the agent faces an incomplete market setting with three independent sources of risks (financial, morbidity, and mortality), and only two traded assets (financial, and sickness insurance).

2.2 Optimal dynamic policies

The presence of endogenous exposure to death risk implies that the previous model has no closed-form solution. However, Hugonnier et al. (2013) rely on a two-step analytical approximation. First, they rewrite the agent's incomplete market, and stochastic finite horizon problem as an equivalent one with complete markets, infinite horizon, and endogenous health-decreasing discounting. Second, they show that, under regularity and transversality conditions restated in Appendix A.1, a closed-form solution exists in the restricted case of exogenous mortality (corresponding to $\lambda_{m1} = 0$, and referred to as order-0 solution). They then perform an asymptotic expansion to calculate the first-order effect of endogenous mortality, and use this expansion to obtain approximate solutions (referred to as order-1 solution).

Adapting the results of Hugonnier et al. (2013) to our setting shows that the optimal investment in health can be written as:

$$I(W, H) = \underbrace{K_0 B H}_{\text{Order-0 demand}} + \underbrace{\bar{K}_m \lambda_{m1} H^{-\xi_m} N_0(W, H)}_{\text{Death risk hedging demand}} \quad (10)$$

where K_0 and $K_m = \bar{K}_m \lambda_{m1}$ are positive constants defined in equations (34), and (35) in Appendix A.2, and where the net total wealth is:

$$N_0(W, H) = W + BH + C. \quad (11)$$

It can further be shown that optimal consumption, risky asset holdings, insurance and welfare are given by:

$$c(W, H) = a + AN_0(W, H) \quad (12)$$

$$\pi(W, H) = L_0 N_0(W, H) \quad (13)$$

$$X(H) = \phi BH \quad (14)$$

$$V(W, H) = \Theta N_0(W, H) \quad (15)$$

where the positive marginal propensity to consume $A > 0$, the portfolio share L_0 and the marginal value of net total wealth $\Theta > 0$ are also constant functions of the deep parameters that are defined in Appendix A.2.

In addition to financial assets W , total wealth (11) includes the shadow value of the human capital BH , for which B is the marginal (and average)- Q of health solving equation (36). This shadow value can be interpreted as the capitalized value of the health-dependent capacity to generate income βH in (4), where B is an increasing function of β . Finally, $C = (y_0 - a)/r$ is the net present value (NPV) of base income y_0 minus subsistence consumption a .

The first term in (10) is the order-0 investment that is proportional to the shadow value. The second term captures the additional demand for health that arises from its death risk hedging capacity; that demand is increasing in the endogenous component $\lambda_{m1} H^{-\xi_m}$ of the death intensity (3). Importantly, health investment (10) is thus the only optimal rule responding to endogenous death risk; all other variables in (12)–(15) have no first-order effect of $\lambda_{m1} H^{-\xi_m}$, and encompass at most only the exogenous component of mortality, λ_{m0} . 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.

The optimal rules for investment, consumption, portfolio, and insurance 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 (11). Indeed, observe from optimal consumption (12) that admissibility is required to ensure that consumption c_t is above subsistence a , and that the continuation utility of living V_t is positive in (15). Otherwise, negative total wealth entails negative continuation utility, and from preferences (7), a lower utility of living ($V_t < 0$), than of dying ($V_t = 0$). More precisely, we can rely on (11) to define:

Definition 1 (admissible) The admissible region \mathcal{A} is characterized by positive net total wealth:

$$\begin{aligned}\mathcal{A} &= \{(W, H) : N_0(W, H) \geq 0\}, \\ &= \{(W, H) : W > x(H) = -C - BH\},\end{aligned}$$

with complement non-admissible set denoted $\mathcal{N}\mathcal{A}$.

3 Optimal health and wealth dynamics

The joint health and wealth system composed of the the laws of motion (1), and (6), and evaluated at the optimal rules (10), and (12)–(14) has complex nonlinear dynamics whose analysis is made even more challenging by the combination of the Brownian financial with the two Poisson health shocks. Indeed, the presence of the latter makes analytical solutions of the pair of stochastic differential equations (dH, dW) intractable, and we will therefore restrict our analysis to conditional (upon observing H) expected local changes instead.¹⁷ Noting that the expected net return dM_{st} on actuarially fair insurance contracts (5) is zero reveals that the expected changes in health and wealth are:

$$E[dH] = \left[I^h(W, H)^\alpha - \tilde{\delta} \right] H dt, \tag{16}$$

$$E[dW] = [rW + Y(H) - c(W, H) - I(W, H) + \pi(W, H)\sigma_S\theta] dt, \tag{17}$$

where $I^h(W, H) = I(W, H)/H$ is the investment-to-health capital ratio, and $\tilde{\delta} = \delta + \phi\lambda_{s0}$ is the sickness-adjusted expected depreciation rate. Since our main focus concerns end-of-

¹⁷See also Laporte and Ferguson (2007) for analysis of expected local changes of the Grossman (1972) model with Poisson shocks.

life decumulation, the local expected changes (16), and (17) can be relied upon to define *depletion* regions of the admissible state space where health and wealth are (locally) expected to fall:

Definition 2 (depletion) Health, and wealth depletion regions $(\mathcal{D}_H, \mathcal{D}_W) \subseteq \mathcal{A}$ are characterized by optimal expected depletion of the health and wealth stocks:

$$\mathcal{D}_H = \{(W, H) \in \mathcal{A} : E[dH] < 0\},$$

$$\mathcal{D}_W = \{(W, H) \in \mathcal{A} : E[dW] < 0\}.$$

The following result relies on intuitive conditions to further characterize the depletion regions of the state space.

Theorem 1 (depletion) *Assume that the regularity, and transversality conditions (31), (32), and (33) hold.*

1. *If the following conditions hold:*

$$y_0 < a, \tag{18}$$

$$BK_0 < \tilde{\delta}^{1/\alpha}. \tag{19}$$

Then the health depletion zone is given by:

$$\mathcal{D}_H = \{(W, H) \in \mathcal{A} : W < y(H) = x(H) + DH^{1+\xi_m}\}, \tag{20}$$

where,

$$D = K_m^{-1} [\tilde{\delta}^{1/\alpha} - BK_0] > 0. \tag{21}$$

2. *If, in addition the following conditions hold:*

$$\beta < B(r + K_0), \tag{22}$$

$$\frac{\theta^2}{\gamma} + r < A, \tag{23}$$

then the wealth depletion zone is given by:

$$\mathcal{D}_W = \mathcal{A}. \quad (24)$$

The conditions (18), (19), (22), and (23) are economically plausible and particularly relevant for end-of-life analysis. Conditions (18), and (23) both refer to high consumption patterns, with the former implying that base income y_0 in (4) is insufficient to cover subsistence consumption a in (8), and the latter implying high marginal propensity to consume A in (12). We can use the closed-form expression (38) to rewrite condition (23) as:

$$A - r - \frac{\theta^2}{\gamma} = \varepsilon(\rho - r) - (1 - \varepsilon)\frac{\lambda_{m0}}{1 - \gamma_m} - (1 + \varepsilon)\frac{\theta^2}{2\gamma} > 0.$$

Since $\gamma_m \in [0, 1)$, and assuming (as will be verified later) that the elasticity of intertemporal substitution $\varepsilon > 1$, the condition (23) of a high marginal propensity to consume obtains when the agent is impatient, i.e. ρ is high, and/or the unconditional risk of dying λ_{m0} is high, and/or the aversion to death risk γ_m is high.

Condition (19) states that expected health depreciation $\tilde{\delta}$ is high, while condition (22) requires a low ability β of healthier agents to generate labor revenues. Intuitively, the expression $(\tilde{\delta}^{1/\alpha} - BK_0)$ in (21) captures the order-0 expected depletion, i.e. in the absence of endogenous mortality. When the latter is reintroduced, optimal investment in (10) is larger, reflecting the additional demand for death risk hedging provided by health capital. If condition (19) 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 in the admissible range.

The optimal health and wealth dynamics characterized by Theorem 1 are plotted in Figure 2. First, the admissible region \mathcal{A} is bounded below by the $W = x(H)$ locus in red, with complementary non-admissible area $\mathcal{N}\mathcal{A}$ in shaded red region. The W -intercept is given by the NPV of base income deficit $-C$ which is positive under assumption (18). The H -intercept is given by $\bar{H}_1 = -C/B > 0$. Second, equation (20) in Theorem 1 states that the health depletion region \mathcal{D}_H is the shaded green area located below the green $W = y(H)$ locus. Both $x(H), y(H)$ loci intersect at the same $-C$ intercept. A sufficiently high depreciation $\tilde{\delta}$ in (19) entails a positive constant $D > 0$ in (20). Consequently the

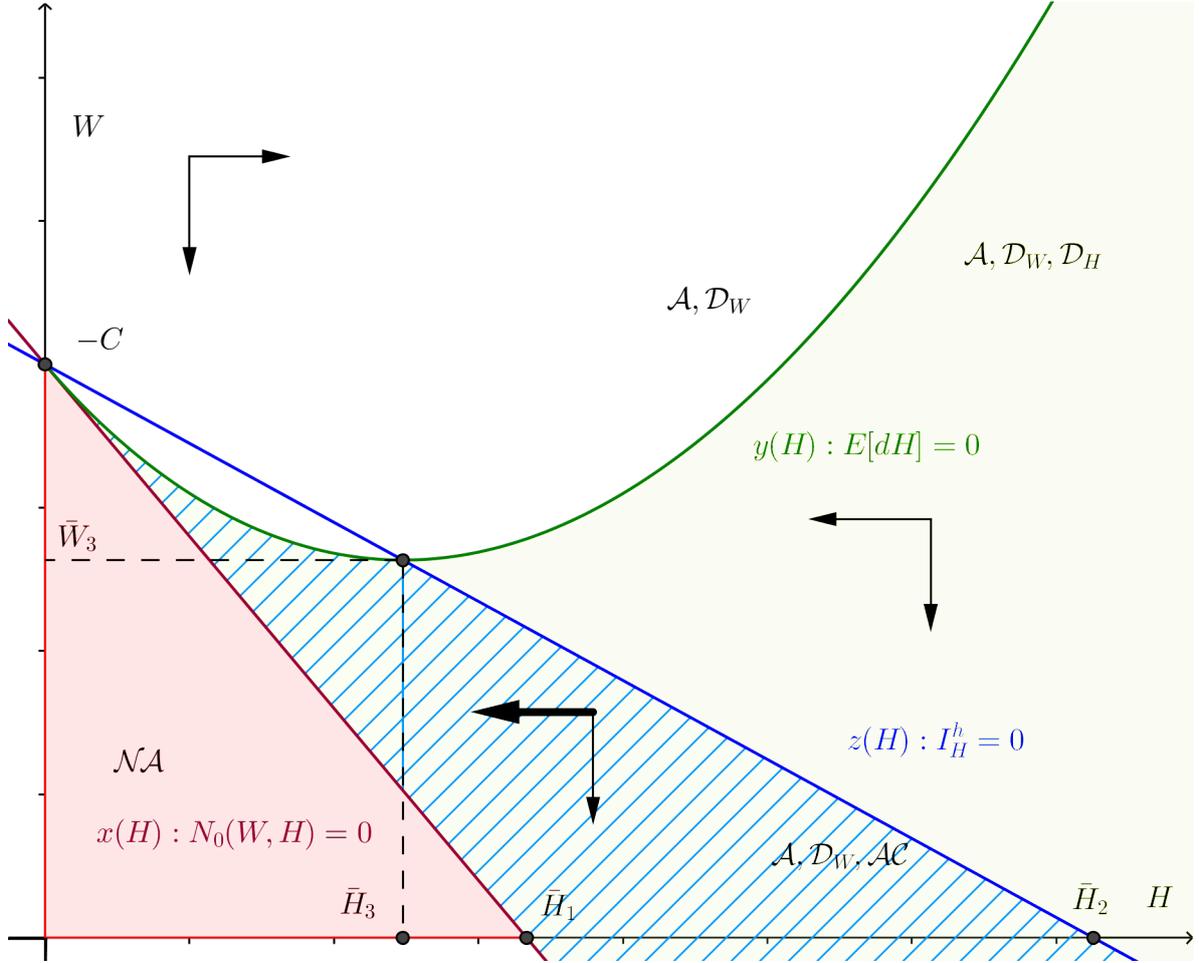


Figure 2: Health and wealth dynamics

Notes: The shaded area in red is the non-admissible set \mathcal{NA} (Definition 1). The depletion area \mathcal{D} (Theorem 1) is the shaded green area under the green curve. The accelerating region \mathcal{AC} (Theorem 2) corresponds to the shaded green area hatched with blue lines.

$y(H)$ locus is U-shaped, and attains a unique minimum at \bar{H}_3 given by:

$$\bar{H}_3 = \left(\frac{B}{D(1 + \xi_m)} \right)^{\frac{1}{\xi_m}} > 0. \quad (25)$$

Finally, equation (24) in Theorem 1 states that the wealth depletion region \mathcal{D}_W boils down to the entire admissible set \mathcal{A} . We will return to the interpretation of the third locus $W = z(H)$ plotted in blue in Theorem 2 below.

To see why the $W = y(H)$ locus is non-monotone, observe from (16) that expected change in health $E[dH]$ increases in the investment-to-health ratio, where optimal invest-

ment (10) reveals that the latter is:

$$I^h(W, H) = BK_0 + K_m H^{-\xi_m - 1} N_0(W, H). \quad (26)$$

This ratio 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 $K_m H^{-\xi_m - 1} = \bar{K}_m \lambda_{m1} H^{-\xi_m - 1}$, and therefore also reduces I^h . Therefore, constant zero growth requires increasing W . The analysis of the $W = y(H)$ locus in (39) 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 ($H > \bar{H}_3$).

The local expected dynamics are represented by the directional arrows in Figure 2. First, equation (20) implies that only agents who are sufficiently rich (i.e. $W > y(H)$) can expect a growth in health; all others are located in the \mathcal{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, under equation (24), the wealth depletion is the entire admissible set such that *all* agents, regardless of their health or wealth levels, expect wealth to fall. Taken together, these results suggest an optimal depletion of both human and financial capital with wealth eventually falling into the \mathcal{D}_H region, and ensuing health depletion. From endogenous death intensity (3), falling health is invariably accompanied by an increase in mortality, and a decline towards the non-admissible locus $W = x(H)$ characterized by zero net total wealth, and indifference between life and death, i.e. $V(W, H) = 0$.

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 (10) 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.

Interestingly, it is also possible to characterize differences in how fast the health capital is allowed to deplete. To do so, we can define an *acceleration* subset in the

health depletion region whereby the investment-to-health ratio is an increasing function of health. Consequently, a depletion of the health capital leads to a decrease in I^h , and thus accelerating health depletion in (16). More precisely,

Definition 3 (acceleration) An accelerating zone $\mathcal{AC} \subset \mathcal{D}_H$ is a health depletion subset where the investment to health ratio $I^h(W, H)$ increases in health:

$$\mathcal{AC} = \{(W, H) \in \mathcal{D}_H : I^h(W, H) > 0\}$$

Relying on the optimal investment-to-health ratio (26) allows us to obtain the following result:

Theorem 2 (acceleration) *Assume that the conditions of Theorem 1 hold. Then the accelerating region is given by:*

$$\mathcal{AC} = \begin{cases} \mathcal{D}_H, & \text{if, } H < \bar{H}_3 \\ \left\{ (W, H) \in \mathcal{D}_H : W < z(H) = x(H) + \frac{BH}{1+\xi_m} \right\}, & \text{otherwise} \end{cases} \quad (27)$$

The accelerating locus $W = z(H)$ is plotted as the blue line in Figure 2; the accelerating region is the dashed blue subset of \mathcal{D}_H . It is straightforward to show that this locus intersects the $x(H), y(H)$ loci at the same $-C$ intercept, that it intersects the H -axis at $\bar{H}_2 = \bar{H}_1(1 + \xi_m)/\xi_m$, and finally and that it also intersects the health depletion locus $y(H)$ at its unique minimal value \bar{H}_3 in (25).

It follows from the expected health growth (16), and the characterization of the accelerating region \mathcal{AC} in equation (27) that agents in the health depletion region \mathcal{D}_H optimally slow down (but do not reverse) the depreciation of their health capital only if sufficiently rich and healthy ($W > z(H)$). Otherwise, for $(W, H) \in \mathcal{AC}$, the health depletion accelerates (illustrated by the thick directional vector) as falling health is accompanied by further cuts in the investment-to-health ratio.

These dynamics thus suggest optimal *closing down the shop* behavior whereby falling health is initially optimally fought back, before eventually being accelerated. Importantly, regardless of whether it is accelerating or not, the optimal descent of health and increased exposure to death risk for those agents in the health depletion region obtains even when life is strictly preferred. Indeed, as shown in Hugonnier et al. (2013), and discussed earlier,

the non-separable preferences (7) 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.

Such a closing-down strategy of optimal wealth depletion, and eventual health depletion is arguably more appropriate for agents nearing death, than for younger ones. Indeed, a base income deficit relative to subsistence consumption (condition (18)), and a high marginal propensity to consume (condition (23)) are suitable for elders nearing end of life, with large demand for comfort care, and without deliberate bequest motives. Moreover, a high sickness-augmented depreciation rate for the health capital (condition (19)), and a low ability to generate labor revenues (condition (22)) both seem legitimate for old agents in the last period of life, yet less so for younger ones. The next section verifies empirically whether or not these conditions are valid.

4 Empirical evaluation

The structural econometric model that we rely upon to (i) estimate the deep parameters and (ii) evaluate the induced parameters (B, C, D, \bar{H}_i for $i = 1, 2, 3$, and \bar{W}_3) 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 $j = 1, 2, \dots, n$ is the optimal investment (10), and the risky asset holdings (13), to which we append the income equation (4) :

$$I_j = K_0 B H_j + K_m H_j^{-\xi_m} N_0(W_j, H_j) + u_{Ij}, \quad (28)$$

$$\pi_j = L_0 N_0(W_j, H_j) + u_{\pi j}, \quad (29)$$

$$Y_j = y_0 + \beta H_j + u_{Yj}, \quad (30)$$

where the u_j are (potentially correlated) error terms. Data limitations discussed below explain why optimal consumption (12), and insurance (14) 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 does not appear unreasonable to the extent that we are considering a relatively homogeneous subset of old individuals, thereby ruling out potent cohort effects. The joint estimation of (28), (29) and (30) is undertaken with respect to the deep parameters, under the theoretical restrictions governing K_0, K_m, L_0 , as well as B, C , and also subject to the regularity conditions (31), (32), and (33).

The identification of the deep parameters is complicated by the significant nonlinearities that are involved. Consequently, not all the parameters can be estimated, and a subset was therefore calibrated. Of those, certain parameters could be set at standard values from the literature. For others however, scant information was available, and we relied on thorough robustness analysis, especially with respect to γ_m , and ϕ . These alternative estimates, which are available upon request, are reasonably robust, with main interpretations qualitatively unaffected.

The estimation approach is an iterative two-step procedure. In a first step, the convexity parameters (ξ_m, ξ_s) are fixed and a maximum likelihood approach is conducted on the remaining structural parameters. In a second step, the structural parameters are fixed and the maximum likelihood function is maximized with respect to ξ_m and ξ_s . 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 (investment, portfolio, and income). 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 risky holdings π_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.¹⁸

¹⁸Note however that our structural model neither rules out zero holdings, nor does predict a Tobit-based specification for the portfolio equation.

4.2 Data

The data base used for estimation is the 2002 wave of the Health and Retirement Study (HRS, Rand data files). A main reason of using this HRS wave is that it is the last one with detailed information on 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, and we therefore resort to the 2002 HRS wave. Also, even though the HRS contains individuals aged 51 and over, we restrict our analysis to elders (i.e. agents aged 65 and more). In doing so, we avoid endogenizing the insurance choice X_t in (5) which, under Medicare coverage, can be considered as exogenous. 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) π_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 \$1,000,000 applied to all nominal variables (I_j, W_j, π_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,

¹⁹Self-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).

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

and falls sharply in health.²¹ Conversely, risky holdings increase sharply in wealth, and are also higher for healthier agents.²²

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 1, and 2. The standard errors indicate that all the estimates 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 ($1 - \exp(-\lambda_{s0}) = 25\%$). The death intensity (3) parameters are realistic, with an expected lifetime of 79.0 years for an individual with an average (i.e. good) health.²³ Importantly, the null of exogenous exposure to death risk is rejected ($\lambda_{m1}, \xi_m \neq 0$), indicating that agent’s health decisions are consequential for their expected life horizon. Taken together, these law of motion and risk exposure parameters compare well to estimates in Hugonnier et al. (2013), and are consistent with expectations regarding an elders’ population.

²¹Similar 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.

²²Similar 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).

²³In particular, Hugonnier et al. (2013) show that an age- t person’s remaining life expectancy can be computed using:

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

The average age in our HRS sample is 75.3 years, and the expected remaining life horizon is 3.7 years for an individual with good health. The unconditional expected lifetime was 77.3 years in 2002, with 74.5 for males, and 79.9 for females (Arias, 2004).

Third, the returns parameters (μ, r, σ_S) are calibrated at standard values in panel c. The income parameters of equation (4) 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 larger than base income y_0 . Both subsistence, and base income values are realistic.²⁴ Our estimate of the inter-temporal elasticity ε is larger than one, as identified by others using micro data.²⁵ Aversion to financial risk is realistic ($\gamma = 2.78$), whereas aversion to mortality risk is calibrated in the admissible range ($0 < \gamma_m < 1$), and close to the value set by Hugonnier et al. (2013) ($\gamma_m = 0.75$). The aversion to morbidity risk γ_s is both unidentifiable and irrelevant under the exogenous morbidity risk assumption, and in the absence of endogenous demand for insurance. 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

Panel e of Table 2 reports the induced parameters that are relevant for the admissible, depletion and accelerating subsets; panel f shows that three out of the four corresponding conditions in Theorem 1 are verified, with the fourth being non-significant. These composite parameters allow us to evaluate the position of the loci $x(H), y(H), z(H)$, and thus of the various subsets in Figure 3. The H axis also records the positions associated with Poor ($H = 0.5$), and Fair ($H = 1.25$) self-reported health statuses, where the scaling is the one used in the estimation. The W axis is reported in \$M, using the same scaling as for the estimation. The health and wealth joint distribution for the HRS data is indicated by plotting the median for wealth associated with each quintiles, as blue points for Q_2, Q_3, Q_4 , (Q_1, Q_5 omitted) for each health statuses.

First, we identify a relatively large marginal- Q of health $B = 0.1148$ in panel e, suggesting that health is very valuable.²⁶ Second, the large negative value for C corre-

²⁴For example, the 2002 poverty threshold for elders above 65 was \$8,628 (source: U.S. Census Bureau).

²⁵For example, Gruber (2013) finds estimates centered around 2.0, relying on CEX data. In our case, the recourse to elders' data, and the assumption of no bequest function could explain a relatively strong consumption reaction to interest rates movements.

²⁶Adapting 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 $w_h(0.10, W, H) = 0.10 * B * 10^6 = \$11,480$ (\$15,150 in 2016).

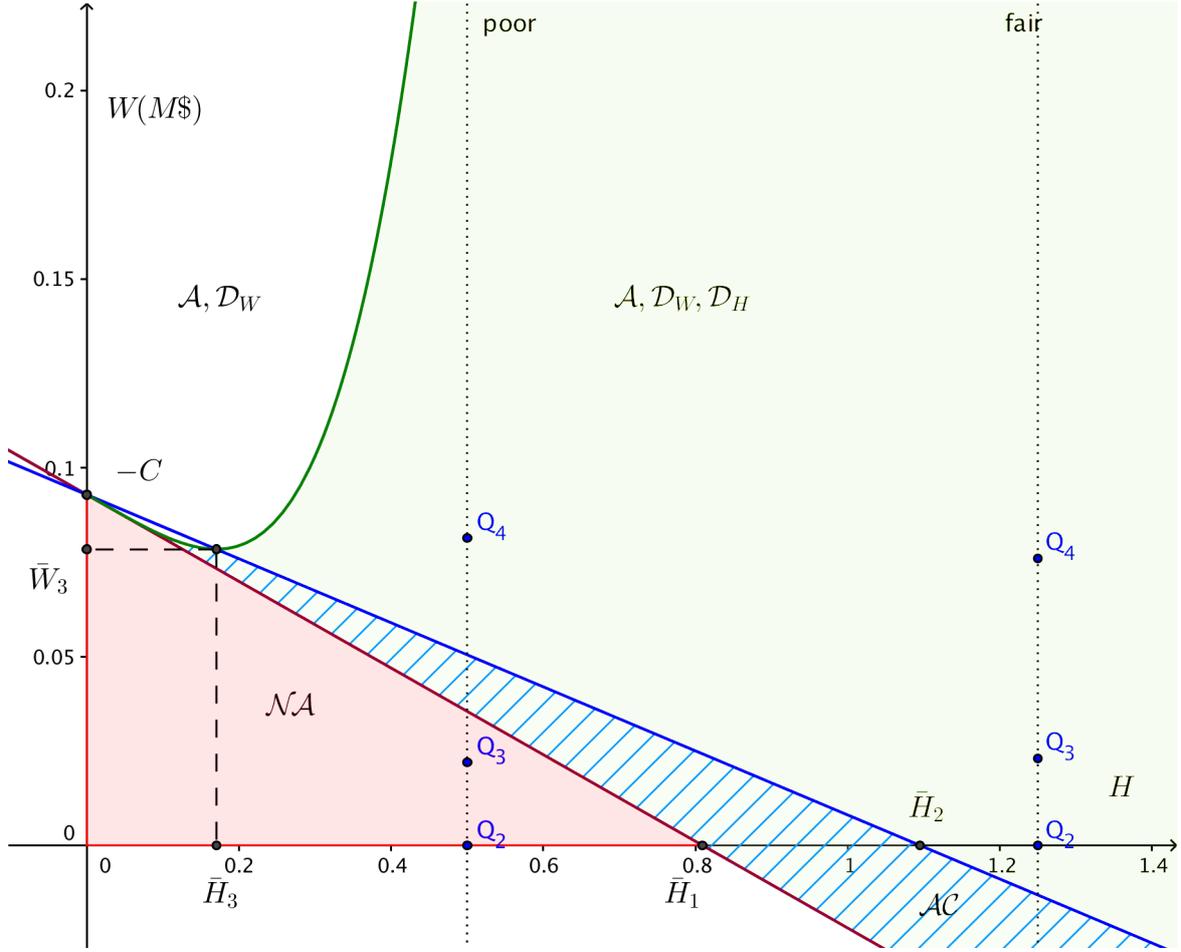


Figure 3: Estimated depletion, accelerating, and non-admissible regions

Notes: The shaded area in red is the non-admissible set \mathcal{NA} (Definition 1). The depletion area \mathcal{D} (Theorem 1) is the shaded green area under the green curve. The accelerating region \mathcal{AC} (Theorem 2) corresponds to the shaded green area hatched with blue lines. Position of loci, and areas evaluated at estimated parameters in Table 2. Median levels for wealth quintiles Q_2, \dots, Q_4 (Q_1, Q_5 not reported) are taken from Table 1, and are reported as blue points for health levels poor, and fair.

sponds to a capitalised base income deficit of 92,900\$ (122,628\$ in 2016), and confirms that condition (18) in Theorem 1 is verified. Third, the value for D is significant which confirms the verification of condition (19). From the definition of $y(H)$ in (20), a large value of D also entails a very steep health depletion locus. It follows that unless very wealthy, and very unhealthy, the bulk of the population would be located in the health and wealth depletion regions. Besides being consistent with expectations regarding agents near the end of life, the results also rationalize better longevity for the rich (Bosworth et

al., 2016; Bosworth and Zhang, 2015). Indeed, very rich and sick agents who are not in the health depletion region will select expected increase in the health capital as long as their wealth maintains them out of the \mathcal{D}_H region.

Finally, our estimates are consistent with a narrow accelerating region \mathcal{AC} . Indeed, the values for B, C, ξ_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. Importantly, it is also consistent with a change in composition in health expenses towards more comfort, and less curative care (De Nardi et al., 2015b; Marshall et al., 2010).

5.3 Simulation analysis

The dynamic analysis presented thus far has focused upon local expected changes for health and wealth $E[dH], E[dW]$. At this stage there is no clear indication that such small anticipated depletions will translate into *bona fide* life cycle declining paths for health and wealth. To verify whether they do, we conduct a Monte-Carlo simulation exercise 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 a common uniform distribution for health, $H_0 \sim U[0.5, 3.5]$, and two different distributions for wealth:
 - Poor: $W_0 \sim U[0.01, 0.10]$;
 - Rich: $W_0 \sim U[0.25, 1.50]$.
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 $t = 75, \dots, 85$.
3. At each time period $t = 75, \dots, 85$, 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_{st}, dZ_t in the stochastic laws of motion dH_t, dW_t .
- (d) We update the health and wealth variables using the Euler approximation:

$$H_{t+1} = H_t + dH_t(H_t, I_t, dQ_{st})$$

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

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

Figure 4 plots the resulting mean values for the optimal life cycles for financial wealth W_t (panel a), net total wealth $N_0(W_t, H_t)$ (panel b), health level H_t (panel c), using only the alive, and admissible agents, with positive financial wealth. We also report the shares of the surviving, admissible population in the health depletion, and accelerating regions (panel d), as well as the exposure to death risk $1 - \exp[-\lambda_m(H_t)]$ (panel e).

Overall, these results provide additional evidence in favor of our previous findings. Consistent with the data, our simulated life cycles feature a rapid 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) as they enter the end of life period. Indeed, recalling that expected longevity 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 \mathcal{A}) at terminal age (panel b), an objective obtained by running down wealth (panel a) very rapidly, consistent with our finding that $\mathcal{D}_W = \mathcal{A}$, and a somewhat slower decline for health (panel c). The resulting subset shares in panel d confirm that virtually all the population is in the health depletion region at an early stage, and a sizable share enters the accelerating region halfway through.

Contrasting rich versus poor cohorts reveals that, as expected, wealth (panel a), and health (panel c) depletion is faster for poor agents, such that low-wealth individuals enter

the depletion, and accelerating regions more rapidly (panel d). Moreover, exposure to death risk is higher for the poor (panel e), consistent with stylized facts (Bosworth et al., 2016; Bosworth and Zhang, 2015), except at very old age where attrition effects imply that only the very healthy poor agents remain alive, and the rich and poor exposures to mortality are converging. Put differently, our simulations indicate that agents entering the last period of life optimally select a short expected lifespan, and 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.

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 they are indifferent between life and death. We rely on closed-form 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 certain 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 or not the agent is. Moreover, this depletion is accelerated below certain levels of health and wealth. Importantly, wealth is expected to fall regardless of the health status, such that all agents eventually close down the shop.

The previous analysis suggests a policy role in reducing the incidence of depletion regions of the state space. In particular, such a reduction is readily achieved by increasing base income (e.g. through enhanced Social Security, Medicaid, or minimal revenue programs),²⁷ or via subsidized improvements in medical technology.²⁸

²⁷To see this, observe from equation (25) that the health threshold \bar{H}_3 is unaffected by the intercept $-C$, whereas the wealth threshold $\bar{W}_3 = z(\bar{H}_3)$ increases in the latter. Consequently, increasing base income y_0 directly lowers the health deficit $-C$, and consequently also the wealth threshold, and therefore the prevalence of health depletion.

²⁸Hence, improvements that result in less sickness-adjusted depreciation $\tilde{\delta} = \delta + \lambda_{s0}\phi$ have a direct effect in lowering D , and therefore how steep the $y(H)$ locus is evaluated. Again the prevalence of \mathcal{D}_H would be reduced.

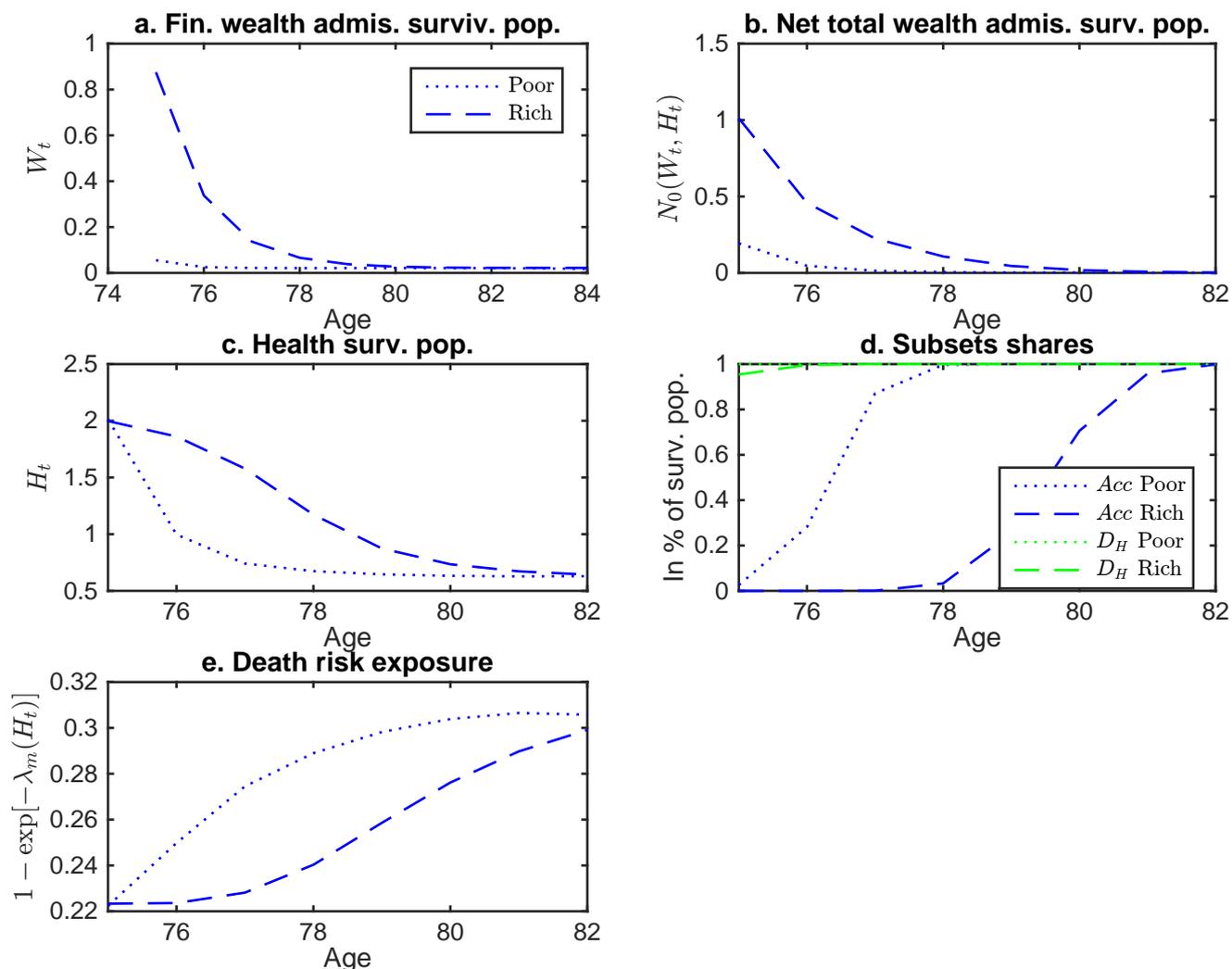


Figure 4: Simulated optimal paths

Notes: Mean values for simulated optimal life cycles taken over an initial population of 1,000 agents with 1,000 replications. Initial draw from rich (dashed lines), and poor (dotted lines) initial populations. Mean values in panels a–c are taken with respect to surviving, admissible agents, with positive financial wealth. Subset shares taken as percentage of admissible surviving population in each subset.

However, whereas the positive arguments are readily obtained, the normative reasons for intervening are 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, assuming away policy changes in base income,

state intervention on the public health domain in order to minimize unmet medical needs may also be questioned if, as the theory, and empirical evaluation suggest, failure to seek treatment is the result of an optimal dynamic decision by individuals.

References

- Acemoglu, Daron, Amy Finkelstein, and Matthew J. Notowidigdo (2013) ‘Income and health spending: Evidence from oil price shocks.’ *Review of Economics and Statistics* 95(4), 1079 – 1095
- Adams, Peter, Michael D. Hurd, Daniel McFadden, Angela Merrill, and Tiago Ribeiro (2003) ‘Healthy, wealthy, and wise? tests for direct causal paths between health and socioeconomic status.’ *Journal of Econometrics* 112(1), 3–56
- American Medical Association (2016) ‘AMA policy on provision of life-sustaining medical treatment.’ AMA website
- Arias, Elizabeth (2004) ‘United States life tables, 2004.’ *National Vital Statistics Report* 53(6), 1–40
- (2014) ‘United States life tables, 2010.’ *National Vital Statistics Report* 63(7), 1–62
- Ayanian, John Z., Joel S. Weissmann, Eric C. Schneider, Jack A. Ginsburg, and Alan M. Zaslasky (2000) ‘Unmet health needs of uninsured adults in the United States.’ *Journal of the American Medical Association* 284(16), 2061–2069
- 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
- Benítez-Silva, Hugo, and Huan Ni (2008) ‘Health status and health dynamics in an empirical model of expected longevity.’ *Journal of Health Economics* 27(3), 564–584
- Benjamins, Maureen Reindl, Robert A Hummer, Isaac W Eberstein, and Charles B Nam (2004) ‘Self-reported health and adult mortality risk: An analysis of cause-specific mortality.’ *Social Science & Medicine* 59(6), 1297 – 1306

- Berkowitz, Michael K., and Jiaping Qiu (2006) ‘A further look at household portfolio choice and health status.’ *Journal of Banking and Finance* 30(4), 1201–1217
- Blau, David M., and Donna B. Gilleskie (2008) ‘The role of employee health insurance in the employment behavior of older men.’ *International Economic Review* 49(2), 475–514
- Bosworth, Barry P., and Kan Zhang (2015) ‘Evidence of increasing differential mortality: A comparison of the HRS and SIPP.’ Working papers, Center for Retirement Research at Boston College
- Bosworth, Barry P., Gary Burtless, and Kan Zhang (2016) ‘Later retirement, inequality in old age, and the growing gap in longevity between rich and poor.’ Economic Studies at Brookings, Brookings Institute
- Bureau of Labor Statistics (2008) ‘Older workers.’ BLS Spotlight on statistics, U.S. Department of Labor, July
- Carroll, Christopher D. (2002) ‘Portfolios of the rich.’ In *Household Portfolios*, ed. Luigi Guiso, Michael Haliassos, and Tullio Jappelli (Cambridge MA: MIT Press) chapter 10, pp. 389–429
- Case, Anne, and Angus Deaton (2005) ‘Broken down by work and sex: How our health declines.’ In *Analyses in the Economics of Aging*, ed. David A. Wise (Chicago: University of Chicago Press) chapter 6, pp. 185 – 205
- Coile, Courtney, and Kevin Milligan (2009) ‘How household portfolios evolve after retirement: The effect of aging and health shocks.’ *Review of Income and Wealth* 55(2), 226 – 248
- Crossley, Thomas F., and Steven Kennedy (2002) ‘The reliability of self-assessed health status.’ *Journal of Health Economics* 21(3), 643–658
- De Nardi, Mariacristina, Eric French, and John Bailey Jones (2009) ‘Life expectancy and old age savings.’ *American Economic Review* 99(2), 110–115
- (2010) ‘Why do the elderly save? The role of medical expenses.’ *Journal of Political Economy* 118(1), 39–75

- (2015a) ‘Savings after retirement: A survey.’ Working Paper 21268, National Bureau of Economic Research, June
- De Nardi, Mariacristina, Eric French, John Bailey Jones, and Angshuman Gooptu (2012) ‘Medicaid and the elderly.’ *Federal Reserve Bank of Chicago Economic Perspectives* 36(1), 17–34
- De Nardi, Mariacristina, Eric French, John Bailey Jones, and Jeremy McCauley (2015b) ‘Medical spending of the U.S. elderly.’ Working Paper 21270, National Bureau of Economic Research, June
- DiMatteo, Livio (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, Darrell, and Larry G. Epstein (1992) ‘Stochastic differential utility.’ *Econometrica* 60(2), 353–394
- Fan, Elliott, and Ruoyun Zhao (2009) ‘Health status and portfolio choice: Causality or heterogeneity?.’ *Journal of Banking and Finance* 33(6), 1079 – 1088
- French, Eric, and John Bailey Jones (2011) ‘The effects of health insurance and self-insurance on retirement behavior.’ *Econometrica* 79(3), 693–732
- French, Eric, Mariacristina De Nardi, John Bailey Jones, Olesya Baker, and Phil Doctor (2006) ‘Right before the end: Asset decumulation at the end of life.’ *Federal Reserve Bank of Chicago Economic Perspectives* 30(3), 2 – 13
- Gilleskie, Donna B., and Thomas A. Mroz (2004) ‘A flexible approach for estimating the effects of covariates on health expenditures.’ *Journal of Health Economics* 23(2), 391 – 418
- Goldman, Dana, and Nicole Maestas (2013) ‘Medical expenditure risk and household portfolio choice.’ *Journal of Applied Econometrics* 28(4), 527 – 550
- Grossman, Michael (1972) ‘On the concept of health capital and the demand for health.’ *Journal of Political Economy* 80(2), 223–255

- Gruber, Jonathan (2013) ‘A tax-based estimate of the elasticity of intertemporal substitution.’ *Quarterly Journal of Finance* 3(1), 1 – 20
- Guiso, Luigi, Tullio Jappelli, and Daniele Terlizzese (1996) ‘Income risks, borrowing constraints, and portfolio choice.’ *American Economic Review* 86(1), 158–172
- Hall, Robert E., and Charles I. Jones (2007) ‘The value of life and the rise in health spending.’ *Quarterly Journal of Economics* 122(1), 39–72
- Heiss, Florian (2011) ‘Dynamics of self-rated health and selective mortality.’ *Empirical Economics* 40(1), 119 – 140
- Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes (1994) ‘Expanding the life-cycle model: Precautionary saving and public policy.’ *American Economic Review* 84(2), 174 – 179
- (1995) ‘Precautionary saving and social insurance.’ *Journal of Political Economy* 103(2), 360 – 399
- Hugonnier, Julien, Florian Pelgrin, and Pascal St-Amour (2013) ‘Health and (other) asset holdings.’ *The Review of Economic Studies* 80(2), 663–710
- Hurd, Michael D. (1987) ‘Savings of the elderly and desired bequests.’ *American Economic Review* 77(3), 298–312
- (2002) ‘Are bequests desired or accidental.’ Working Paper Series DRU-3010, RAND, January
- Hurd, Michael D., and Kathleen McGarry (1995) ‘Evaluation of the subjective probabilities of survival in the Health and Retirement Study.’ *Journal of Human Resources* 30(0), s268–s292. Supplement
- (2002) ‘The predictive validity of subjective probabilities of survival.’ *Economic Journal* 112(482), 966–985
- Hurd, Michael D., Daniel McFadden, and Angela Merrill (2001) ‘Predictors of mortality among the elderly.’ In ‘Themes in the economics of aging’ NBER Conference Report series (Chicago and London: University of Chicago Press) pp. 171–97

- Laporte, Audrey, and Brian S. Ferguson (2007) ‘Investment in health when health is stochastic.’ *Journal of Population Economics* 20(2), 423 – 444
- Lee, Jinkook, and Hyungsoo Kim (2008) ‘A longitudinal analysis of the impact of health shocks on the wealth of elders.’ *Journal of Population Economics* 21(1), 217 – 230
- Legal Advisors Committee of Concerns for Dying (1983) ‘The right to refuse treatment: A model act.’ *American Journal of Public Health* 73(3), 918–921
- Marshall, Samuel, Kathleen M. McGarry, and Jonathan S. Skinner (2010) ‘The risk of out-of-pocket health care expenditure at end of life.’ Working paper 16170, National Bureau of Economic Research
- Meer, Jonathan, Douglas L. Miller, and Harvey S. Rosen (2003) ‘Exploring the health-wealth nexus.’ *Journal of Health Economics* 22(5), 713–730
- Michaud, Pierre-Carl, and Arthur van Soest (2008) ‘Health and wealth of elderly couples: Causality tests using dynamic panel data models.’ *Journal of Health Economics* 27(5), 1312–1325
- National Center for Health Statistics (2012) *Health, United States, 2011: With Special Feature on Socioeconomic Status and Health* (Hyattsville, MD: United States Department of Health and Human Services. Centers for Disease Controls and Prevention)
- Palumbo, Michael G. (1999) ‘Uncertain medical expenses and precautionary saving near the end of the life cycle.’ *Review of Economic Studies* 66(2), 395–421
- 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
- Pelgrin, Florian, and Pascal St-Amour (2016) ‘Life cycle responses to health insurance status.’ *Journal of Health Economics*. Forthcoming
- Philipson, Tomas J., Gary Becker, Dana Goldman, and Kevin M. Murphy (2010) ‘Terminal care and the value of life near its end.’ Working paper 15649, National Bureau of Economic Research

- Poterba, James, Steven Venti, and David A. Wise (2015) ‘What determines end-of-life assets? a retrospective view.’ Working Paper 21682, National Bureau of Economic Research, October
- Rosen, Harvey S., and Stephen Wu (2004) ‘Portfolio choice and health status.’ *Journal of Financial Economics* 72(3), 457–484
- Skinner, Jonathan (2007) ‘Are you sure you’re saving enough for retirement?’ *Journal of Economic Perspectives* 21(3), 59 – 80
- Smith, James P. (1999) ‘Healthy bodies and thick wallets: The dual relation between health and economic status.’ *Journal of Economic Perspective* 13(2), 145–166
- (2007) ‘The impact of socioeconomic status on health over the life-course.’ *Journal of Human Resources* 42(4), 739 – 764
- Toossi, Mitra (2015) ‘Labor force projections to 2024: the labor force is growing, but slowly.’ *Monthly Labor Review*
- Tsai, Yuping (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, Hans, Tom Van Ourti, Owen O’Donnell, and Eddy van Doorslaer (2009) ‘Health and income across the life cycle and generations in Europe.’ *Journal of Health Economics* 28, 818–830
- Wachter, Jessica A., and Motohiro Yogo (2010) ‘Why do household portfolio shares rise in wealth?.’ *Review of Financial Studies* 23(11), 3929 – 3965
- Yogo, Motohiro (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, September

A Parametric restrictions

A.1 Regularity and transversality restrictions

The theoretical model is solved under three regularity and transversality conditions that are reproduced for completeness. To do so, define the following functions:

$$\chi(x) = 1 - (1 - \phi)^{-x},$$

$$F(x) = x(\alpha B)^{\frac{\alpha}{1-\alpha}} - x\delta - \lambda_{s0}\chi(-x),$$

$$L_m(H) = ((1 - \gamma_m)(A - F(-\xi_m)))^{-1} H^{-\xi_m},$$

and assume that the following regularity and transversality conditions hold:

$$\beta < (r + \delta + \phi\lambda_{s0})^{\frac{1}{\alpha}}, \quad (31)$$

$$\max\left(0; r - \frac{\lambda_{m0}}{1 - \gamma_m} + \theta^2/\gamma\right) < A, \quad (32)$$

$$0 < A - \max(0, r - \nu_{m0} + \theta^2/\gamma) - F(-\xi_m), \quad (33)$$

where the shadow price of health B , and the marginal propensity to consume A are defined below.

A.2 Closed-form solutions for optimal rules parameters

The closed-form expression for the parameters in the optimal rules are given as follows.

The positive parameters of the optimal investment in (10) are:

$$K_0 = \alpha^{\frac{1}{1-\alpha}} B^{\frac{\alpha}{1-\alpha}}, \quad (34)$$

$$K_m = \lambda_{m1} \frac{\xi_m K_0 L_m(1)}{1 - \alpha} \quad (35)$$

where the shadow price B of health solves:

$$g(B) = \beta - (r + \delta + \phi\lambda_{s0})B - (1 - 1/\alpha)(\alpha B)^{\frac{1}{1-\alpha}} = 0 \quad (36)$$

subject to $g'(B) < 0$, and the NPV of excess base income in (11) is:

$$C = \frac{y_0 - a}{r}. \quad (37)$$

The other parameters include the marginal propensity to consume in (12):

$$A = \varepsilon\rho + (1 - \varepsilon) \left(r - \frac{\lambda_{m0}}{1 - \gamma_m} + \frac{\theta^2}{2\gamma} \right), \quad (38)$$

the risky portfolio share in (13):

$$L_0 = \frac{\theta}{\gamma\sigma_S},$$

and the positive marginal value of total wealth in (15) :

$$\Theta = \rho(A/\rho)^{1/(1-\varepsilon)}.$$

B Proof Theorem 1

B.1 Health depletion \mathcal{D}_H

First, substituting the optimal investment (10) in the expected local change for health (16), and using the definition of net total wealth (11) shows that:

$$E[dH] = 0 \iff W = x(H) + DH^{1+\xi_m} = y(H).$$

with expected depletion for $W < y(H)$. Second, observe that condition (18) implies that $-C > 0$ in (37), whereas condition (19) implies that $D > 0$ in (20). It follows directly that $y(H) \geq x(H), \forall H$, i.e. the locus $y(H)$ lies everywhere in the admissible zone, and is

characterized by:

$$y_H(H) = -B + (1 + \xi_m)DH^{\xi_m} \begin{cases} < 0, & \text{if } H < \bar{H}_3, \\ = 0, & \text{if } H = \bar{H}_3, \\ > 0, & \text{if } H > \bar{H}_3, \end{cases} \quad \text{and} \quad (39)$$

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

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 (25), with corresponding wealth level $\bar{W}_3 = y(\bar{H}_3)$.

B.2 Wealth depletion \mathcal{D}_W

Substituting the optimal investment (10), consumption (12), risky portfolio (13), and insurance (14) in the expected local change for wealth (17), and using the definition of net total wealth (11) reveals that

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

where

$$l(H) = [A + K_m H^{-\xi_m} - \sigma_S \theta L_0 - r],$$

$$k(H) = (y_0 - a) + H(\beta - BK_0).$$

Observe that since $K_m > 0$, condition (23) is sufficient to guarantee that $l(H) > 0, \forall H$. Consequently, the wealth depletion zone \mathcal{D}_W is delimited by:

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

We now have to show that this locus lies everywhere in the \mathcal{NA} region:

$$w(H) < x(H) \iff x(H)r + k(H) < 0 \iff \beta < B(r + K_0)$$

as indicated by condition (22). Consequently, the wealth depletion $\mathcal{D}_W \subseteq \mathcal{A}$ coincides with the entire admissible set, i.e. $\mathcal{D}_W = \mathcal{A}$.

C Proof Theorem 2

By a similar reasoning, we can observe from optimal investment (10) that the investment-to-health ratio is given by (26) . Taking the derivative with respect to H and setting to zero shows that the accelerating region obtains as:

$$I_H^h(W_H) > 0 \iff W > -C - \frac{BH\xi_m}{1 + \xi_m} = z(H) = x(H) + \frac{BH}{1 + \xi_m}.$$

Since $B > 0$, this locus lies everywhere above the \mathcal{A} locus, and is therefore admissible.

D Tables

Table 1: HRS data statistics

Variable	Wealth quintile				
	1	2	3	4	5
	a. Poor health ($H = 0.5$)				
Financial wealth	0.000	0.030	0.220	0.814	2.930
Investment	0.379	0.417	0.469	0.427	0.615
Risky holdings	0.005	0.079	0.216	0.485	0.800
	b. Fair health ($H = 1.25$)				
Financial wealth	0.000	0.030	0.230	0.760	3.400
Investment	0.255	0.254	0.233	0.252	0.266
Risky holdings	0.000	0.046	0.253	0.514	0.782
	c. Good health ($H = 2.0$)				
Financial wealth	0.000	0.040	0.220	0.770	3.300
Investment	0.157	0.149	0.156	0.129	0.168
Risky holdings	0.002	0.082	0.299	0.510	0.824
	d. Very good health ($H = 2.75$)				
Financial wealth	0.000	0.040	0.230	0.840	3.500
Investment	0.100	0.112	0.106	0.105	0.107
Risky holdings	0.011	0.107	0.368	0.604	0.854
	e. Excellent health ($H = 3.5$)				
Financial wealth	0.000	0.050	0.210	0.800	3.820
Investment	0.137	0.065	0.063	0.105	0.091
Risky holdings	0.010	0.131	0.350	0.520	0.861

Notes: Median (wealth), and mean values (investment, risky holdings), measured in 100'000\$ (year 2002) per health status, and wealth quintiles for HRS data used in estimation.

Table 2: Estimated and calibrated parameter values

Parameter	Value	Parameter	Value	Parameter	Value
a. Law of motion health (1)					
α	0.6940* (0.1873)	δ	0.0723* (0.0366)	ϕ	0.011 ^c
b. Sickness and death intensities (2), (3)					
λ_{s0}	0.2876* (0.1419)	λ_{m0}	0.2356* (0.0844)		
λ_{m1}	0.0280* (0.0108)	ξ_m	2.8338* (1.1257)		
c. Income and wealth (4), (6)					
y_0	0.0082* ^{\$} (0.0029)	β	0.0141* (0.0059)		
μ	0.108 ^c	r	0.048 ^c	σ_S	0.20 ^c
d. Preferences (7), (8)					
a	0.0127* ^{\$} (0.0063)	ε	1.6738* (0.6846)	γ	2.7832* (1.3796)
ρ	0.025 ^c	γ_m	0.75 ^c	γ_s	N.I.
e. State space subsets (36), (37), (21), (25), (10)					
B	0.1148*	C	-0.0929* ^{\$}	D	4.5088*
\bar{H}_1	0.8093*	\bar{H}_2	1.0460*	\bar{H}_3	0.1743*
K_0	0.0022*	K_m	0.0053*	\bar{W}_3	0.0781* ^{\$}
f. Conditions (18), (19), (22), (23) in Theorem 1					
$y_0 - a$	-0.0045*	$BK_0 - \tilde{\delta}^{1/\alpha}$	-0.0239*		
$\beta - B(r + K_0)$	0.0082	$\theta^2/\gamma + r - A$	-0.5533*		

Notes: *: Estimated structural and induced parameters (standard errors in parentheses), significant at 5% level; *c*: calibrated parameters; *\$*: In \$M; N.I.: non-identifiable/irrelevant under the exogenous morbidity restriction.