Health and (other) asset holdings*

Julien Hugonnier† Florian Pelgrin‡ Pascal St-Amour‡

July 8, 2011

Abstract

Despite strong evidence of correlations between financial and medical statuses and decisions, most of the existing models treat financial and health-related choices separately. This paper bridges this gap by proposing a tractable dynamic framework for the joint determination of optimal consumption, portfolio holdings, medical expenditures and health insurance. We solve for the optimal rules in closed form and capitalize on this tractability to gain a better understanding of the conditions under which separation between financial and health-related decisions is sensible, and of the pathways through which wealth and health determine allocations, welfare and other variables of interest such as expected longevity or the value of health. Furthermore we show that the model is consistent with the observed patterns of individual allocations and provide realistic estimate of the parameters that confirm the relevance of all the main characteristics of the model.

Keywords: Consumption, Health Expenditures and Insurance, Mortality and Morbidity Risks, Portfolio, Values of Life and of Health.

JEL Classification. G11, I12.

*We thank Antoine Bommier, Peter Bossaerts, Bernard Dumas, Darrell Duffie, Pierre-Thomas Léger, Loriano Mancini, Pedro Mira, Norman Schüroff, and Costis Skiadas as well as seminar participants at the DIW-Berlin Conference in Aging, Savings and retirement, the European Winter Finance Summit 2010, the Household Finance and Macroeconomics Conference at the Banco de España in Madrid, the European Summer Symposium in Financial Markets (ESSFM), the DEEP-CIRPÉE-TSE meeting at the University of Toulouse, Gerzensee, the European Central Bank and the University of Lausanne for useful comments and discussions. Excellent research assistance from Louis Alain Signé, Yun Wu and Qing Li, as well as financial support by the Swiss Finance Institute and the Swiss National Center of Competence in Research Financial Valuation and Risk Management (NCCR FinRisk, Projects A2, A4 and A5) is gratefully acknowledged.

†Ecole Polytechnique Fédérale de Lausanne and Swiss Finance Institute. Postal address: SFI@EPFL, Quartier Unil Dorigny, CH–1015 Lausanne, Switzerland. Email: Julien.Hugonnier@epfl.ch

‡Both Pelgrin and St-Amour are with the Faculty of Business and Economics, University of Lausanne, Swiss Finance Institute, CIRANO and CIRPÉE. Postal address: University of Lausanne, CH–1015 Lausanne, Switzerland. Email: (Florian.Pelgrin,Pascal.St-Amour)@unil.ch

1 Introduction

A vast literature on the socioeconomic and health nexus shows that how wealthy and healthy we are has a strong impact on both our financial and health-related decisions.[1] In particular, this literature reveals that health is positively correlated with income, consumption and risky asset holdings, and negatively correlated with health expenditures, whereas it has a mixed effect on insurance coverage. On the other hand, an agent's wealth correlates positively will all these choice variables.

Taken together these stylized facts strongly suggest that any theoretical analysis of financial and health related allocations should be undertaken as that of a *joint* decision problem. Yet, aside from rare exceptions, the two are almost always analyzed separately. At the risk of over-simplifying, health models abstract from financial investment choices whereas health-related considerations are usually absent from financial models. This segmentation might not be so problematic if it could be shown that the two types of decisions are indeed separable. Unfortunately, in the absence of encompassing models, separation cannot even be verified and, thus, should not be taken for granted. Otherwise, segmented models can only provide a partial understanding of the intricate pathways through which wealth and health determine allocations and welfare.

This paper bridges this gap by proposing a tractable dynamic model for the joint determination of consumption, portfolio, health investment and insurance coverage. Our modeling strategy innovates by combining two well-accepted, but otherwise segmented, frameworks from the Financial and Health Economics literatures within a unified setup. More precisely, we start from a standard Merton (1971) portfolio and consumption choice problem with IID returns and append to this model an insurance choice model, as well as a costly health investment decision à la Grossman (1972) in which better health improves labor income as well as reduces the agent's morbidity and mortality risks through a decrease in the arrival rates of the corresponding shocks.

We solve the model analytically and show that it can generate patterns of consumption, portfolio, health expenditures and insurance coverage that are consistent with those observed empirically. In addition, our analytic solution allows us to determine the conditions under which it is sensible to separate financial from health related decisions,

[1] See Smith (1999, 2007, 2009) for an enlightening survey and recent evidence. See also Section 4.2 for cross-sectional evidence from PSID data.

and also provides a natural way of estimating the model. Capitalizing on this feature we estimate the key parameters of the model using cross-sectional data from the Panel Study of Income Dynamics (PSID) and find that our predicted rules are able to fit the data with reasonable parameter values that confirm the relevance of all the model's main characteristics. Importantly, these estimates also indicate that the conditions for separation are not met and therefore justify the need for a joint dynamic analysis of financial and health-related decisions.

As is well-known (e.g. Shepard and Zeckhauser, 1984; Rosen, 1988; Bommier and Rochet, 2006; Bommier, 2010, among others), the specification of preferences is delicate in an endogenous mortality setting such as ours. In the standard time-additive framework of Yaari (1965) and Hakansson (1969) utility is computed as a sum of discounted period utilities up to the random time of death. This associates death with a utility level of zero and, therefore, entails a counterintuitive preference for death over life when the period utility is negative.[2] Our approach to this problem innovates by resorting to a class of recursive preferences that measure utility and consumption in the same metric (Epstein and Zin, 1989; Duffie and Epstein, 1992b). With such preferences death is associated with a consumption level of zero whereas life corresponds to strictly positive consumption and, since preferences are monotonic, it follows that life is always preferred to death, regardless of parameter values. Another distinctive feature of our preference specification is that it assigns distinct risk aversion parameters to each of the three types of risk (financial, morbidity and mortality) present in the model.[3] This feature is referred to as source-dependent risk aversion (Skiadas, 2007, 2009) and our paper constitutes the first application of such preferences to the study of individual consumption, portfolio and health-related choices in a dynamic setting.

In our model, health is subject to diminishing returns to scale and enters the agent's decision problem through two channels. The first channel is referred to as the budget

[2] This is in particular the case for power utility functions with relative risk aversion larger than 1, as is often found in the finance literature, and for negative exponential utility functions. To avoid this outcome, existing solutions include adding a sufficiently large positive constant to utility (see Rosen, 1988; Becker et al., 2005; Hall and Jones, 2007, among others) or simply restricting the relative risk aversion of the power utility function to be smaller than one (Shepard and Zeckhauser, 1984). Another possible solution is to equate death with full depreciation of the health stock and impose Inada conditions on the flow utility of health (e.g., Yogo, 2009).

[3] A further benefit of recursive preferences is that it also disentangles sentiment towards risk from attitudes towards time. This appears particularly relevant in a context where longevity risk can be controlled. Indeed, the elasticity of intertemporal substitution is shown to be a strong determinant of the responses of welfare and consumption to mortality risk.

constraint channel and captures the fact that better health increases labor income, e.g. through less frequent sick leaves and/or better access to promotions for more assiduous workers. This explicit modeling of the health dependence of income departs from standard approaches in which it is assumed that agents get direct utility from being healthy.[4] The second channel is referred to as the risk channel and captures the fact that better health lowers morbidity and mortality risks by reducing the arrival intensities of the corresponding discrete shocks. In this dimension our model is more general than other health risks models that typically consider a single endogenous risk.[5] To gain some intuition about the respective impact of these two channels, we start by abstracting from the second by considering a model in which health risks are exogenous.

In this restricted version of the model, the arrival rates of mortality and morbidity shocks are independent from the agent's health and this feature allows us to derive closed form solutions for the optimal rules. These closed form solutions in turn permit an intuitive interpretation of the underlying economic mechanisms, and show that separating financial and health-related decisions is sensible under exogenous health risks. More precisely, our results show that the agent's problem can be split into two parts: First, solve for the optimal health expenditures by maximizing the present value of the agent's income net of health investments to determine the agent's human capital. Second, compute the optimal consumption, portfolio and insurance coverage to maximize the agent's utility given that his total wealth is equal to the sum of his financial wealth and human capital.

The model with exogenous health mortality and morbidity risks is very tractable and captures some of the determinants of the agent's decisions but, unfortunately, it also displays some important shortcomings when confronted to the data. In particular, it counter-factually entails that both health expenditures and insurance coverage are wealth-independent as well as increasing in health, and that health and wealth are perfect substitutes, contrary to recent evidence suggesting that the marginal utility of wealth increases with health (e.g., Finkelstein et al., 2008, 2009, among others). Motivated by theses shortcomings we then turn to an unrestricted version of the model in which the agent's health influences his decisions through both the budget constraint channel and the risk channel.

[4]See for example Grossman (1972); Hall and Jones (2007); Edwards (2008) and Yogo (2008)

[5]See Hall and Jones (2007); Chang (2005) for models with endogenous mortality but exogenous morbidity, and Picone et al. (1998); Edwards (2008) and Laporte and Ferguson (2007) among others for models with endogenous morbidity but exogenous mortality.

Allowing for health-dependent arrival rates of mortality and morbidity shocks endogenizes the agent's health risks, and implies that the model can no longer be solved in closed form. To circumvent this difficulty, we resort to a perturbation analysis that uses the explicit solution of the restricted model as the starting point of a first order expansion with respect to the parameters that govern the health dependence of the intensities associated with mortality and morbidity shocks. This approach delivers an explicit solution for the approximate optimal rules and thereby allows for a clear interpretation of the marginal impact of endogenous morbidity and mortality risks on the agent's decisions. In particular, separating financial and health-related decisions is optimal as long as mortality remains exogenous, but not otherwise. Furthermore, we show that the unrestricted model fixes the shortcomings of the model with exogenous health risks and can potentially explain the cross-sectional patterns found in the data.

To verify whether this is the case we estimate the quadrivariate system of optimal rules derived from the theoretical model to identify a set of key parameters. The estimation results, obtained using a sample of individuals drawn from PSID, attest that the model with endogenous health risks offers a good in-sample fit of the observed allocations with realistic parameter values and allow us to confirm the relevance of the main characteristics of the model. In particular, our parameter estimates clearly show that agents' preferences are non time additive and display source dependent risk aversion. To investigate the out-of-sample performance of the model we derive explicit expressions for life expectancy, as well as the values of health and life, and then use the estimated parameters to compute the prediction of the models regarding these quantities. The corresponding results are realistic and compare favorably with received estimates in the literature. In particular, we find that at the estimated parameters the model matches almost exactly the life expectancy of an average agent. Overall, both in- and out-of-sample results convey a similar message: Whereas a non-negligible part of morbidity and mortality risks is attributable to endowed factors, agents can (and do) adjust both health-related risks through health investments. However, our estimates confirm that these adjustments are constrained by powerful convexities.

The three papers that are most closely related to our work are those of Edwards (2008), Yogo (2009) and Hall and Jones (2007). Edwards (2008) studies financial decisions in the presence of health risks, but he completely abstracts from health-dependent income and endogenous health risks. Moreover, his distributional assumptions on health are quite different from ours since in his model sickness is uninsurable and requires constant expenditures once incurred. Yogo (2009) is closer to us in that he also considers the portfolio implications of a model where health investments are subject to diminishing returns to scale. However, his focus on housing and the welfare gains of actuarially fair annuities is quite different. Moreover, he models health as generating direct utility flows instead of our health-dependent labor income approach. Similar to us Hall and Jones (2007) also consider an endogenous mortality model with costly health investment and positive service flows of health. However, they do not consider portfolio allocations and their focus on the time series of aggregate health spending and longevity is very different from ours. Importantly, these papers provide neither joint analytical solutions for consumption, portfolio, health expenditures and insurance in the presence of endogenous health risks, nor a structural estimation of these allocations.

The rest of this paper is organized as follows. We introduce the theoretical model in Section 2. The solution to the model is discussed in Section 3. We present the empirical evaluation of the model in Section 4, and provide concluding remarks in Section 5. The proofs of all results are gathered in Appendix A. Appendix B outlines a general version of the model where all coefficients can depend on the agent's age, and Appendix C presents an overview of the cross-sectional PSID data that we use in our estimation.

2 The model

This section describes an economic environment in which the agent has preferences over lifetime consumption plans in the presence of partially controllable mortality and morbidity risks.

2.1 Survival and health dynamics

Let T_m denote the random duration of the agent's lifetime or, equivalently the agent's age at death, and H_t represent his health status at age t. Following Ehrlich (2000); Ehrlich and Yin (2005) and Hall and Jones (2007), we model the agent's mortality as the first jump of a Poisson process Q_m whose intensity depends on the agent's health status.

Specifically, the agent's death intensity is defined to be

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

for some nonnegative constants λ_{m0}, λ_{m1} and $\xi_m \geq 1$ where $P_t(\cdot)$ is a conditional probability and $H_{t-} = \lim_{s \uparrow t} H_s$. The fact that the intensity function is decreasing in health ensures that the survival probability:

$$P_t[T_m > t + s] = 1_{\{T_m > t\}} E_t \left[e^{-\int_t^{t+s} \lambda_m(H_{\tau-}) \mathrm{d}\tau}\right] \tag{2}$$

is monotone increasing in the agent's health status up to an exogenous ceiling that is determined by the constant $\lambda_{m0} > 0$. Intuitively, an agent may increase his survival probability by investing in his health and still die from an exogenous shock that does not depend on controllable health (e.g., an accident or certain types of cancer). Alternatively, this incompressible part of the intensity can be interpreted as an endowed death probability that is determined by environmental and/or biological factors.

The specification of the survival probability in (2) differs from those proposed in the literature along three important dimensions. First, the incompressible part of the death intensity is made constant rather than age-varying for tractability reasons (see however Remark 1 below for time-varying extensions). Second, the endogenous part of the death intensity is a function of the agent's current health status rather than of his current health investment. This assumption implies that the agent cannot freely alter his survival probability by investing large amounts in times of sickness and, thus, reflects the path dependence of health-related decisions. Third, the death intensity in (1) is a function of a stochastic rather than a deterministic health process.

To describe the dynamics of the agent's health status, let Q_s denote a Poisson process whose jumps capture shocks to the agent's health, and I be a nonnegative predictable process that represents the agent's health investment.[6] We assume that the agent's health

[6]The constraint that health investment cannot be negative is standard in the health economics literature. See for example Grossman (1972); Ehrlich and Chuma (1990); Chang (1996); Picone et al. (1998); Ehrlich (2000); Edwards (2008); Hall and Jones (2007). It reflects the irreversibility of health related expenditures and the fact that health is not a traded asset.

status evolves according to

$$\mathrm{d}H_t = \left((I_t/H_{t-})^\alpha - \delta\right) H_{t-}\mathrm{d}t - \phi H_{t-}\mathrm{d}Q_{st}, \quad H_0 > 0, \tag{3}$$

for some constants $\delta \geq 0$, and $\alpha, \phi \in (0,1)$ that represent the decay rate of health in the absence of shocks, the degree of health adjustment costs and the fraction of health that is lost upon suffering a shock. The above dynamics imply that the expected instantaneous growth rate of health

$$E_{t-}\left[\mathrm{d}H_t/H_{t-}\right] = \left((I_t/H_{t-})^\alpha - \delta - \phi\lambda_s(H_{t-})\right)\mathrm{d}t \tag{4}$$

is concave in the investment-to-health ratio. This implies that a given amount of health investment has a larger impact on the agent's health when he is currently unhealthy and thus models decreasing returns to health investment.[7]

To capture the fact that morbidity shocks are less likely for healthier agents we assume that the intensity with which these shocks occur is decreasing in the agent's health status and given by

$$\lambda_s(H_{t-}) = \eta + \frac{\lambda_{s0} - \eta}{1 + \lambda_{s1}H_{t-}^{-\xi_s}} \tag{5}$$

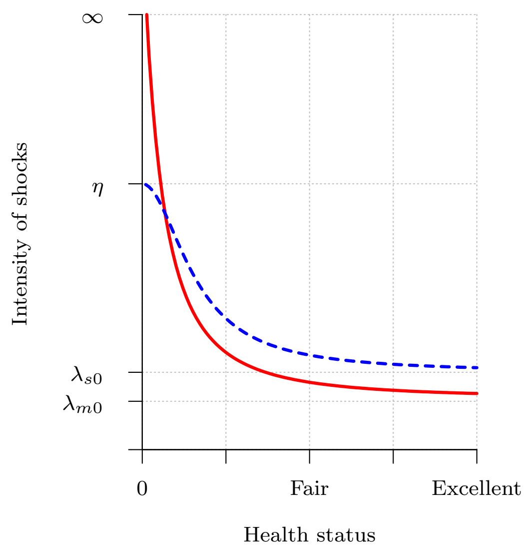
for some nonnegative constants such that $\lambda_{s0} \leq \eta$ and $\xi_s \geq 1$. Similar to (1) this functional form implies that while the agent can lower the likelihood of health shocks by investing in his health, he cannot reduce it further than

$$\lambda_{s0} = \lim_{H\to\infty} \lambda_s(H)$$

which can be interpreted as an endowed probability of health shocks. Note that the intensities of mortality and morbidity shocks imply very different risk characteristics as the agent's health deteriorates. In particular, and as illustrated by Figure 1, the agent's death intensity diverges to infinity, thus leading to certain death, as his health decreases

[7] Similar decreasing returns to health investments can be found in Ehrlich and Chuma (1990) and Ehrlich (2000); Ehrlich and Yin (2005). An equivalent interpretation of (3) is that the agent is endowed with a health production function that is linear in gross health investment $I_g = I^\alpha H^{1-\alpha}$ but faces convex adjustment costs that are given by $I = H^{1-b}I_g^b$ with $b = 1/\alpha > 1$.

Figure 1: Intensities of mortality and morbidity shocks



Notes: Instantaneous intensities of mortality shocks (solid) and morbidity shocks (dashed) as functions of the agent's health status.

to zero whereas the intensity of morbidity shocks remains bounded and reaches a finite maximal value given by $\eta = \lambda_s(0)$.

2.2 Income, traded assets and budget constraint

We assume that the agent's flow rate of labor income is given by an increasing function of his current health status:

$$Y_t = Y(H_{t-}) = y_0 + \beta H_{t-}$$

for some constants $y_0, \beta \geq 0$. A natural interpretation of this specification is that employers offer higher wages to agents who are in better health and thus less subject to be absent from work. Equivalently, a healthier agent misses less workdays and hence receives higher labor income.[8]

[8] Since the benchmark model does not allow for age-dependent parameters (see however Remark 1 below) our income specification implies that the agent's income depends on his health status even at old age. This feature of the model is consistent with the findings of French (2005) that many elders find it profitable to continue working after retirement.

The financial market is frictionless and consists in two continuously traded securities: a riskless bond and a risky stock. The price of the bond is e^{rt} for some constant rate of interest $r > 0$ and the price of the stock evolves according to

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t + \sigma_S S_t \mathrm{d}Z_t, \qquad S_0 > 0,$$

for some constant growth rate $\mu \geq r$ and constant volatility $\sigma_S > 0$ where the process Z is a standard Brownian motion.

In addition to the bond and the stock, we assume that the agent can invest in an instantaneous health insurance contract. Specifically, at every point in time the agent may purchase an actuarially fair contract that pays one unit of consumption if a health shock occurs at the next instant and zero otherwise. The net pay-off of such a contract to the agent is thus given by

$$x_t \mathrm{d}M_{st} = x_t \mathrm{d}Q_{st} - x_t \lambda_s(H_{t-})\mathrm{d}t$$

where the predictable process x_t represents the chosen amount of coverage chosen, $x_t \mathrm{d}Q_{st}$ is the amount paid by the insurer in case of a shock, and $x_t \lambda_s(H_{t-})\mathrm{d}t$ represents the instantaneous insurance premium paid by the agent. Since the agent should not be allowed sell insurance contracts on his own health, the amount of coverage x_t is constrained to be nonnegative at all times.

Assume that the agent has some initial financial wealth W_0, and let the predictable processes $c \geq 0$ and $\pi \in \mathbb{R}$ represent the amount he consumes and the amount he invests in the stock. Under the usual self-financing requirement, the agent's financial wealth then evolves according to

$$dW_t = (rW_{t-} + \beta H_{t-} - c_t - I_t)\,\mathrm{d}t + \pi_t \sigma_S\,(\mathrm{d}Z_t + \theta \mathrm{d}t) + x_t \mathrm{d}M_{st} \tag{6}$$

where the constant $\theta = (\mu - r)/\sigma_S \geq 0$ is the market price of financial risk. This budget constraint reveals two additional channels through which the agent's health status influences his decisions: An unhealthy individual faces not only a lower labor income but also a higher health insurance premium because of the higher probability of health shock occurrence. Both of these channels concur to reduce resources available for consumption, health expenditures and financial investments.

Remark 1 (Age dependent parameters) The model presented above assumes that all the agent-specific parameters are constant. This assumption is imposed in order to facilitate the exposition and interpretation of our results but can be relaxed at the cost of more involved notation. We present in Appendix B a general version of the model in which the intensity of shocks λ_{m0}, λ_{m1}, λ_{s1}, λ_{s_0}, η, the depreciation rate of health δ, the fraction of health ϕ that is lost upon experiencing a health shock, and the health sensitivity β of labor income are allowed to vary with the agent's age.

2.3 Preferences

Starting with the seminal contributions of Yaari (1965) and Hakansson (1969), the standard way of specifying preferences in the presence of mortality risk has been to define the utility to an agent of age t of a consumption plan c as

$$U_t = 1_{\{T_m > t\}} E_t \int_t^{T_m} e^{-\rho(s-t)} u(c_s) \mathrm{d}s \tag{7}$$

for some nonnegative subjective rate of time preference ρ and some concave period utility function u satisfying the usual regularity conditions.[9]

As pointed out by Shepard and Zeckhauser (1984) and Rosen (1988), the level of the period utility has important implications in such a specification since adding a constant to u changes the value that the agent places on longevity relative to consumption. Put differently, in the presence of an uncertain and endogenous horizon, preferences are not invariant to affine transformations as they are in the standard setting where the horizon is non random and exogenous. This undesirable feature is due to the fact that (7) attributes utility zero to death and, hence, implies that the utility of any consumption schedule must be compared to zero to determine whether the agent is better off living or dying. In particular, if the period utility is of the iso-elastic type:

$$u(c; \vartheta) = c^{1-\vartheta}/(1-\vartheta); \quad c \geq 0 \tag{8}$$

[9]See for example Richard (1975); Shepard and Zeckhauser (1984); Rosen (1988); Ehrlich and Chuma (1990); Ehrlich (2000); Becker et al. (2005); Edwards (2008); Hall and Jones (2007) and Yogo (2009).

for some nonnegative constant $\vartheta \neq 1$ then the agent's preferences towards mortality depend on whether the risk aversion parameter ϑ is smaller or larger than unity. In the former, the utility of any consumption schedule is positive and it follows that the agent prefers life to death. On the contrary, if $\vartheta > 1$, as is often found in empirical studies, then the utility of any consumption schedule is negative and the agent thus counterintuitively prefers death to life irrespective of his current consumption level.[10]

In addition to this non-invariance, the time additive specification in (7) suffers from two other important limitations. First, by summing up the utility of period consumption up to the time of death, the time additive specification counter-intuitively assumes that the agent is risk neutral towards mortality risk (see Bommier (2006)). Second, this specification supposes that the agent's risk preferences are entirely summarized by the period utility function u and thus does not allow for different attitudes towards different sources of risk. This last restriction is particularly important in the context of our model because there is no *ex-ante* reason to believe that agents should be equally averse to mortality, morbidity, and financial risks.

Motivated by the above discussion, and in particular by the fact that (7) cannot reconcile an empirically plausible level of risk aversion with a sensible behavior towards longevity risk, we will forgo the time additive specification and assume instead that the agent has recursive preferences of the type proposed by Kreps and Porteus (1979); Epstein and Zin (1989); Weil (1989) and Duffie and Epstein (1992b). As we show below, an appropriate generalization of these preferences allows to remedy the above deficiencies of the time additive specification while maintaining a tractable setup.

Let $U_t = U_t(c)$ be the continuation utility to an agent of age t of a consumption schedule c, denote the instantaneous volatility of this process by

$$\sigma_t = \frac{1}{\mathrm{d}t} d\langle U, Z\rangle_t$$

[10] A similar problem arises for the negative exponential utility given by $u(c) = -\exp(-ac)$ for some $a > 0$. To ensure sensible results, many authors consider nonnegative period utility functions for which life is always preferred. Following this approach, Rosen (1988); Becker et al. (2005) and Hall and Jones (2007) use a utility of the form $v(c) = u(c) + b$ where b is chosen in such a way as to guarantee that v is nonnegative. Unfortunately, such a constant exists only if u is bounded from below and it follows that this approach cannot be used to accommodate the case where u is given by (8) for some $\vartheta > 1$.

and let

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

represent the predictable jump in the agent's continuation utility that is triggered by a jump in either the mortality process ($k = m$), or the health shock process ($k = s$). Generalizing the continuous-time recursive preference specification of Duffie and Epstein (1992b) we assume that the continuation utility process, its volatility and its jumps satisfy the recursive integral equation

$$U_t = 1_{\{T_m > t\}} E_t \int_t^{T_m} \left(f(c_\tau, U_{\tau-}) - \frac{\gamma|\sigma_\tau|^2}{2U_{\tau-}} - \sum_{k=m}^{s} F_k(U_{\tau-}, H_{\tau-}, \Delta_k U_\tau) \right) \mathrm{d}\tau \tag{9}$$

where the constant $\gamma > 0$ measures the agent's local risk aversion over static financial gambles; the function

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

is the standard Kreps-Porteus aggregator with elasticity of intertemporal substitution $\varepsilon > 0$, subjective rate of time preference $\rho > 0$ and subsistence consumption level $a \geq 0$; and we have set

$$F_k(v, h, \Delta) = v\lambda_k(h) \left[\frac{\Delta}{v} + u(1; \gamma_k) - u\left(1 + \frac{\Delta}{v}; \gamma_k\right) \right],$$

where $u(x; \gamma_k)$ is the constant relative risk aversion utility function of equation (8) with curvature indices $0 \leq \gamma_m < 1$ and $\gamma_s \geq 0$.

The first two terms inside the integral on the right hand side of (9) correspond to standard Kreps-Porteus preferences in a Brownian setting and encode, respectively, the agent's substitution behavior and his risk aversion towards the Brownian motion driving financial market returns. By contrast, the last two terms are associated with mortality ($k = m$) and morbidity shocks ($k = s$) and penalize the agent's utility for exposure of these sources of risks due to the fact that F_k is nonnegative. In particular, since (9) does not include bequests[11] the continuation utility vanishes at death, and it follows that the

[11] This assumption is imposed for tractability and can be justified by noting that while bequest motives are potentially relevant in an endogenous mortality setting, panel data evidence suggests that their role in

penalization for mortality risk satisfies

$$\Phi_m = \frac{F_m(U_{\tau-}, H_{\tau-}, \Delta_m U_\tau)}{\lambda_m(H_{\tau-})U_{\tau-}} = u(1;\gamma_m) - \lim_{x\to 0} u(x;\gamma_m) - 1.$$

This expression reveals why the risk aversion parameter γ_m associated with mortality risk must be strictly smaller than unity. Indeed, the penalization associated with death would otherwise be infinite due to the fact that $\lim_{x\to 0} u(x;\gamma_m) = -\infty$ for $\gamma_m \geq 1$, and the agent's utility would therefore be undefined.

An important feature of our preference specification is that, since the risk aversion parameters γ, γ_m and γ_s can be different, it not only disentangles the agent's attitude toward intertemporal substitution from his attitude towards risk but also allows to discriminate among various sources of risk.[12] This feature is referred to as *source dependent* risk aversion (see e.g., Lazrak and Quenez, 2003; Skiadas, 2007, 2009) and our model constitutes one of the first applications of such preferences to the study of portfolio, consumption, and health-related choices.

A second essential property of our specification is that it guarantees unconditional preference for life. Indeed, following Duffie and Epstein (1992b) it can be shown that the homogeneity of the aggregator f and the penalty functions (F_m, F_s) implies that continuation utility is homogenous of degree one so that utility and excess consumption are is measured in the same units. In particular, the utility associated with a nonnegative consumption schedule is nonnegative and, since death is by definition associated with zero consumption in the absence of bequests, it follows that the agent sees his own mortality as detrimental irrespective of whether his risk aversion towards financial risks (γ) and morbidity risk (γ_s) are smaller or larger than unity.

2.4 The decision problem

The agent's decision problem consists in choosing a portfolio, consumption, health insurance and health investment strategy to maximize his lifetime utility. Accordingly, the

explaining the behavior of retired agents is debatable. In particular, Hurd (2002) finds no clear evidence of a bequest motive behind savings decisions and Hurd (1987) finds no differences in the saving behavior of the elderly who have children compared to those who don't.

[12] Our specification is equivalent to the continuous-time Kreps-Porteus specification of Duffie and Epstein (1992b) when $\gamma_s = \gamma_m = \gamma$ and to time-additive iso-elastic utility when $\gamma_s = \gamma_m = \gamma = 1/\varepsilon$.

agent's indirect utility is defined by

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

subject to the death intensity (1), the health process (3), the health shock intensity (5), and the budget constraint (6), where $U_t(c)$ is the continuation utility associated with the lifetime consumption plan c through the recursive integral equation (9).

Since the uncertain duration of his lifetime cannot be hedged by trading in the available assets, the agent faces incomplete markets. However, under the assumption of Poisson mortality, his decision problem can be conveniently recast as an equivalent infinite horizon problem with endogenous discounting and complete markets. Indeed, using (2) and the law of iterated expectations, the continuation utility writes

$$U_t(c) = 1_{\{T_m>t\}} V_t(c)$$

where the modified utility process $V_t = V_t(c)$ solves

$$V_t = E_t \int_t^\infty e^{-\int_t^\tau \nu_m(H_{v-})\mathrm{d}v} \left(f(c_\tau, V_{\tau-}) - \frac{\gamma|\sigma_\tau|^2}{2V_{\tau-}} - F_s(V_{\tau-}, H_{\tau-}, \Delta_s V_\tau) \right) \mathrm{d}\tau. \tag{10}$$

with

$$\nu_m(H) = (1+\Phi_m)\lambda_m(H) = \frac{\lambda_m(H)}{1-\gamma_m}. \tag{11}$$

This formulation of the agent's objective brings to light the channels through which health enters the decision problem. First, health can be interpreted as a durable good that generates service flows through the income $Y(H)$ net of insurance premium $x\lambda_s(H)$. Second, health determines the instantaneous probability of morbidity shocks and the rate $\nu_m(H)$ at which the agent discounts future consumption and continuation utilities. We show in the next sections how these two channels, that we refer to as the budget constraint and the risk channel, interact to generate the optimal rules.

Remark 2 (Health-dependent preferences) Our formulation of the agent's problem closely parallels the widely used approach of specifying a health dependent utility and omitting health-dependent income.[13] To see this, let $\bar{c} = c - \beta H$ denote the agent's consumption in excess of his income, and rewrite the problem as

$$V(W_t, H_t) = \sup_{(\bar{c},\pi,x,I)} U_t(\bar{c} + \beta H),$$

subject to (1), (3), (5), and the modified budget constraint

$$dW_t = (rW_{t-} - \bar{c}_t - I_t)\,\mathrm{d}t + \pi_t\sigma\,(\mathrm{d}Z_t + \theta\mathrm{d}t) + x_t\mathrm{d}M_{st}.$$

Hence, abstracting from health-dependent income and solving the agent's problem with the non separable, health-dependent intertemporal aggregator

$$\bar{f}(c, H, v) = f(c + \beta H, v),$$

is equivalent to solving our formulation of the agent's problem with health-independent intertemporal aggregator and health-dependent income.

3 Optimal rules

This section derives the solution to our model. As explained above, the agent's health enters the problem through two channels: the risk channel and the budget constraint channel. In order to gain intuition on the respective impact of these pathways, Section 3.1 starts by abstracting from the first channel to consider the budget constraint effects only. Section 3.2 then turns to the more general case where the agent's health influences both his mortality and morbidity intensities in addition to his budget constraint.

3.1 Health independent mortality and morbidity

When $\lambda_{m1} = \lambda_{s1} = 0$ the mortality and morbidity intensities are constant and, as a result, the agent's objective function (10) is independent from his health. In conjunction with market completeness, this implies that the problem is separable and can be solved in two steps as in Bodie et al. (1992). First, the optimal health investment is computed

[13] Examples of studies that follow this approach include Grossman (1972), Ehrlich and Chuma (1990), Picone et al. (1998), Ehrlich (2000), Edwards (2008), Hall and Jones (2007) and Yogo (2009).

by maximizing the agent's human wealth defined as the present value of his income net of health investments. Second, the optimal portfolio, optimal health insurance and optimal consumption schedule are obtained by solving the problem of an hypothetical agent who has no income, but whose initial wealth is replaced by the total (i.e., financial plus human) wealth of the original agent.

Since markets are complete, the present value of the health dependent part of the agent's income net of health investments can be computed as

$$P(H_t) = \sup_{I \geq 0} E_t \int_t^\infty m_{t,\tau} \left(\beta H_{\tau-} - I_\tau \right) \mathrm{d}\tau$$

subject to the law of motion for health (3), where the nonnegative process

$$m_t = \exp\left(-rt - \theta Z_t - \frac{\theta^2}{2} t \right) \tag{12}$$

is the stochastic discount factor induced by the prices of the bond, the stock and the insurance contract, and we have set $m_{t,\tau} = m_\tau / m_t$. The following proposition derives an analytical solution to this first-step problem.

Proposition 1 *Let $\lambda_{m1} = \lambda_{s1} = 0$, assume that*

$$\beta < (r + \delta + \phi \lambda_{s0})^{\frac{1}{\alpha}} \tag{13}$$

and define

$$g(x) = \beta - (r + \delta + \phi \lambda_{s0}) x - (1 - 1/\alpha) (\alpha x)^{\frac{1}{1-\alpha}}.$$

Then the present value of the agent's income and the optimal health investment strategy are explicitly given by

$$P_0(H) = BH, \tag{14}$$

$$I_{0t} = (\alpha P_{0H}(H_{t-}))^{\frac{1}{1-\alpha}} H_{t-} = (\alpha B)^{\frac{1}{1-\alpha}} H_{t-} = K P_0(H_{t-}), \tag{15}$$

where B is the unique nonnegative constant such that $g(B) = 0$ and $g'(B) < 0$.

The restriction imposed by equation (13) is a transversality condition that limits the health sensitivity of the agent's income rate in order to guarantee that the present value of his income is finite. The fact that this present value is linear in the agent's health status implies that the constant

$$B = P_0(1) = P_{0H}(1)$$

gives both the average and the marginal value of health and follows from the linearity of income, the restriction to constant intensities, and the Cobb-Douglas specification of the health adjustment technology.[14]

Having computed the present value of the agent's income and the optimal health investment strategy, we now turn to the determination of the optimal consumption, portfolio and insurance strategy. Let

$$N_t = N_0(W_t, H_t) = W_t + P_0(H_t) + \frac{y_0 - a}{r} \tag{16}$$

denote the agent's total (i.e., financial plus human) wealth net of minimal consumption expenditures. Using the result of Proposition 1 together with the budget constraint (6) and the definition of the constant B, it can be shown that the agent's total wealth evolves according to

$$\mathrm{d}N_t = (rN_{t-} - c_t)\mathrm{d}t + \pi_t \sigma_S(\mathrm{d}Z_t + \theta \mathrm{d}t) + \bar{x}_t \mathrm{d}M_{st} \tag{17}$$

where $\bar{x}_t = x_t - \phi P_0(H_{t-})$ and $\bar{c}_t = c_t - a$ represent the agent's net exposure to health shocks and his excess consumption. This implies that under exogenous mortality and morbidity the indirect utility of an alive agent is given by

$$V_0(W_t, H_t) = G(N_0(W_t, H_t)) = \sup_{(\bar{c},\pi,\bar{x})} V_t(c) \tag{18}$$

subject to the budget constraint for total wealth in equation (17). The solution to this portfolio, insurance and consumption choice problem with recursive utility and source dependent preferences can be obtained as a generalization of the results in Svensson

[14] See Uzawa (1969), Hayashi (1982) and Abel and Eberly (1994) for similar results in the investment literature where this property is referred to as the equivalence between marginal and average q.

(1989); Obstfeld (1994) and Smith (1996) among others. Using this solution to construct the agent's optimal rules delivers the following theorem.

Theorem 1 *Let* $\lambda_{m1} = \lambda_{s1} = 0$*, assume that the transversality condition* (13) *as well as*

$$A = \varepsilon\rho + (1-\varepsilon)(r - \nu_{m0} + \theta^2/(2\gamma)) > \max\left(0; r - \nu_{m0} + \theta^2/\gamma\right) \tag{19}$$

hold true with $\nu_{m0} = \lambda_{m0}/(1-\gamma_m)$ *and define* $\Theta = \rho(A/\rho)^{1/(1-\varepsilon)} > 0$*. Then the indirect utility function of an alive agent is*

$$V_0(W, H) = \Theta N_0(W, H), \tag{20}$$

and generates the optimal consumption, portfolio, health insurance and health investment strategies given by

$$c_{0t} = a + A\, N_0(W_{t-}, H_{t-}), \tag{21}$$

$$\pi_{0t} = (\theta/(\gamma\sigma_S))N_0(W_{t-}, H_{t-}), \tag{22}$$

$$x_{0t} = \phi P_0(H_{t-}), \tag{23}$$

and equation (15) *where the agent's human wealth* $P_0(H)$ *and total wealth* $N_0(W, H)$ *are defined in equations* (14) *and* (16).

As explained by Smith (1996) in a simpler context, the restriction imposed by (19) serves two purposes. On the one hand, it guarantees that the marginal propensity to consume A is strictly positive and, hence, that the optimal consumption plan is feasible. On the other hand, it insures that the indirect utility coincides with the continuation utility of the optimal consumption schedule as defined in (9) and, thus, constitutes a transversality condition. The parametric form of the restriction is entirely standard (e.g., Svensson, 1989; Obstfeld, 1994; Smith, 1996), except for the presence of the constant $\lambda_{m0}/(1-\gamma_m) > 0$ that reflects the combined impact of mortality risk and the agent's aversion to that risk on the optimal consumption schedule.

Proposition 1 and Theorem 1 show that, due to the separation between health investment and the agent's other decisions, exogenous morbidity and mortality have very different effects on the optimal rules. Indeed, the morbidity parameters (ϕ, λ_{s0}) govern the marginal value of health and thereby determine the agent's total wealth so that their impact on the optimal rules must be analyzed through their effect on available resources. By contrast, the mortality parameter λ_{m0} does not affect the agent's total wealth but determines the sensitivity of the optimal rules to changes in the available resources.

To understand the effect of exogenous morbidity risk, consider the expected instantaneous growth rate of health (4). As shown by this equation, an increase in either the health shock intensity λ_{s0} or the fraction of health ϕ lost from such a shock is equivalent to an increase in the rate $\delta + \phi\lambda_{s0}$ at which the agent's health status is expected to depreciate in the absence of investment. Faster expected depreciation reduces the agent's human wealth by lowering the marginal value of health,[15] and thus implies both lower health investments and a lower amount of health insurance. In addition, a lower human wealth leads to a lower total wealth and, thereby, triggers a decrease in the agent's welfare, consumption and risky portfolio since all are proportional to total wealth.

The human wealth (14) and health insurance (23) reveal that with exogenous mortality and morbidity, it is always optimal for the agent to fully hedge health shocks. Indeed, the dynamics of the health status and the expression for the optimal insurance coverage imply that the net exposure to health shocks is:

$$\begin{aligned}\Delta_s N_{0t} &= \Delta_s N_0(W_t, H_t) \\ &= 1_{\{\mathrm{d}Q_{st}\neq 0\}}(N_0(W_{t-} + x_{0t}, H_{t-}(1-\phi)) - N_0(W_{t-}, H_{t-})) \\ &= 1_{\{\mathrm{d}Q_{st}\neq 0\}}(x_{0t} + \Delta_s P_0(H_{t-})) = 0,\end{aligned}$$

so that the agent's total wealth is insensitive to health shocks at the optimum. To understand this result note that with exogenous mortality and morbidity the agent's only exposure to health shock risk comes from his income and observe that this risk does carry a risk premium as the insurance contract is assumed to be actuarially fair. Since the agent is risk averse he will not willingly expose himself to a risk for which he is not remunerated, and it follows that he will choose his health insurance coverage in such a way as to eliminate any exposure to that risk.

Turning to the impact of exogenous mortality risk, (19) and (21) reveal that an increase in either the mortality risk parameter λ_{m0} or the mortality risk aversion γ_m is

[15]See the appendix for a proof that the unique nonnegative constant B satisfying $g(B) = 0$ and $g'(B) < 0$ decreases as the expected depreciation rate $\delta + \phi\lambda_{s0}$ increases.

equivalent to a decrease in the interest rate and thereby leads to two conflicting effects. First, it implies that more resources are needed to fund a given level of future consumption and thus encourages the agent to consume less today in order to maintain the same level of future consumption. Second, it makes current consumption less costly relative to future consumption and therefore leads to consume more today through a substitution effect. When the agent's elasticity of intertemporal substitution (EIS) ε is smaller than unity, the first effect dominates and the agent reduces his consumption in response to an increase in either mortality risk or his aversion to that risk. Conversely, when $\varepsilon > 1$ the substitution effect dominates and the agent increases his consumption. Exact cancelation of the two effects occurs when $\varepsilon = 1$ in which case mortality risk has no impact on the optimal rules. By contrast, equation (20) and the definition of Θ imply that an increase in either mortality risk or the agent's aversion to that risk decreases the agent's indirect utility irrespective of his EIS. This result reflects the unconditional preference for life implied by our preference specification and stands in stark contrast to the corresponding result for time additive iso-elastic preferences where the impact of mortality on welfare depends on whether risk aversion is greater or smaller than unity.

The expression for the optimal risky portfolio in (22) shows that the fraction of total wealth invested in the risky stock depends neither on mortality risk nor on the agent's aversion to that risk, and only involves the market Sharpe ratio and the agent's aversion to financial risk. To understand this result observe that with exogenous mortality and morbidity the agent's investment opportunity set is constant, and recall from Richard (1975) that in such a setting the optimal investment in risky assets is independent of the distribution of the agent's exogenous planning horizon. Consequently, the optimal fraction of total wealth invested in the stock is given by the myopic demand $\theta/(\gamma\sigma_S)$ and decreases with the agent's financial risk aversion but remains unaffected by changes in either mortality risk or the agent's aversion to that risk.

The optimal rules associated with exogenous mortality and morbidity capture some of the determinants of the agent's decisions but also display some significant shortcomings when confronted to the data. In particular, recent evidence surveyed in Finkelstein et al. (2008, 2009) indicates that the marginal utility of wealth is positively affected by health, i.e. $V_{WH} > 0$, but this property cannot obtain in the restricted version of model because, under exogenous mortality and morbidity, the marginal rate of substitution

$$\left.\frac{V_H(W,H)}{V_W(W,H)}\right|_{\lambda_{m1}=\lambda_{s1}=0} = \frac{\Theta P'(H)}{\Theta} = B$$

is constant so that health and wealth are perfect substitutes. Similarly, there is ample evidence to the facts that health investment and insurance are both increasing in wealth and non-increasing in health (e.g Smith, 1999; Wu, 2003; Barros et al., 2008). But these properties cannot be obtained within the restricted model as it predicts that the optimal health investment and health insurance

$$I_{0t} = (\alpha B)^{\frac{1}{1-\alpha}} H_{t-},$$
$$x_{0t} = \phi P_0(H_{t-}) = \phi B H_{t-}$$

are independent of the agent's wealth and increase with his health. To verify whether these stylized facts can be compatible with a richer model we now relax the assumption of exogenous shocks by considering the general case in which the intensity of mortality and morbidity shocks is allowed to depend on the agent's health status.

3.2 Health dependent mortality and morbidity

When λ_{m1} and λ_{s1} are non zero the arrival rates of shocks In this case, one can no longer determine the optimal heath investment independently of the optimal portfolio, consumption and insurance strategies since the objective function in (10) now depends on the agent's health status through both the endogenous discount rate ν_m, and the health shock penalty function F_s.

Resorting instead to the Hamilton-Jacobi-Bellman (HJB) and assuming sufficient smoothness, the agent's indirect utility solves

$$0 = \max_{(c,\pi,x,I)} \mathcal{D}^{(c,\pi,x,I)} V(W,H) + f(c, V(W,H)) - \frac{\gamma(\pi\sigma_S V_W(W,H))^2}{2V(W,H)} - \lambda_s(H)V(W,H)\left(u(1;\gamma_s) - u\left(\kappa(x,W,H);\gamma_s\right)\right) - \nu_m(H)V(W,H) \quad (24)$$

where the differential operator

$$\mathcal{D}^{(c,\pi,x,I)} = ((\pi\sigma_S)^2/2)\partial_{WW} + H((I/H)^\alpha - \delta)\partial_H \\ + (rW + \pi\sigma_S\theta - c + y_0 + \beta H - I - x\lambda_s(H))\partial_W$$

is the continuous part of the infinitesimal generator of the state variables under the strategy (π, c, x, I), and

$$\kappa(x, W, H) = \frac{V(W + x, H(1-\phi))}{V(W, H)}$$

represents the relative jump in the agent's indirect utility induced by the occurrence of a health shock. Maximizing the right hand side of the HJB equation reveals that, given the indirect utility function, the optimal consumption, portfolio and health investment can be computed as

$$c^* = a + V(W,H)\left(\frac{\rho}{V_W(W,H)}\right)^{\varepsilon}, \tag{25}$$

$$\pi^* = \frac{(\theta/\sigma_S)V(W,H)V_W(W,H)}{\gamma V_W(W,H)^2 - V(W,H)V_{WW}(W,H)}, \tag{26}$$

$$I^* = H\left(\frac{\alpha V_H(W,H)}{V_W(W,H)}\right)^{\frac{1}{1-\alpha}}, \tag{27}$$

whereas the optimal health insurance is implicitly defined by

$$\frac{V_W(W,H)}{V_W(W + x^*, H(1-\phi))} = \kappa(x^*, W, H)^{-\gamma_s}. \tag{28}$$

Substituting these first order conditions into the HJB equation and simplifying the result produces a nonlinear partial differential equation for the indirect utility. Unfortunately, no closed form solution to this equation can be obtained except for the case of exogenous mortality and morbidity considered in Section 3.1. Nonetheless, and as we now explain, one can use the solution to this special case together with an asymptotic expansion to construct an approximate solution to the general case.

Assume that $\lambda_{k1} = \epsilon\bar{\lambda}_{k1}$ for some fixed constants $\bar{\lambda}_{m1}, \bar{\lambda}_{s1} \geq 0$ and let us expand the indirect utility of an alive agent around $\epsilon = 0$ as

$$V(W, H) \approx V_n(W, H) = V_0(W, H) + \sum_{k=1}^{n} \frac{\epsilon^k}{k!} V^{(k)}(W, H)$$

where n is an integer that represents the order of the expansion, V_0 is the indirect utility for the case of exogenous mortality and morbidity and the derivative

$$V^{(k)}(W, H) = \left. \frac{\partial^k V(W, H)}{\partial \epsilon^k} \right|_{\epsilon=0},$$

represents the k-th order correction to the indirect utility induced by the presence of health-dependent mortality and morbidity. Substituting this approximation into the HJB equation and expanding the result in powers of ϵ gives a sequence of partial differential equations that can be solved recursively starting from the known function V_0. Once the correction terms have been computed up to the desired order, one can obtain an approximation of the optimal portfolio, consumption, health investment, and insurance coverage by substituting the above expansion into the first-order conditions (26), (25) (27) and (28) and again expanding the result in powers of ϵ.

In order to implement this solution method it is necessary to select the accuracy of the approximation by fixing the number of terms n to include in the expansion. Since the intensity parameters $\lambda_{m1}, \lambda_{s1}$ are expected to be small,[16] we can be reasonably confident that the expansion method already delivers good approximations of the indirect utility and optimal rules at the first order ($n = 1$). While higher order approximations can also be computed, we will restrict ourselves to this first order solution because it allows for an intuitive analysis of the optimal rules.

Theorem 2 *Let*

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

[16] The estimated value of the parameters λ_{m1} and λ_{s1} obtained through a structural estimation of the optimal rules predicted by the model are of the order of 10^{-3}. See Section 4.1 for details.

assume that the transversality conditions (13) *and* (19)*, as well as*

$$\min(\nu_{m0}, r) - F(1-\xi_s) > 0, \tag{29}$$

$$A - \max\left(0, r - \nu_{m0} + \theta^2/\gamma\right) - F(-\xi_m) > 0, \tag{30}$$

hold true, and define a pair of negative functions by setting

$$L_s(H) = \lambda_{s1}\phi(\eta - \lambda_{s0})\left(F(1-\xi_s) - r\right)^{-1} H^{-\xi_s}$$
$$L_m(H) = \lambda_{m1}\left((1-\gamma_m)(F(-\xi_m) - A)\right)^{-1} H^{-\xi_m}$$

where the constants B*,* A *and* Θ *are defined as in Proposition 1 and Theorem 1. Up to a first order approximation the indirect utility of an alive agent is*

$$V_1(W, H) = V_0(W, H) + \Theta L_m(H) N_0(W, H) + \Theta L_s(H) P_0(H) \tag{31}$$

and generates the approximate optimal consumption, portfolio, health insurance and health investment strategy given by

$$c_{1t} = c_{0t} + A(1-\varepsilon)L_m(H_{t-})N_0(W_{t-}, H_{t-}) + AL_s(H_{t-})P_0(H_{t-}),$$

$$\pi_{1t} = \pi_{0t} + (\theta/(\gamma\sigma_S))L_s(H_{t-})P_0(H_{t-}), \tag{32}$$

$$x_{1t} = x_{0t} + \chi(\xi_m)(1 - 1/\gamma_s)L_m(H_{t-})N_0(W_{t-}, H_{t-}) + \chi(\xi_s - 1)L_s(H_{t-})P_0(H_{t-}) \tag{33}$$

$$I_{1t} = I_{0t} - (\xi_m K/(1-\alpha))L_m(H_{t-})N_0(W_{t-}, H_{t-}) - ((\xi_s - 1)K/(1-\alpha))L_s(H_{t-})P_0(H_{t-}) \tag{34}$$

where the constant K *is defined as in* (15)*.*

The restrictions imposed by equations (29) and (30) are both transversality conditions associated with the agent's indirect utility and total wealth. The negative functions L_k are first-order corrections to the optimal rules induced by endogenous health-related risks and can be shown to be increasing in the agent's health status.

Given complete markets, it is also possible to derive an approximation for the agent's total wealth at the optimum under endogenous morbidity and morbidity as the present value of the optimal consumption plan:

Proposition 2 *Assume that the conditions of Theorem 2 hold and let*

$$N_t^* = E_t \int_t^\infty m_{t,\tau} c_\tau^* \mathrm{d}\tau = W_t^* + E_t \int_t^\infty m_{t,\tau}(\beta H_{\tau-}^* - I_\tau^*)\mathrm{d}\tau$$

denote the agent's total wealth at the optimum. Up to a first order approximation

$$N_t^* \approx N_{1t} = N_0(W_t, H_t) + L_s(H_t)P_0(H_t)$$

where the functions $P_0(H)$ and $L_s(H)$ are defined as in Proposition 1 and Theorem 2.

The marginal effect of endogenous morbidity on the agent's welfare and decisions can be isolated by imposing the restriction $\lambda_{m1} = L_{m1} \equiv 0$ in Theorem 2 and is entirely summarized by the induced change in the agent's total wealth highlighted in Proposition 2. Indeed, a close inspection of Theorem 2 shows that under this restriction the approximate indirect utility and optimal rules are the same as those of Proposition 1 and Theorem 1 except that the zero order total wealth $N_0(W, H)$ and human wealth $P_0(H)$ are replaced by their first order counterparts:[17]

$$N_1(W_t, H_t) = N_0(W_t, H_t) + L_s(H_t)P_0(H_t)$$
$$P_1(H_t) = N_1(W_t, H_t) - W_t = (1 + L_s(H_t))P_0(H_t).$$

This shows that the separation between financial and health related decisions which was obtained in the model with exogenous health risks carries over to the model with endogenous morbidity shocks provided that mortality risk remains exogenous. It follows that to understand the marginal impact of endogenous morbidity it suffices to study to study its impact on the agent's first order total wealth $N_1(W, H)$.

[17] This is immediate for the optimal consumption, optimal portfolio, and optimal insurance coverage. For the optimal health investment the result follows by noting that under exogenous mortality the approximate optimal health investment I_1 in (34) is given by a first order approximation of the function $(\alpha P_{1H})^{1/(1-\alpha)}H$ with respect to the morbidity parameter λ_{s1} in a neighborhood of zero.

Recall from Figure 1 that allowing for $\lambda_{s1} > 0$ pushes the intensity of morbidity shocks above its minimal level λ_{s0}. As can be seen from the fact that

$$N_1(W_t, H_t) - N_0(W_t, H_t) = P_1(H_t) - P_0(H_t) = L_s(H_t)P_0(H_t) \leq 0,$$

this higher likelihood of morbidity shocks lowers the agent's total wealth and thereby induces a proportional reduction in the indirect utility, the optimal consumption schedule and the optimal portfolio. In addition, allowing for $\lambda_{s1} > 0$ implies that the agent can now partially control the likelihood of morbidity shocks by investing in his health. This increases the benefits of health expenditures and prompts the agent to increase his investment relative to the restricted model as can be seen from the fact that

$$I_{1t} - I_{0t} = \frac{K}{\alpha - 1}\left(\xi_m L_m(H_{t-})N_{0t-} + (\xi_s - 1)L_s(H_{t-})P_0(H_{t-})\right) \geq 0.$$

Comparing (23) to (33) shows that, even though it reduces the agent's total wealth, endogenous morbidity triggers an increase in the optimal amount of insurance coverage. To understand this feature observe that the occurrence of a morbidity shock not only lowers current health, but also increases the likelihood of even lower health levels in the future. As a result, the amount

$$\begin{aligned} -\Delta_s P_1(H_t) &= 1_{\{\mathrm{d}Q_{st} \neq 0\}}\left(P_1(H_{t-}) - P_1(H_{t-}(1-\phi))\right) \\ &= -\Delta_s P_0(H_t) + 1_{\{\mathrm{d}Q_{st} \neq 0\}}\chi(\xi_s - 1)L_s(H_{t-})P_0(H_{t-}), \end{aligned}$$

that the agent stands to loose from a morbidity shock is larger than its zero order counterpart $-\Delta_s P_0(H_t)$ and this induces him to increase his demand for insurance relative to the restricted version of the model with exogenous health shocks.

Imposing the restriction $\lambda_{s1} = L_{s1} \equiv 0$ in Theorem 2 and Proposition 2 allows to single out the marginal impact of endogenous mortality and reveals that up to a first order approximation the agent's total wealth is independent of mortality risk, be it endogenous or not. Unlike endogenous morbidity, the effects of endogenous mortality therefore cannot be traced back to changes in total wealth, but rather in how the optimal rules respond to available resources. In that respect, endogenous mortality invalidates the separation between financial and health-related decisions that was obtained in both the restricted model, and the model with exogenous mortality but endogenous morbidity.

As shown by the solid curve in Figure 1 allowing for $\lambda_{m1} > 0$ raises the agent's death intensity above its minimal level λ_{m0}. Because our preference specification entails a strict preference for life, this higher death probability is unambiguously welfare reducing i.e. $V_1 < V_0$ in (31). For the same reason, and because longevity risk can now be partially adjusted, the agent will try to mitigate a higher death intensity by increasing his health investment. As in the restricted version of the model, the agent may choose to offset a lower quantity of life by a higher quality of life, i.e. by increasing current consumption. Whether or not this occurs depends entirely on the EIS: the agent increases consumption relative to the restricted model with exogenous health risks if $\varepsilon > 1$, and decreases it otherwise. Exact cancellation of the income and substitution effects occurs when $\varepsilon = 1$ in which case the optimal consumption is independent of mortality risk.

Comparing (22) and (32) shows that mortality risk, be it endogenous or not, has no first order impact on the optimal portfolio. As in the restricted model this can be explained by the fact that optimal portfolios are independent of discounting in the absence of hedging motives. Indeed, the marginal impact of endogenous mortality is computed by performing an expansion around the case of exogenous mortality and morbidity. Since the optimal health status associated to that case does not covary with the Brownian motion driving the stock returns it follows that the dynamic hedging demand is zero and hence that the optimal stock holdings remain unaffected by mortality risk. On the contrary, the fact that health is subject to morbidity shocks and influences the agent's discount rate gives rise to a dynamic hedging component in the optimal demand for insurance. This dynamic hedging component is given by

$$x_{1t} - x_{0t} = \chi(\xi_m)(1 - 1/\gamma_s)L_m(H_{t-})N_{0t-}$$

and implies that, in the presence of endogenous mortality, the agent will not select his insurance coverage so as to make his total wealth insensitive to morbidity shocks. Indeed, upon the occurrence of a morbidity shock the agent's total wealth experiences a jump that is given by

$$\Delta_s(W_t + P_0(H_t)) = 1_{\{dQ_{st}\neq 0\}}(x_{1t} - \phi P_0(H_{t-}))$$
$$= 1_{\{dQ_{st}\neq 0\}}(x_{1t} - x_{0t}) = \chi(\xi_m)(1 - 1/\gamma_s)L_m(H_{t-})N_{0t-}.$$

This jump is negative, indicating that the agent remains optimally exposed to health shocks, whenever his risk aversion to morbidity risk γ_s is lower than one, positive whenever γ_s is larger than one, and zero if and only $\gamma_s = 1$ in which case the agent's preferences are myopic towards morbidity risk.[18]

The unrestricted version of the model with endogenous mortality and morbidity generates much richer comparative statics than the restricted version of the model and can potentially address all of its shortcomings. In particular, the unrestricted model predicts that, in accordance with the data, the optimal health investment increases with the agent's financial wealth, and that the marginal utility of wealth increases with the agent's health. Moreover, the unrestricted model can be consistent with the empirical comparative statics of the optimal health insurance coverage, the optimal consumption and the optimal portfolio under appropriate parametric restrictions. Verifying whether or not these restrictions are satisfied requires an empirical analysis of the model to which we now turn.

4 Empirical performance

We adopt a dual approach in order to assess the empirical performance of the endogenous health risks model. First, we perform in Section 4.1 a structural estimation of the model and use the resulting parameter estimates to compute predicted rules that are compared to the observed allocations in Section 4.2. Second, we use our estimated parameters to compute the expected longevity, and the values of health and life, and contrast these results with received estimates in the literature in Section 4.3.

[18]Note that while the exposure of the agent's total wealth to morbidity shocks can be positive, negative or zero depending on the value of his morbidity risk aversion parameter, the exposure of his indirect utility $\Delta_s V_1(W_t, H_t) = -1_{\{dQ_{st}\neq 0\}}\chi(\xi_m)(1-\chi(\xi_m))\Theta L_m(H_{t-})^2 N_{0t-}$ is strictly negative for all positive values of the morbidity risk aversion parameter.

Throughout our empirical analysis, we rely on a sample of 30'961 individuals drawn from the Panel Study of Income Dynamics (PSID). The construction of our sample is detailed in Appendix C and summary statistics are presented in Table 1.

4.1 Parameter identification

The econometric model corresponding to the approximate optimal rules in Theorem 2 can be written as:

$$D_j = \left(\beta_N + \beta_m H_j^{-\xi_m}\right)\left(W_j + BH_j + \frac{y_0 - a}{r}\right) + \left(\beta_P + \beta_s H_j^{-\xi_s}\right) BH_j + \epsilon_j, \tag{35}$$

where $D_j' = (c_j, \pi_j, x_j, I_j)$ denotes the decisions of agent $j = 1$ in our sample, $\epsilon_j \in \mathbb{R}^4$ is an error term, and the theoretical restrictions on the coefficients $\beta_N, \beta_m, \beta_P, \beta_s \in \mathbb{R}^4$ are summarized as follows:

$$\begin{array}{c|ccccc} D & \beta_N & \beta_m & \beta_P & \beta_s \\ \hline c & A & A(1-\varepsilon)L_m(1) & 0 & AL_s(1) \\ \pi & \frac{\theta}{\gamma\sigma_S} & 0 & 0 & \frac{\theta}{\gamma\sigma_S}L_s(1) \\ x & 0 & \chi(\xi_m)\left(1-\frac{1}{\gamma_s}\right)L_m(1) & \phi & \chi(\xi_s-1)L_s(1) \\ I & 0 & \frac{-\xi_m K}{1-\alpha}L_m(1) & K & \frac{-(\xi_s-1)K}{1-\alpha}L_s(1) \end{array} \tag{36}$$

The structural estimation of the above model is challenging for two reasons. First the model is characterized by a large number of parameters (4 budget constraint parameters, 5 preference parameters and 10 survival and health dynamics parameters). Second, the model involves important nonlinearities in both variables and parameters as can be seen from (35) and (36). As a result of these difficulties it appears that all the parameters of the model cannot be simultaneously identified and we therefore decided to calibrate a subset of the parameters. Specifically, we calibrate the financial parameters (r, μ, σ_S) as well as the subjective discount rate ρ, the minimal consumption a, and the health independent part of income y_0 at realistic values from the financial economics literature. Next, a thorough search procedure was used for the calibration of the health sensitivity of income β, the risk aversion to health shocks γ_s, the maximum illness intensity η, and the two convexity parameters ξ_m, ξ_s.

Given this calibrated subset of eleven parameters, we resort to a Maximum Likelihood estimation of the quadrivariate system given by (35) and (36) for the remaining eight parameters. In estimating this system we use the wealth of individuals scaled by 10^{-4} and encode the individual self-reported health levels using a scale from 1.5 (Poor health) to 3.5 (Excellent health) with an increment of 0.5. To guarantee theoretical consistency, the estimation is performed subject to the transversality conditions (13), (19), (29), (30), and the auxiliary restriction (41) that is discussed below, and ensures that life expectancy predicted by the model is finite. We report the calibrated and estimated parameters in Table 2. Importantly, all our estimates are significant at standard levels, and all the sign and required theoretical restrictions are satisfied.

The calibrated financial parameters in Panel A are conventional. In particular, we set the interest rate to $r = 0.048$, the expected stock return to $\mu = 0.108$ and the stock volatility to $\sigma_S = 0.20$ so that the market price of risk is $\theta = 0.30$. Turning to the preference parameters in Panel B, we fix the minimal consumption $a = 0.69$ at a slightly higher level than health-independent revenues $y_0 = 0.68$. The calibrated value of the subjective discount rate $\rho = 0.05$ is standard for PSID studies,[19] and we set the aversion to morbidity risks $\gamma_s = 7.4$ at a relatively high, but nonetheless reasonable, level. Since the latter is greater than one, our estimation results are consistent with a positive dynamic hedging demand in health insurance. Our estimate for the aversion to financial risk, $\gamma = 2.6$, is very realistic (e.g., Mehra and Prescott, 1985), and much lower than γ_s. The estimated aversion to mortality risk, $\gamma_m = 0.68$, is lower than one as required by the model and implies an important penalization for mortality risk $\Phi_m = \gamma_m/(1-\gamma_m) = 2.16$ that corresponds to a threefold increase of the endogenous discount rate in (11). Finally, the low estimate of the elasticity of intertemporal substitution $\varepsilon = 0.65$ is consistent with previous estimates in the literature[20] and indicates that agents tend to decrease consumption in response to an increase in mortality risk.

In Panel C, the calibrated values of the parameters $\xi_m = 1.8$ and $\xi_s = 4.9$ point towards significant convexities in adjustment of health-related risks. The maximal health shock intensity $\eta = 50$ ensures that the agent is near certain to become sick as health falls to zero (see Figure 1). The exogenous death intensity parameter $\lambda_{m0} = 0.024$ reproduces

[19] See Alan and Browning (2010, Tab. 7) or Alan et al. (2009, Tab. IV) for recent examples.

[20] See Engelhardt and Kumar (2009); Lee (2008); Biederman and Goenner (2008); Saltari and Ticchi (2007); Vissing-Jørgensen (2002), among others for recent EIS estimates.

a maximal remaining longevity of 42 years (see the discussion in Section 4.3.1), whereas the exogenous sickness intensity $\lambda_{s0} = 1.21$ corresponds to one event every 10 months for perfectly healthy agents. Given these calibrated values our estimates for the the endogenous intensity parameters $\lambda_{m1} = 0.002$ and $\lambda_{s1} = 0.02$ are low, but both significant. This validates the approximation method we use to solve the model, and also confirms that the model with endogenous health risk is preferable. Furthermore, the fact that both $\xi_m < \xi_s$ and $\lambda_{m1} < \lambda_{s1}$ is consistent with the intuition that mortality risk is more difficult to adjust than morbidity risk. The estimated share of health lost upon experiencing a health shock $\phi = 1.1\%$ is nontrivial and represents twice the depreciation rate of health $\delta = 0.55\%$. Finally, the estimated value of the Cobb-Douglas parameter $\alpha = 0.77$ points towards strong convexities in the health adjustment process.

These structural parameter estimates provide key insights into the main features of the model. First, the parameters of the mortality and morbidity intensities allow us to gauge the relevance of the endogenous and exogenous health risks. While a sizeable share of these risks is captured by the incompressible part of the arrival rates, we find that agents can indeed adjust both types of risks through health improving investments. Second, the parameters that govern the dynamics of health confirm that health is subject to both proportional depreciation and morbidity shocks, and show that while agents can adjust their health through investment they face strongly diminishing returns in doing so. Finally, the estimated risk aversion parameters confirm that agents have non time-additive preferences that display source dependent risk aversion.

4.2 Predicted and observed allocations

In order to compare the predictions of the estimated model to the observed rules we proceed as follows. First, we use the parameter values of Table 2 to calculate the predicted consumption, portfolio, insurance and health investment at the observed health and wealth levels for all agents in our sample. Second, we compute the predicted sample average in each health and wealth quintile and contrast those with the data averages. Table 3 shows the results for consumption and portfolio holdings while Table 4 shows the results for health insurance and health investment.

The observed consumption schedules in Panel A.1 are clearly increasing in both health and wealth.[21] The estimated consumption in Panel A.2 reproduces the signs of the gradients and provides a reasonable fit, keeping in mind the caveats for the implied PSID consumption data.[22] Similarly, the observed stock holdings in Panel B.1 are increasing in both health and wealth.[23] Both the levels of observations and the signs of the health and wealth gradients are well captured by the estimated model in Panel B.2. Interestingly, the estimated model predicts negative stock holdings positions for poor and unhealthy agents and large stock holdings for poor but healthy agents. While the former may indicate that poor and unhealthy agents engage is risk shifting activities, the latter is likely a reflection of the well-known "participation puzzle" according to which low-wealth individuals do not take active positions in stock markets (e.g. Vissing-Jørgensen, 2002; Brav et al., 2002; Gormley et al., 2010). Health-related risks alone are apparently unable to account for this salient feature of the data.

Turning to medical variables, Panel C.1 of Table 4 shows that in accordance with previous studies the observed health insurance increases in wealth but is non monotonic in health.[24] The predicted insurance levels in Panel C.2 correctly capture these features but are lower than observed, indicating that other elements that we abstracted from, such as employer-provided health plans, are possibly at stake. Finally, the observed health expenditures in Panel D.1 fall sharply in health and increase in wealth.[25] Once again the estimated model in Panel D.2 performs reasonably well in reproducing both the range of observations and the signs of gradients.

Overall, we find that the calibrated and estimated parameters satisfy the theoretical restrictions, and are realistic. Furthermore, the comparison of fitted versus actual data shows that the model is able to reproduce both observed allocations and comparative statics with respect to health and wealth.

[21] See among others Smith (1999); Gertler and Gruber (2002); Domeij and Johannesson (2006) for the effect of health on consumption, and Gourinchas and Parker (2002); Dynan et al. (2004); Jappelli and Pistaferri (2010) for the effect of wealth.

[22] As discussed in Appendix C, the observed consumption is implied from a small number of measures (food, utility and transportation), and therefore likely measured with considerable error.

[23] See among others Rosen and Wu (2004); Berkowitz and Qiu (2006); Coile and Milligan (2009) for the effect of health on portfolio holdings, and Brunnermeier and Nagel (2008); Calvet and Sodini (2010); Wachter and Yogo (2010) for the effect of wealth.

[24] Cardon and Hendel (2001); Kaestner and Kaushal (2003); Barros et al. (2008); Yang et al. (2009); Khwaja (2010) discuss health and wealth effects on the demand for insurance.

[25] See Smith (1999); Wu (2003); Gilleskie and Mroz (2004); Smith (2007); Barros et al. (2008); Yang et al. (2009); Marshall et al. (2010) for evidence and discussion of health and wealth effects on health expenditures.

4.3 Additional performance measures

To further assess the quantitative performance of our estimated model, we now investigate its predictions concerning the value of health, the expected longevity of agents and the value that they attribute to additional years of life expectancy.

4.3.1 Value of health, expected longevity and value of life

The explicit expression for the indirect utility in (31) makes it possible to compute the implied value of health and longevity by determining the amount of wealth that an agent would be willing to give-up to improve either his health or his life expectancy. In the spirit of the Hicksian compensating variation, we define the value of n additional units of health as the solution

$$w_h = w_h(n, W_t, H_t)$$

to the indifference equation

$$V(W_t - w_h, H_t + n) = V(W_t, H_t). \tag{37}$$

The following proposition relies on an expansion technique similar to that of Theorem 2 to derive a first order approximation for the value of health.

Proposition 3 (Value of health) *Assume that the conditions of Theorem 2 hold true and define a pair of nonnegative functions by setting*

$$J_m(n, H) = L_m(H + n) - L_m(H),$$
$$J_s(n, H) = L_s(H + n)P_0(H + n) - L_s(H)P_0(H).$$

Then up to a first order approximation the value of n units of health is

$$w_h(n, W_t, H_t) = nB + J_m(n, H_t)N_{0t} + J_s(n, H_t), \tag{38}$$

where the constant B and total wealth N_{0t} are defined as in Proposition 1.

To determine the value of longevity that is implied by the model, we compute the amount of wealth w_ℓ that the agent would be willing to give-up to increase his life expectancy by a fixed amount. More precisely, if

$$\ell(W,H) = \ell(W,H;(\lambda_{m0},\lambda_{m1})) = E[T_m]$$

denotes the life expectancy of an alive agent with wealth W, health status H and mortality parameters $(\lambda_{m0},\lambda_{m1})$, then we define the value of n additional years of expected lifetime as the solution

$$w_\ell = w_\ell(n,W,H)$$

to the indifference equation

$$V(W-w_\ell,H;(\lambda^*_{m0},\lambda_{m1})) = V(W,H;(\lambda_{m0},\lambda_{m1})) \tag{39}$$

where the modified incompressible death intensity $\lambda^*_{m0} = \lambda^*_{m0}(n,W,H)$ is computed in such a way as to guarantee that the agent's life expectancy after the transfer has increased by exactly n years:

$$\ell(W-w_\ell,H;(\lambda^*_{m0},\lambda_{m1})) = n + \ell(W,H;(\lambda_{m0},\lambda_{m1})). \tag{40}$$

The following proposition relies on an expansion technique similar to that of Theorem 2 to derive first order approximations for both the life expectancy and the value of life implied by the theoretical model.

Proposition 4 (Life expectancy and value of life) *Assume that the conditions of Theorem 2 hold true and that*

$$1/\kappa_0 = \lambda_{m0} - F(-\xi_m) > 0. \tag{41}$$

Then up to a first order approximation, the agent's life expectancy and the value of n additional years of life expectancy are given by

$$\ell(W_t,H_t;(\lambda_{m0},\lambda_{m1})) = (1/\lambda_{m0})(1-\kappa_0\lambda_{m1}H_t^{-\xi_m}), \tag{42}$$

and

$$w_\ell(n, W_t, H_t) = (1 - q^*(n))N_{1t} + q^*(n)Q^*(n, H_t)N_{0t} \tag{43}$$

where N_{1t} *is the approximation of the agent's total wealth given in Proposition 2 and the functions* $q^*(n) \in (0, 1)$ *and* $Q^*(n, H)$ *are defined by*

$$\begin{aligned} q^*(n) &= \left(\frac{A}{A^*(n)}\right)^{\frac{1}{1-\varepsilon}}, \\ Q^*(n, H) &= \lambda_{m1}\left(R^*(n) + \frac{1}{(1-\gamma_m)(F(-\xi_m) - A^*(n))}\right) H^{-\xi_m} - L_m(H), \end{aligned}$$

with

$$\begin{aligned} A^*(n) &= A + \frac{1-\varepsilon}{1-\gamma_m}\left(\frac{n\lambda_{m0}^2}{1+n\lambda_{m0}}\right), \\ R^*(n) &= \frac{n}{(1-\gamma_m)A^*(n)}\left(\frac{\lambda_{m0}}{1+n\lambda_{m0}}\right)^2 \frac{\kappa_0 + \kappa_0\lambda_{m0}(n+\kappa_0)}{1+n\lambda_{m0}(1-\kappa_0\lambda_{m0})}, \end{aligned}$$

and the functions $F(-\xi_m)$ *and* $L_m(H)$ *are defined in Theorem 2.*

4.3.2 Empirical estimates

Taking the parameter estimates of Table 2, we compute the value of health (38), remaining expected lifetime (42), and the value of one year of additional life expectancy for all agents in our sample at the observed health and wealth levels. The sample averages are then computed per wealth quintiles and health status and are reported in Table 5. We report comparisons with other estimates in the literature in Table 6.

The first panel of Table 5 shows that the willingness to pay for a unit increment in health status is non trivial, with an average value of 8% of annual income, and that consistent with economic intuition the value of health increases in wealth and falls rapidly as the agent's wealth deteriorates. When compared to our estimates for the marginal value of health $B =$ \$3'066, the reported values of health indicate that between 0.50 (for healthy agent) and 0.75 (for unhealthy agent) of the value of health can be attributed to the capacity in adjusting health-related risks (J_m, J_s). These estimates are also quite realistic in view of received estimates in the literature. For example, Smith (2005) uses survey data to compute the willingness to pay in percent of annual income to prevent a given relative decrease in health from an excellent health state. In our model, the same willingness to pay can be computed as

$$w_h\left(\alpha H_e, W, (1-\alpha)H_e\right)/Y$$

where α is a given percentage change, $H_e = 3.5$ denotes the benchmark state of excellent health and Y denotes the observed annual income level. As reported in Panel A of Table 6, the estimated values obtained from the model provide a close match of both the observed levels and the observed gradients.

The remaining life expectancies reported in Panel B of Table 5 are also very realistic. Indeed, the average age in our sample is 44 and when restricted to the 789 agents of that age, the unconditional expected lifetime[26] is 80.84 years, halfway between the national values of 78.22 years for males and 82.17 for females aged 44 (Social Security Administration, 2007). Moreover, we find that longevity is independent of the wealth level, and increases when health improves. Both facts are consistent with previous empirical findings.[27] The fact that stock holdings and life expectancy both increase with the agent's health implies that, in accordance with the horizon effects documented by **?** among others, stock holdings and longevity are positively related. Finally, we note that the magnitude of the health gradient implied by our estimated parameters is realistic. In particular, Panel B of Table 6 compares our estimates of longevity to those obtained by Lubitz et al. (2003) for agents aged 70 and shows that the estimated health gradients provide a close match of the observations.

The third panel of Table 5 shows that the willingness to pay for an additional year of life expectancy is significant, with an average value of 12% of annual income. Interestingly, this panel shows that while the value of life is increasing in wealth, it is either increasing or decreasing in health depending on whether agents are poor or rich. To understand this mixed effect recall from Section 3 that a decrease in the exogenous mortality intensity impacts the agent's utility through two channels. First, a decrease in λ_{m0} implies an increase in the marginal utility of wealth and health at order zero and thereby prompts healthier agents to pay more for a given reduction. Second, a decrease in λ_{m0} lowers the

[26]The expected lifetime of an agent is obtained by summing the agent's age and his remaining life expectancy computed according to (42).

[27]See Benitez-Silva and Ni (2008, Table 4) and Hurd et al. (2001, Table 20).

utility penalty associated to endogenous mortality and, because of convexity, this effect becomes relatively more pronounced as the agent's heath deteriorates. As shown by the figures in Table 5 the net effect increases with the agent's health for poor agents and decreases for rich agents. This result shows that the mortality control feature associated with endogenous mortality is more important for rich agents than it is for poor agents who are more focused on the budget constraint effects of health.

Overall, we conclude that the model offers a remarkable empirical performance. In particular, the predicted rules evaluated at realistic estimated parameters are plausible, both in terms of levels and of comparative statics, and the model also generates accurate predictions for the values of life and longevity, as well as for life expectancy.

5 Conclusion

This paper shows that the complex interactions between financial and health-related statuses and allocations can be jointly explained by a parsimonious model that combines two baseline frameworks from the Health and Financial Economics literature with a novel specification of health-related risks, and preferences.

The analytical solutions that we obtain and estimate are easy to interpret and confirm that endogenous mortality and morbidity risks, a positive health elasticity of labor income as well as convex health adjustment costs, and recursive preferences with source-dependent risk aversion are all key ingredients in better understanding how risks and resources condition financial and health-related choices.

References

Abel, Andrew B., and Janice C. Eberly (1994) 'A unified model of investment under uncertainty.' *American Economic Review* 84(5), 1369–1384

Alan, Sule, and Martin Browning (2010) 'Estimating intertemporal allocation parameters using synthetic residual estimation.' *Review of Economic Studies* 77(4), 1231 – 1261

Alan, Sule, Orazio Attanasio, and Martin Browning (2009) 'Estimating euler equations with noisy data: Two exact gmm estimators.' *Journal of Applied Econometrics* 24(2), pp. 309–324

Barros, Pedro Pita, Matilde P. Machado, and Anna Sanz-de Galdeano (2008) 'Moral hazard and the demand for health services: A matching estimator approach.' *Journal of Health Economics* 27(4), 1006 – 1025

Becker, Gary S., Tomas J. Philipson, and Rodrigo R. Soares (2005) 'The quantity and quality of life and the evolution of world inequality.' *American Economic Review* 95(1), 277–291

Benitez-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

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

Biederman, Daniel K., and Cullen F. Goenner (2008) 'A life-cycle approach to the intertemporal elasticity of substitution.' *Journal of Macroeconomics* 30(1), 481–498

Bodie, Zvi, Robert C Merton, and William F Samuelson (1992) 'Labor supply flexibility and portfolio choice in a life cycle model.' *Journal of Economic Dynamics and Control* 16(3-4), 427–449

Bommier, Antoine (2006) 'Uncertain lifetime and intertemporal choice: Risk aversion as a rationale for time discounting.' *International Economic Review* 47(4), 1223 – 1246

——— (2010) 'Portfolio choice under uncertain lifetime.' *Journal of Public Economic Theory* 12(1), 57 – 73

Bommier, Antoine, and Jean-Charles Rochet (2006) 'Risk aversion and planning horizons.' *Journal of the European Economic Association* 4(4), 708 – 734

Brav, Alon, George M. Constantinides, and Christopher C. Gezcy (2002) 'Asset pricing with heterogeneous consumers and limited participation: Empirical evidence.' *Journal of Political Economy* 110(4), 793–824

Brunnermeier, Markus K., and Stefan Nagel (2008) 'Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals' asset allocation.' *American Economic Review* 98(3), 713 – 736

Calvet, Laurent E., and Paolo Sodini (2010) 'Twin picks: Disentangling the determinants of risk-taking in household portfolios.' Working Paper 15859, National Bureau of Economic Research

Campbell, John Y., and Luis M. Viceira (2002) *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors* (Oxford University Press)

Cardon, James H., and Igal Hendel (2001) 'Asymmetric information in health insurance: Evidence from the national medical expenditure survey.' *RAND Journal of Economics* 32(3), 408 – 427

Chang, Fwu-Ranq (1996) 'Uncertainty and investment in health.' *Journal of Health Economics* 15(3), 369–376

——— (2005) 'A theory of health investment under competing mortality risks.' *Journal of Health Economics* 24(3), 449–463

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

Domeij, David, and Magnus Johannesson (2006) 'Consumption and health.' *B.E. Journal of Macroeconomics: Contributions to Macroeconomics* 6(1), 1 – 30

Duffie, Darrell, and Larry G. Epstein (1992a) 'Asset pricing with stochastic differential utility.' *Review of Financial Studies* 5(3), 411–436

——— (1992b) 'Stochastic differential utility.' *Econometrica* 60(2), 353–394

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes (2004) 'Do the rich save more?.' *Journal of Political Economy* 112(2), 397 – 444

Edwards, Ryan D. (2008) 'Health risk and portfolio choice.' *Journal of Business and Economic Statistics* 26(4), 472–485

Ehrlich, Isaac (2000) 'Uncertain lifetime, life protection and the value of life saving.' *Journal of Health Economics* 19(3), 341–367

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

Ehrlich, Isaac, and Yong Yin (2005) 'Explaining diversities in age-specific life expectancies and values of life saving: A numerical analysis.' *Journal of Risk and Uncertainty* 31(2), 129–162

Engelhardt, Gary V., and Anil Kumar (2009) 'The elasticity of intertemporal substitution: New evidence from 401(k) participation.' *Economics Letters.* forthcoming

Epstein, Larry G., and Stanley E. Zin (1989) 'Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework.' *Econometrica* 57(4), 937–969

Finkelstein, Amy, Erzo F. P. Luttmer, and Matthew J. Notowidigdo (2009) 'Approaches to estimating the health state dependence of the utility function.' *American Economic Review* 99(2), 116 – 121

Finkelstein, Amy, Erzo F.P. Luttmer, and Matthew J. Notowidigdo (2008) 'What good is wealth without health? the effect of health on the marginal utility of consumption.' Working Paper 14089, National Bureau of Economic Research, June

French, Eric (2005) 'The effects of health, wealth, and wages on labour supply and retirement behaviour.' *Review of Economic Studies* 72(2), 395–427

Gertler, Paul, and Jonathan Gruber (2002) 'Insuring consumption against illness.' *American Economic Review* 92(1), 51 – 76

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

Gormley, Todd, Hong Liu, and Guofu Zhou (2010) 'Limited participation and consumption-saving puzzles: A simple explanation and the role of insurance.' *Journal of Financial Economics* 96(2), 331 – 344

Gourinchas, Pierre-Olivier, and Jonathan A. Parker (2002) 'Consumption over the life cycle.' *Econometrica* 70(1), 47–89

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

Guo, Sheng (2010) 'The superior measure of PSID consumption: An update.' *Economics Letters* 108(3), 253 – 256

Hakansson, Nils H. (1969) 'Optimal investment and consumption strategies under risk, an uncertain lifetime, and insurance.' *International Economic Review* 10(3), 443–466

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

Hayashi, Fumio (1982) 'Tobin's marginal q and average q: A neoclassical interpretation.' *Econometrica* 50(1), 213–224

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., 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

Jappelli, Tullio, and Luigi Pistaferri (2010) 'The consumption response to income changes.' Working paper 15739, National Bureau of Economic Research

Kaestner, Robert, and Neeraj Kaushal (2003) 'Welfare reform and health insurance coverage of low-income families.' *Journal of Health Economics* 22(6), 959 – 981

Khwaja, Ahmed (2010) 'Estimating willingness to pay for medicare using a dynamic life-cycle model of demand for health insurance.' *Journal of Econometrics* 156(1), 130 – 147

Kreps, David M., and Evan L. Porteus (1979) 'Temporal von Neumann-Morgenstern and induced preferences.' *Journal of Economic Theory* 20(1), 81–109

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

Lazrak, Ali, and Marie Claire Quenez (2003) 'A generalized stochastic differential utility.' *Mathematics of Operations Research* 28(1), 154–180

Lee, Chul-In (2008) 'On-the-job human capital investment and intertemporal substitution: New evidence on intertemporal substitution elasticity.' *Journal of Economic Dynamics and Control* 32(10), 3350–3375

Lubitz, James, Liming Cai, Ellen Kramarow, and Harold Lentzner (2003) 'Health, life expectancy, and health care spending among the elderly.' *New England Journal of Medicine* 349(11), 1048–1055

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

Mehra, Rajnish, and Edward C. Prescott (1985) 'The equity premium: A puzzle.' *Journal of Monetary Economics* 15(2), 145–161

Merton, Robert C. (1971) 'Optimum consumption and portfolio rules in a continuous-time model.' *Journal of Economic Theory* 3(4), 373–413

Obstfeld, Maurice (1994) 'Risk-taking, global diversification, and growth.' *American Economic Review* 84(5), 1310–1329

Picone, Gabriel, Martín Uribe, and R. Mark Wilson (1998) 'The effect of uncertainty on the demand for medical care, health capital and wealth.' *Journal of Health Economics* 17(2), 171–185

Richard, Scott F. (1975) 'Optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous time model.' *Journal of Financial Economics* 2(2), 187–203

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

Rosen, Sherwin (1988) 'The value of changes in life expectancy.' *Journal of Risk and Uncertainty* 1(3), 285–304

Saltari, Enrico, and Davide Ticchi (2007) 'Risk aversion, intertemporal substitution, and the aggregate investment–uncertainty relationship.' *Journal of Monetary Economics* 54(3), 622–648

Shepard, Donald S., and Richard J. Zeckhauser (1984) 'Survival versus consumption.' *Management Science* 30(4), 423–439

Skiadas, Costis (2007) 'Dynamic portfolio choice and risk aversion.' In *Financial Engineering,* ed. John R. Birge and Vadim Linetsky, vol. 15 of *Handbooks in Operations Research and Management Science* (Elsevier) chapter 19, pp. 789 – 843

——— (2009) 'Smooth ambiguity aversion toward small risks and continuous-time recursive utility.' manuscript, Kellogg School of Management, Department of Finance, Northwestern University

Skinner, Jonathan (1987) 'A superior measure of consumption from the panel study of income dynamics.' *Economics Letters* 23(2), 213 – 216

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

——— (2009) 'The impact of childhood health on adult labor market outcomes.' *Review of Economic and Statistics* 91(3), 478–489

Smith, Richard D. (2005) 'Sensitivity to scale in contingent valuation: the importance of the budget constraint.' *Journal of Health Economics* 24(3), 515 – 529

Smith, William T. (1996) 'Feasibility and transversality conditions for models of portfolio choice with non-expected utility in continuous time.' *Economics Letters* 53(2), 123–131

Social Security Administration (2007) 'Period life tables.' In 'Actuarial Publications' (http://www.ssa.gov/oact/STATS/table4c6.html)

Svensson, Lars E. O. (1989) 'Portfolio choice with non-expected utility in continuous time.' *Economics Letters* 30(4), 313–317

Uzawa, Hirofumi (1969) 'Time preference and the penrose effect in a two-class model of economic growth.' *Journal of Political Economy* 77(4), 628–652. Part 2: Symposium on the Theory of Economic Growth

Vissing-Jørgensen, Annette (2002) 'Limited asset market participation and the elasticity of intertemporal substitution.' *Journal of Political Economy* 110(4), 825–853

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

Weil, Philippe (1989) 'The equity premium puzzle and the risk-free rate puzzle.' *Journal of Monetary Economics* 24(3), 401–421

Wu, Stephen (2003) 'Sickness and preventive medical behavior.' *Journal of Health Economics* 22(4), 675 – 689

Yaari, M. (1965) 'Uncertain lifetime, life insurance and the theory of the consumer.' *Review of Economic Studies* 52, 137–150

Yang, Zhou, Donna B. Gilleskie, and Edward C. Norton (2009) 'Health insurance, medical care, and health outcomes: A model of elderly health dynamics.' *Journal of Human Resources* 44(1), 47 – 114

Yogo, Motohiro (2008) 'Portfolio choices in retirement: Health risk and the demand for annuities, housing and risky assets.' manuscript, Finance Department, The Wharton School, University of Pennsylvania, January

——— (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

Table 1: Descriptive statistics of the PSID sample

Variable	Mean	Std. Dev.	Min	Max
Age	44.18	15.51	16	101
Health	2.79	0.54	1.5	3.5
Wealth	$28'356	$79'293	$0	$1'183'728
Income	$41'199	$58'502	$0	$3'290'784
Consumption	$9'838	$10'126	$1	$433'838
Stocks	$15'139	$70'194	$0	$1'157'273
Health insurance	$522	$903	$0	$15'879
Health expenditures	$541	$2'195	$0	$120'704

Notes: This table presents summary statistics for the main variables in the sample of 30'961 individuals that we use in our estimation. Nominal variables are in dollars while the self-reported health status is encoded using a discrete scale between 1.5 (Poor health) and 3.5 (Excellent health) with an increment of 0.5 between two consecutive health status.

Table 2: Calibrated and estimated structural parameters

Symbol	Interpretation	Calibrated	Estimated Value	(Std. Error)
		Panel A. Budget constraint		
β	health dependence in income	0.020		
y_0	health independent income	0.68		
r	risk-free rate interest	0.048		
μ	expected risky return	0.108		
σ_S	std. error risky returns	0.20		
		Panel B. Preferences		
a	minimal consumption	0.69		
ρ	subjective discount rate	0.05		
γ_s	aversion to morbidity risk	7.40		
γ	aversion to financial risk		2.5968	(0.0292)
γ_m	aversion to mortality risk		0.6834	(0.0384)
ε	elasticity intertemp. subst.		0.6465	(0.0523)
		Panel C. Survival and health dynamics		
ξ_m	mortality intensity convexity	1.80		
ξ_s	morbidity intensity convexity	4.90		
η	max. exog. morbidity intensity	50.00		
λ_{m0}	min. exog. mortality intensity		0.0237	(0.0043)
λ_{s0}	min. exog. morbidity intensity		1.2098	(0.1125)
λ_{m1}	endo. mortality intensity parameter		0.0017	(0.0008)
λ_{s1}	endo. morbidity intensity parameter		0.0198	(0.0012)
ϕ	depreciation upon health shock		0.0110	(0.0012)
δ	deterministic health depreciation rate		0.0055	(0.0013)
α	Cobb-Douglas param. health process		0.7742	(0.0085)

Notes: The estimated parameters are Maximum Likelihood estimates for the quadrivariate system (35) subject to the theoretical restrictions (36), (13), (19), (29), (30), (41).

Table 3: Actual and predicted financial variables (in $)

Health	Wealth quintiles				
	1	2	3	4	5
	Panel A.1 Consumption: Data				
Poor	3'635	4'899	8'374	9'328	12'048
Fair	4'441	6'084	9'603	11'475	13'764
Good	5'919	7'109	9'702	12'139	14'327
Very good	6'403	7'299	10'057	11'815	15'065
Excellent	6'697	7'130	10'284	12'164	15'516
	Panel A.2 Consumption: Predicted				
Poor	6'801	6'803	6'836	7'080	10'661
Fair	6'955	6'958	6'993	7'246	10'287
Good	7'031	7'035	7'072	7'327	10'347
Very good	7'087	7'091	7'130	7'395	10'579
Excellent	7'137	7'141	7'181	7'444	10'739
	Panel B.1 Stock holdings: Data				
Poor	0	3	30	1'687	54'529
Fair	0	2	83	2'261	64'712
Good	0	2	186	3'062	60'876
Very good	0	4	170	3'310	70'450
Excellent	0	8	201	3'893	82'408
	Panel B.2 Stock holdings: Predicted				
Poor	−1'809	−1'750	−995	4'572	86'380
Fair	1'277	1'347	2'093	7'566	73'278
Good	2'780	2'857	3'640	9'018	72'591
Very good	3'892	3'978	4'784	10'289	76'350
Excellent	4'876	4'962	5'791	11'184	78'949

Notes: The observed rules are sample averages using pooled data from PSID (30'961 individuals) described in Appendix C. Predicted rules are sample averages of the optimal rules of Theorem 2 evaluated at the parameter values of Table 2 and using individual PSID data on wealth and health. All reported values are expressed in dollars.

Table 4: Actual and predicted health-related variables (in $)

Health	Wealth quintiles				
	1	2	3	4	5
Panel C.1 Health insurance: Actual					
Poor	260	279	560	818	1'192
Fair	209	226	573	897	1'112
Good	223	280	445	689	1'040
Very good	244	280	448	660	958
Excellent	213	262	429	556	917
Panel C.2 Health insurance: Predicted					
Poor	315	316	324	382	1'245
Fair	163	164	168	203	616
Good	132	132	136	158	426
Very good	130	131	133	150	350
Excellent	139	139	141	153	309
Panel D.1 Health expenditures: Data					
Poor	951	655	2'368	1'933	5'218
Fair	448	440	789	808	1'746
Good	306	322	402	616	1'001
Very good	215	271	362	503	915
Excellent	176	198	277	343	612
Panel D.2 Health expenditures: Predicted					
Poor	610	611	613	782	2'990
Fair	249	250	262	350	1'407
Good	152	153	161	219	903
Very good	120	121	127	169	681
Excellent	109	110	115	146	544

Notes: The observed rules are sample averages using pooled data from PSID (30'961 individuals) described in Appendix C. Predicted rules are sample averages of the optimal rules of Theorem 2 evaluated at the parameter values of Table 2 and using individual PSID data on wealth and health. All reported values are expressed in dollars.

Table 5: Predicted life expectancy and predicted values of health and life

Health	Wealth quintiles				
	1	2	3	4	5
	Panel A. Value of health (in $)				
Poor	5'699	5'713	5'898	7'263	27'324
Fair	2'882	2'890	2'979	3'634	11'491
Good	2'143	2'149	2'202	2'566	6'868
Very good	1'878	1'881	1'915	2'148	4'937
Excellent	1'758	1'761	1'784	1'936	3'847
	Panel B. Remaining life expectancy (in years)				
Poor			27.76		
Fair			33.53		
Good			36.35		
Very good			37.84		
Excellent			38.93		
	Panel C. Value of longevity (in $)				
Poor	445	476	877	3'837	47'322
Fair	729	754	1'021	2'976	26'458
Good	834	855	1'068	2'533	19'849
Very good	900	919	1'100	2'335	17'153
Excellent	959	975	1'136	2'184	15'352

Notes: The value of health w_h in Panel A is the willingness to pay for $n = 0.5$ additional unit of health (corresponding to a change in polytomial status) computed according to (38). The remaining life expectancy ℓ in Panel B is the life expectancy computed according to (42). The value of longevity w_ℓ in Panel C is the willingness to pay for $n = 1$ year of additional life expectancy computed according to (43). All the reported quantities are sample averages computed at the parameters values of Table 2 using individual PSID data on wealth and health.

Table 6: Comparisons of the model predictions with other estimates

Panel A. WTP in % of income to avoid reduction in health		
α	Smith (2005)	Estimated median w_h/Y
6.0%	1.8%	2.2%
13.0%	3.4%	4.9%
22.0%	9.4%	8.9%
28.0%	14.5%	11.9%
40.0%	18.8%	19.7%

Panel B. Expected conditional longevity		
Health	Lubitz et al. (2003) (base age 70)	Estimated mean ℓ (base age 44.18)
Poor	79.2	71.9
Fair	81.3	77.7
Good	82.6	80.5
Very Good	83.4	82.1
Excellent	83.8	83.1

Notes: Panel A: Smith (2005, Tab. 2 and 3, p. 518 and 521) and the estimated median of $w_h(\alpha H_e, W, (1-\alpha)H_e)/Y$ conditional upon non-zero income and base health $H_e = 3.5$. Panel B: Lubitz et al. (2003, Fig. 2, p. 1052) and the estimated mean of the sum of the base age and the expected remaining lifetime $\ell(H)$ calculated from (42).

A Proofs

To simplify the presentation of the proofs we assume throughout this appendix that the agent's subsistence consumption a and the health independent part of his income y_0 are both equal to zero. Since the agent faces complete markets when solving the modified problem (10), the general case can be obtained from this one by adding a to the optimal consumption and the present value

$$E_t \int_t^\infty m_{t,\tau}(y_0 - a)d\tau = \frac{y_0 - a}{r}$$

of the corresponding cash flow streams to the agent's financial wealth.

Proof of Proposition 1. Let Q denote the risk neutral measure defined by

$$\left.\frac{\mathrm{d}Q}{\mathrm{d}P}\right|_t = e^{rt} m_t.$$

where m is the state price density process of equation (12). Using the independence between market and morbidity shocks it is immediate to show that the function P_0 is given by

$$P_0(H_t) = \sup_{I \geq 0} E^Q \int_0^\infty e^{-rs}(\beta H_{\tau-} - I_\tau)d\tau = \sup_{I \geq 0} E \int_0^\infty e^{-rs}(\beta H_{\tau-} - I_\tau)d\tau$$

and satisfies the Hamilton-Jacobi-Bellman equation

$$rP_0 = \beta H + \lambda_{s0}(P_0((1-\phi)H) - P_0) + \max_{I \geq 0}(((I/H)^\alpha - \delta)HP_{0H} - I)$$

subject to the transversality condition

$$\lim_{t \to \infty} E^Q[e^{-rt}P_0(H_{0t})] = 0$$

where H_0 denotes the path of the agent's health under the optimal strategy. The dynamics of H and the linearity of the objective function imply that P_0 is increasing and homogenous of degree one with respect to health so that the value function and optimal investment policy are given by

$$P(H) = BH,$$
$$I_0(H) = H(\alpha P_{0H}(H))^{\frac{1}{1-\alpha}} = H(\alpha B)^{\frac{1}{1-\alpha}},$$

for some nonnegative constant B that solves

$$\begin{aligned} 0 &= \beta - (r + \delta + \phi\lambda_{s0})B + \max_{x\geq 0}(x^\alpha B - x) \\ &= \beta - (r + \delta + \phi\lambda_{s0})B - (1 - 1/\alpha)(\alpha B)^{\frac{1}{1-\alpha}} = g(B), \end{aligned}$$

subject to the transversality condition

$$\lim_{t\to\infty} E[e^{-rt} B H_{0t}] = 0,$$

where

$$\mathrm{d}H_{0t} = H_{0t-}\left((\alpha B)^{\frac{\alpha}{1-\alpha}} - \delta\right)\mathrm{d}t - \phi H_{0t-} dN_t$$

denotes the path of the agent's health under the candidate optimal strategy. Using the above dynamics in conjunction with basic properties of Poisson processes we obtain

$$E[e^{-rt} B H_{0t}] = e^{g'(B)t} B H_{00}, \qquad t \geq 0,$$

and it follows that the transversality condition is equivalent to $g'(B) < 0$. Straightforward analysis shows that g satisfies $g(0) = r + \delta + \phi\lambda_{s0} > 0$ as well as

$$g'(0) = -(r + \delta + \phi\lambda_{s0}) < 0$$

and attains a unique minimum over the positive real line whose value is given by

$$\min_{x\geq 0} g(x) = \beta - (r + \delta + \phi\lambda_{s0})^{\frac{1}{\alpha}}.$$

Under condition (13), this minimal value is negative and it follows that there exists a unique nonnegative B such that $g(B) = 0$ and $g'(B) < 0$. ■

Properties of B. Let $R = r + \delta + \phi\lambda_{s0}$. By definition $B = B(R) > 0$ solves the non linear equation given by

$$g(R, B(R)) = \beta - RB(R) - (1 - 1/\alpha)(\alpha B(R))^{\frac{1}{1-\alpha}} = 0.$$

and satisfies $g_B(R, B(R)) < 0$. In particular, we have

$$B(R) \leq B^o(R) = \operatorname*{argmin}_{x \geq 0} g(R, x) = R^{\frac{1-\alpha}{\alpha}}/\alpha$$

and since

$$B'(R) = \frac{g_R(R, B(R))}{g_B(R, B(R))} = \frac{B(R)/R}{(B(R)/B^o(R))^{\frac{\alpha}{1-\alpha}} - 1} \leq 0$$

we conclude that B is decreasing in R. ■

Proof of Theorem 1. When the intensity of Poisson shocks is health-independent, the agent's problem is equivalent to that of equation (18) with initial capital $N_{0t} = N_0(W_t, H_t)$. In particular, the value function and optimal controls are given by

$$V_0(W, H) = G(N_0(W, H))$$

and

$$\begin{aligned}
c_{0t} &= k_t^*, \\
\pi_{0t} &= p_t^* N_{0t}, \\
x_{0t} &= x_t^* + \phi P_0(H_{t-}), \\
I_{0t} &= (\alpha B)^{\frac{1}{1-\alpha}},
\end{aligned}$$

where (p^*, x^*, k^*) denote the optimal portfolio proportion, optimal insurance coverage and optimal consumption for the problem defined by

$$G(N_t) = \sup_{(p,x,k)} V_t(k)$$

subject to

$$\mathrm{d}N_t = rN_{t-}\mathrm{d}t + p_t N_{t-}\sigma(\mathrm{d}Z_t + \theta\mathrm{d}t) + x_t\mathrm{d}M_{st}. \tag{44}$$

Following Svensson (1989) and Duffie and Epstein (1992a,b) among others we have that the Hamilton-Jacobi-Bellman equation associated with the latter problem is

$$0 = \max_{(p,x,k)} \mathcal{D}_N^{(p,x,k)} G(N) + f(c, G(N)) - \frac{\gamma(pN\sigma_S G_N(N))^2}{2G(N)} - \lambda_{s0} G(N) \left(u(1;\gamma_s) - u\left(b(x,N);\gamma_s\right)\right) - \nu_{m0} G(N)$$

subject to the transversality conditions

$$\lim_{t\to\infty} E[e^{-\nu_{m0}t} G(N_{0t})] = \lim_{t\to\infty} E^Q[e^{-rt}N_{0t}] = 0 \tag{45}$$

where N_0 denotes the path of the process N under the optimal strategy, the second order differential operator

$$\mathcal{D}_N^{(p,x,k)} = ((pN\sigma_S)^2/2)\partial_{NN} + (rN + pN\sigma_S\theta - k - x\lambda_{s0}(H))\partial_N$$

is the continuous part of the infinitesimal generator of the process N under the portfolio, insurance and consumption strategy (p, x, k) and we have set

$$b(x,N) = \frac{G(N+x)}{G(N)}.$$

The specification of the agent's preferences and the dynamics of the controlled process in equation (44) imply that G is increasing and homogenous of degree one. Using these properties in conjunction with the HJB equation, we obtain that the value function and the optimal strategy are given by $G(N) = \Theta N$ and

$$p_t^* = \theta/(\gamma\sigma_S),$$
$$x_t^* = 0,$$
$$k_t^* = \rho^\varepsilon \Theta^{1-\varepsilon} N_{0t-},$$

for some nonnegative constant such that

$$\rho^{\varepsilon}\Theta^{1-\varepsilon} = \varepsilon\rho + (1-\varepsilon)\left(r - \nu_{0m} + \theta^2/(2\gamma)\right)$$

This equation admits a well-defined solution if and only if the constant A of equation (19) is strictly positive. In this case, $\Theta = \rho(A/\rho)^{\frac{1}{1-\varepsilon}}$ and substituting this into the definition of the optimal consumption plan we conclude that

$$\begin{aligned} c_{0t} &= AN_{0t-}, \\ \pi_{0t} &= (\theta/(\gamma\sigma_S))N_{0t-}, \\ x_{0t} &= \phi P_0(H_{t-}) \end{aligned}$$

as required. To complete the proof we need to show that under condition (19) the above solution satisfies (45). Using equation (44) and the definition of the candidate optimal strategy we obtain that the agent's disposable wealth evolves according to

$$\begin{aligned} \mathrm{d}N_{0t} &= N_{0t}(r-A)\mathrm{d}t + N_{0t}(\theta/\gamma)(\mathrm{d}Z_t + \theta\mathrm{d}t) \\ &= N_{0t}(r-A)\mathrm{d}t + N_{0t}(\theta/\gamma)\mathrm{d}\hat{Z}_t \end{aligned}$$

where $\hat{Z}$ is a risk neutral Brownian motion. Combining this expression with well-known results on the expectation of the geometric Brownian motion gives

$$\begin{aligned} E^Q[e^{-rt}N_{0t}] &= e^{-At}N_{00}, \\ E_0[e^{-\nu_{0m}t}G(N_{0t})] &= e^{(r-\nu_{m0}-A+\theta^2/\gamma)t}\Theta N_{00}, \end{aligned}$$

and it follows that condition (19) is necessary and sufficient for both the feasibility of c_0 and the validity of the transversality conditions. ■

Proof of Theorem 2. The Hamilton-Jacobi-Bellman equation associated with the agent's optimization problem is given by equation (24) subject to

$$\lim_{t\to\infty} E^Q[e^{-rt}W_t^*] = 0, \tag{46}$$

and

$$\lim_{t\to\infty} E\left[e^{-\int_0^t \nu_m(H^*_{\tau-})\mathrm{d}\tau} V(W^*_\tau, H^*_\tau)\right] = 0, \tag{47}$$

where the processes (W^*, H^*) denote the agent's wealth and health status under the optimal strategy. Maximizing the HJB equation gives the candidate optimal strategy of equations (25)–(28) and substituting these back into equation (24) shows that the HJB equation can be written as

$$\begin{aligned}\nu_m(H)V = \mathcal{D}^*V + f(c, V) - \frac{\gamma\theta^2 V V_W^4}{2(\gamma V_W^2 - V V_{WW})^2} \\ - \lambda_s(H)(u(1;\gamma_s) - u(\kappa(x^*, W, H);\gamma_s))V(W, H)\end{aligned}$$

where

$$\begin{aligned}\mathcal{D}^* = ((I^*/H)^\alpha - \delta)H\partial_H + \frac{1}{2}(\pi^*\sigma_S)^2\partial_{WW} \\ + (rW + \pi^*\sigma_S\theta + \beta H - c^* - I^* - x^*\lambda_s(H))\partial_W\end{aligned}$$

is the continuous part of the differential operator associated to the process (H, W) under the candidate optimal strategy, and x^* is implicitly defined by

$$\kappa(x^*, W, H)^{-\gamma_s} = \frac{V_W(W + x^*, (1-\phi)H)}{V_W(W, H)}. \tag{48}$$

Consider the first order approximations given by

$$V(W, H) \approx V_1(W, H) = V_0(W, H) + \epsilon V_\epsilon(W, H) \tag{49}$$

and

$$x^*(W, H) \approx x_1(W, H) = x_0(W, H) + \epsilon x_\epsilon(H, W) \tag{50}$$

where V_0 is the value function for the case of health independent mortality and morbidity, and the unknown functions

$$(x_\epsilon, V_\epsilon)(W,H) = \left(x^{(1)}, V^{(1)}\right)(W,H) = \frac{\partial(x^*, V)}{\partial \epsilon}(W,H)\bigg|_{\epsilon=0}$$

are the first order corrections induced by the presence of health-dependent mortality and morbidity. Substituting these approximation into equation (48) and expanding the resulting expression to the first order in ϵ shows that the first order correction to the optimal insurance coverage is given by

$$x_\epsilon = \frac{1}{\Theta}\left(V_\epsilon - V_\epsilon(W + x_0, H(1-\phi)) + \frac{N_0}{\gamma_s}(V_{\epsilon W}(W + x_0, H(1-\phi)) - V_{\epsilon W})\right).$$

On the other hand, substituting the approximations (49), (50) into the HJB equation and expanding the result to the first order in ϵ shows that the first order correction to the value function solves

$$\begin{aligned}\nu_{m0} V_\epsilon = \mathcal{D}^0 V_\epsilon + f_V(c_0, V_0)V_\epsilon + \frac{\theta^2}{2\gamma}(V_\epsilon - 2N_0(W,H)V_{\epsilon W}) - \bar{\nu}_{m1} H^{-\xi_m} V_\epsilon \\ - \bar{\lambda}_{s1}\Theta(\eta - \lambda_{s0})\phi B H^{1-\xi_s} + \lambda_{s0}(V_\epsilon(W + x_0, H(1-\phi)) - V_\epsilon)\end{aligned} \tag{51}$$

where

$$\begin{aligned}\mathcal{D}^0 = ((I_0/H)^\alpha - \delta)H\partial_H + \frac{1}{2}(\pi_0\sigma_S)^2\partial_{WW} \\ + (rW + \pi_0\sigma_S\theta + \beta H - c_0 - I_0 - x_0\lambda_{s0})\partial_W\end{aligned}$$

is the continuous part of the differential operator associated to the optimal strategy of the health-independent intensity case, and we have set

$$\bar{\nu}_{m1} = \frac{\bar{\lambda}_{m1}}{1-\gamma_m}.$$

Similarly, substituting the approximations (49), (50) into equations (25)–(27) and expanding the resulting expressions shows that up to a first order approximation

$$\pi^* = \pi_0 + \frac{\epsilon\theta}{\gamma^2\sigma_S\Theta}(\gamma V_\epsilon + N_0(V_{\epsilon WW}N_0 - \gamma V_{\epsilon W})) \tag{52}$$

$$c^* = c_0 + \epsilon(\rho/\Theta)^\varepsilon(V_\epsilon - \varepsilon V_{\epsilon W}N_0), \tag{53}$$

$$I^* = I_0 + \frac{\epsilon}{(1-1/\alpha)\Theta}I_0^\alpha H^{1-\alpha}(BV_{\epsilon W} - V_{\epsilon H}) \tag{54}$$

where the functions π_0, c_0 and I_0 are defined as in Proposition 1 and Theorem 1. An educated guess suggests that the first order correction to the agent's value function should be of the form

$$V_\epsilon(W,H) = C_{m1}N_0(W,H)H^{-\xi_m} + C_{s1}P_0(H)H^{-\xi_s}$$

for some constants C_{m1}, C_{s1}. Substituting this ansatz into equation (51), matching terms and solving for the constants shows that

$$C_{k1} = \Theta\bar{L}_k(1), \qquad k = m, s,$$

where we have set $\bar{L}_k = L_k/\epsilon$ and the functions L_k are defined as in the statement. Using these constants together with equations (52), (53), (54) then gives the approximate optimal policy reported in the statement and it only remains to show that a suitable approximation of the transversality conditions is satisfied.

Consider first the transversality condition for the value function in equation (47) and expand the quantity inside the expectation to the first order in ϵ. This gives

$$\begin{aligned} e^{-\int_0^t \nu_m(H^*_{s-})ds}V(W^*_t, H^*_t) &\approx e^{-\nu_{m0}t}\Theta N_{0t} \\ &+ \epsilon e^{-\nu_{m0}t}\left(V_\epsilon(W_{0t}, H_{0t}) + \Theta\nabla N_{0t} + \bar{\nu}_{m1}\Theta N_{0t}\int_0^t H_{0\tau-}^{-\xi_m}\mathrm{d}\tau\right) \end{aligned} \tag{55}$$

where the processes (W_0, H_0) denote the agent's wealth and health under the optimal strategy of the benchmark case in which $\epsilon = 0$, and

$$\nabla N_{0t} = \lim_{\epsilon\to 0}\frac{N_0(W^*_t, H^*_t) - N_0(W_{0t}, H_{0t})}{\epsilon}$$

denotes the derivative of the process $N_0(W_t^*, H_t^*)$ with respect to ϵ at the origin. Using the definition of the approximate optimal strategy in conjunction with straightforward (but lengthy) algebra it can be shown that

$$\begin{aligned}
\mathrm{d}\nabla N_{0t} &= \nabla N_{0t}\frac{\mathrm{d}N_{0t}}{N_{0t}} - (A - r + F(1-\xi_s))\bar{L}_s(H_{0t-})P_0(H_{0t-})\mathrm{d}t \\
&\quad - A(1-\varepsilon)\bar{L}_m(H_{0t-})N_{0t}\mathrm{d}t + (\theta/\gamma)\bar{L}_s(H_{0t-})P_0(H_{0t-})(\mathrm{d}Z_t + \theta\mathrm{d}t) + \mathrm{d}M_t
\end{aligned}$$

for some discontinuous martingale M with initial value equal to zero. Integrating this equation and using the fact that

$$\frac{\mathrm{d}N_{0t}}{N_{0t}} = (r - A)\mathrm{d}t + (\theta/\gamma)(\mathrm{d}Z_t + \theta\mathrm{d}t) \tag{56}$$

we find that

$$\nabla N_{0t} = N_{0t}\left(\int_0^t \frac{\bar{L}_s(H_{0\tau-})P_0(H_{0\tau-})}{N_{0\tau}}(C_1\mathrm{d}\tau + C_2\mathrm{d}Z_\tau) - \int_0^t C_3\bar{L}_m(H_{0\tau-})\mathrm{d}\tau + \hat{M}_t\right)$$

where $\hat{M}$ is a discontinuous martingale with initial value equal to zero and

$$\begin{aligned}
C_1 &= r - A - F(1-\xi_s) + (\theta^2/\gamma)(1 - 1/\gamma), \\
C_2 &= \theta/\gamma, \\
C_3 &= (1-\varepsilon)A.
\end{aligned} \tag{57}$$

Taking expectations on both sides and using equation (56) together with basic properties of Poisson processes, the definition of F and the fact that

$$E\left[N_{0t}\int_0^t (X_\tau\mathrm{d}\tau + Y_\tau dZ_\tau)\right] = E\int_0^t e^{(r-A+\theta^2/\gamma)(t-\tau)}N_{0\tau}(X_\tau + Y_\tau(\theta/\gamma))\mathrm{d}\tau$$

for any sufficiently integrable predictable processes, we obtain

$$\begin{aligned}
e^{-\nu_{m0}t}E[\nabla N_{0t}] &= \frac{C_3N_{00}\bar{L}_m(H_{00})}{F(-\xi_m)}e^{(r-\nu_{m0}-A+\theta^2/\gamma)t}\left(e^{F(-\xi_m)t} - 1\right) \\
&\quad + \bar{L}_s(H_{00})P_0(H_{00})\left(e^{(r-\nu_{m0}-A+\theta^2/\gamma)t} - e^{-(\nu_{m0}-F(1-\xi_s))t}\right).
\end{aligned}$$

Similarly, using the definition of the functions N_0 and V_ϵ together with equation (56) and basic properties of Poisson processes we obtain

$$\begin{aligned} e^{-\nu_{m0}t}E\left[N_{0t}\right] &= e^{(r-\nu_{m0}-A+\theta^2/\gamma)t}N_{00} \\ e^{-\nu_{m0}t}E\left[V_\epsilon(W_{0t},H_{0t})\right] &= e^{-(\nu_{m0}-F(1-\xi_s))t}\Theta\bar{L}_s(H_{00})P_0(H_{00}) \\ &\quad + e^{(r-\nu_{m0}-A+\theta^2/\gamma+F(-\xi_m))t}\Theta N_{00}\bar{L}_m(H_{00}) \\ e^{-\nu_{m0}t}E\left[N_{0t}\int_0^t H_{0\tau-}^{-\xi_m}\mathrm{d}\tau\right] &= \frac{N_{00}H_{00}^{-\xi_m}}{F(-\xi_m)}e^{(r-\nu_{m0}-A+\theta^2/\gamma)t}\left(e^{F(-\xi_m)t}-1\right) \end{aligned}$$

and it now follows from equation (55) that the transversality condition for the approximate value function holds if and only if

$$r-\nu_{m0}-A+\theta^2/\gamma<0, \tag{58}$$

$$r-\nu_{m0}-A+\theta^2/\gamma+F(-\xi_m)<0, \tag{59}$$

$$F(1-\xi_s)-\nu_{m0}<0. \tag{60}$$

Let us now turn to the agent's wealth. To verify that an approximate version of the transversality condition (46) holds we start by observing that

$$W_t^* = N_0(W_t^*,H_t^*)-P_0(H_t^*). \tag{61}$$

Expanding both sides of this identity as ϵ approaches zero shows that up to a first order approximation the agent's optimal wealth is given by

$$W_t^* \approx N_{0t}-P_0(H_{0t})+\epsilon(\nabla N_{0t}-B\nabla H_{0t})$$

where the process defined by

$$\nabla H_{0t}=\lim_{\epsilon\to 0}\left(\frac{H_t^*-H_{0t}}{\epsilon}\right)=H_{0t}\int_0^t\left(C_4\frac{N_{0\tau}\bar{L}_m(H_{0\tau-})}{H_{0\tau}}+C_5\frac{\bar{L}_s(H_{0\tau-})P_0(H_{0\tau-})}{H_{0\tau-}}\right)\mathrm{d}\tau$$

with

$$\begin{aligned} C_4 &= -\xi_m K/((1-\alpha)B), \\ C_5 &= -(\xi_m-1)K/((1-\alpha)B), \end{aligned}$$

represents the directional derivative of the agent's health process along the optimal strategy as $\epsilon \to 0$. Taking expectations under the risk neutral probability measure on both sides of equation (61) and using the fact that

$$\frac{\nabla N_{0t}}{N_{0t}} = \int_0^t \frac{\bar{L}_s(H_{0\tau-})P_0(H_{0\tau-})}{N_{0\tau}}((C_1 - \theta C_2)\mathrm{d}\tau + C_2 \mathrm{d}\hat{Z}_\tau) - \int_0^t C_3 \bar{L}_m(H_{0\tau-})\mathrm{d}\tau + \hat{M}_t$$

for some risk neutral Brownian motion $\hat{Z}$ and some discontinuous risk neutral martingale $\hat{M}$ with initial value zero together with the same arguments as above we deduce that

$$\begin{aligned}
E^Q[e^{-rt}W_{0t}] &= E^Q[e^{-rt}(N_{0t} - BH_{0t})] = e^{-At}N_{00} - e^{-(r-F(1))t}BH_{00}, \\
E^Q[e^{-rt}\nabla N_{0t}] &= \bar{L}_s(H_{00})P_0(H_{00})\left(e^{-At} - e^{-(r-F(1-\xi_s))t}\right) \\
&\qquad - C_3 \frac{N_{00}\bar{L}_m(H_{00})}{F(-\xi_m)}\left(e^{-(A-F(-\xi_m))t} - e^{-At}\right), \\
E^Q[e^{-rt}\nabla H_{0t}] &= C_4' N_{00}\bar{L}_m(H_{00})\left(e^{-(A-F(-\xi_m))t} - e^{-(r-F(1))t}\right) \\
&\qquad + C_5' \bar{L}_s(H_{00})P_0(H_{00})\left(e^{-(r-F(1-\xi_s))t} - e^{-(r-F(1))t}\right)
\end{aligned}$$

for some constants C_4', C_5' and it follows that the approximate transversality condition for wealth holds if and only if

$$A > 0, \tag{62}$$

$$r - F(1) > 0, \tag{63}$$

$$r - F(1 - \xi_s) > 0, \tag{64}$$

$$A - F(-\xi_m) > 0. \tag{65}$$

Combining the restrictions (58), (59), (60), (62), (63), (64), (65) with those imposed in Theorem 1 produces the restrictions of the statement and completes the proof. ■

Proof of Proposition 2. Under the conditions of the statement we have that the agent's total is given by

$$N_t^* = W_t^* + E_t \int_t^\infty m_{t,\tau}(\beta H_{\tau-}^* - I_\tau^*)\mathrm{d}\tau = E_t^Q \int_t^\infty e^{-r(\tau-t)} c_\tau^* \mathrm{d}\tau.$$

Expanding both sides of the above expression as ϵ approaches zero we find that up to a first order approximation the agent's total wealth is given by

$$N_t^* \approx N_{0t} + \epsilon A \int_t^\infty e^{-r(\tau-t)} E_t^Q[\nabla N_{0\tau}]\mathrm{d}\tau$$
$$+ \epsilon A \int_t^\infty e^{-r(\tau-t)} E_t^Q[(1-\varepsilon)N_{0\tau}\bar{L}_m(H_{0\tau-}) + \bar{L}_s(H_{0\tau-})P_0(H_{0\tau-})]\mathrm{d}\tau. \quad (66)$$

Using the same arguments as above we obtain

$$E_t^Q[e^{-r(\tau-t)}\nabla N_{0\tau}] = \bar{L}_s(H_{0t})P_0(H_{0t})\left(e^{-A(\tau-t)} - e^{-(r-F(1-\xi_s))(\tau-t)}\right)$$
$$- C_3 \frac{N_{0t}\bar{L}_m(H_{0t})}{F(-\xi_m)}\left(e^{-(A-F(-\xi_m))(\tau-t)} - e^{-A(\tau-t)}\right)$$

and

$$E_t^Q[e^{-r(\tau-t)}N_{0\tau}\bar{L}_m(H_{0\tau-})] = e^{-(A-F(-\xi_m))(\tau-t)}N_{0t}\bar{L}_m(H_{0t}),$$
$$E_t^Q[e^{-r(\tau-t)}\bar{L}_s(H_{0\tau-})P_0(H_{0\tau-})] = e^{-(r-F(1-\xi_s))(\tau-t)}N_{0t}\bar{L}_s(H_{0t})P_0(H_{0t})$$

where the constant C_3 is defined as in (57). Plugging these expressions into equation (66) and computing the resulting integrals then shows that the agent's total wealth satisfies

$$N_t^* \approx N_{0t} + \epsilon\bar{L}_s(H_{0t})P_0(L_s H_{0t}) = N_{0t} + L_s(H_{0t})P_0(H_{0t})$$

and completes the proof. ∎

Proof of Proposition 3. Consider an agent with wealth W, health H and intensity parameters $(\lambda_{m0}, \lambda_{s0}, \epsilon\lambda_{m1}, \epsilon\lambda_{s1})$ and denote by

$$w_h(\epsilon) = w_h(n, W, H, \epsilon)$$

the value to this agent of n additional units of health. Expanding equation (37) to the first order as ϵ decreases to zero and using the definition of V_0 we obtain

$$0 \approx \Theta(nB - w_h(0)) + \epsilon\Theta(V_\epsilon(W - w_h(0), H+n) - V_\epsilon(W,H) - w_h'(0)).$$

Setting both terms on the right to zero and using the definition of V_ϵ shows that up to a first order approximation the value of n additional units of heath is

$$w_h(n, W, H, \epsilon) \approx nB + J_m(n, W, H)N_0(W, H) + J_s(n, W, H)$$

and completes the proof. ■

Proof of Proposition 4. Consider first the computation of the expected lifetime. Using basic properties of point processes we have that

$$\ell(W, H) = E \int_0^\infty e^{-\int_0^\tau \lambda_m(H^*_{s-})\mathrm{d}s} \mathrm{d}\tau$$

where H^* denotes the agent's health along the optimal path. Expanding both sides of the above expression to the first order as ϵ approaches zero gives

$$\ell(W, H) \approx E \int_0^\infty e^{-\lambda_{m0}\tau} \left(1 + \epsilon\bar{\lambda}_{m1} \int_0^\tau H_{0s}^{-\xi_m} \mathrm{d}s\right) \mathrm{d}\tau = \frac{1 - \kappa_0 \lambda_{m1} H^{-\xi_m}}{\lambda_{m0}}$$

where H_0 denotes the agent's health under the optimal strategy of the benchmark case with health independent intensities and the second equality follows from the assumptions of the statement and basic properties of Poisson processes.

Let us now turn to the computation of the value of life. Consider an agent with intensity parameters $(\lambda_{m0}, \lambda_{s0}, \epsilon\lambda_{m1}, \epsilon\lambda_{s1})$, denote by

$$w_\ell(\epsilon) = w_\ell(n, W, H, \epsilon)$$

the value to this agent of n units of additional life expectancy and by

$$\lambda^*_{m0}(\epsilon) = \lambda^*_{m0}(n, W, H, \epsilon)$$

the solution to equation (39). Expanding equations (39) and (40) to the first order and using the approximation of the life expectancy derived in the first part gives

$$\begin{aligned} 0 &\approx n + 1/\lambda_{m0} - 1/\lambda^*_{m0}(0) - \epsilon \\ &- \epsilon \left(\frac{\mathrm{d}}{\mathrm{d}\epsilon}(1/\lambda^*_{m0}(\epsilon))\bigg|_{\epsilon=0} + \frac{\kappa_0 \bar{\lambda}_{m1} H^{-\xi_m}}{\lambda_{m0}} - \frac{\bar{\lambda}_{m1} H^{-\xi_m}}{\lambda^*_{m0}(0)(\lambda^*_{m0}(0) - F(-\xi_m))} \right), \end{aligned}$$

and

$$0 \approx \Theta^*(W - w_l(0) + P_0(H)) - V_0(W,H) - \epsilon\left(V_\epsilon(W,H) + \Theta^* w_l'(0)\right)$$
$$+ \epsilon \left(\frac{\mathrm{d}\Theta^*}{\mathrm{d}\lambda_{m0}^*(0)} \frac{\mathrm{d}\lambda_{m0}^*(\epsilon)}{\mathrm{d}\epsilon}\bigg|_{\epsilon=0} + \Theta^*(W - w_l(0) + P_0(H))\bar{L}_m(H) + \Theta^* \bar{L}_s(H) P_0(H) \right),$$

where we have set

$$\Theta^* = \rho^{\frac{\varepsilon}{1-\varepsilon}} A(n)^{\frac{1}{1-\varepsilon}} = \rho^{\frac{\varepsilon}{1-\varepsilon}} \left[\varepsilon\rho + (1-\varepsilon)\left(r - \frac{\lambda_{m0}^*(0)}{1-\gamma_m} + \frac{\theta^2}{2\gamma} \right) \right]^{\frac{1}{1-\varepsilon}}.$$

Setting the terms on the right to zero and using the definition of V_ϵ allows to solve for the unknowns $w_\ell(0)$, $w_\ell'(0)$, $\lambda_{m0}^*(0)$, $(\lambda_{m0}^*)'(0)$ and simplifying the resulting expansion of the value life gives the formula reported in the statement. ■

B Age dependent parameters

In this appendix we briefly discuss a generalization of the model in which the intensity parameters λ_{m0}, λ_{s_0}, $\bar{\lambda}_{m1}$, $\bar{\lambda}_{s1}$, η, the depreciation rate of health δ, the fraction of health ϕ that is lost upon experiencing a health shock, and the health sensitivity β of labor income are allowed to vary with the agent's age.

The difference between such a model and the one we considered in the text is that instead of depending only on wealth and health the value function and optimal strategy now also depends on the agent's age. Despite this added dependence the model can still be solved using an first order approximation but the functions P_0, L_m and L_s will now be age and health-dependent rather than just health-dependent. In particular, the analog of Theorem 1 is given by:

Theorem 3 *Let* $\lambda_{m1} = \lambda_{s1} = 0$*, define*

$$\nu_{m0}(t) = \lambda_{m0}(t)/(1-\gamma_m)$$

and assume that there exist strictly positive solutions A and B to the ordinary differential equations

$$A'(t) = A(t)^2 - \left(\varepsilon\rho + (1-\varepsilon)\left(r - \nu_{m0}(t) + \theta^2/(2\gamma)\right)\right) A(t), \tag{67}$$

$$B'(t) = (r + \delta(t) + \phi(t)\lambda_{s0}(t))B(t) + (1 - 1/\alpha)(\alpha B(t))^{\frac{1}{1-\alpha}} - \beta. \tag{68}$$

such that

$$\lim_{t\to\infty}(r - \nu_0(t) + \theta^2/(2\gamma) - A(t)) < 0, \tag{69}$$

$$\lim_{t\to\infty}((\alpha B(t))^{\frac{\alpha}{a-\alpha}} - r - \delta(t) - \phi(t)\lambda_{s0}(t)) < 0. \tag{70}$$

Then the indirect utility function of an alive agent is

$$V_0(t, W, H) = \Theta(t)N_0(t, W, H) = \Theta(t)\left(W + B(t)H + \frac{y_0 - a}{r}\right),$$

and generates the optimal consumption, portfolio, health insurance and health investment strategies given by

$$c_{0t} = a + A(t)N_0(t, W_{t-}, H_{t-}),$$
$$\pi_{0t} = (\theta/(\gamma\sigma_S))N_0(t, W_{t-}, H_{t-}),$$
$$x_{0t} = \phi(t)B(t)H_{t-},$$
$$I_{0t} = (\alpha B(t))^{\frac{1}{1-\alpha}}H_{t-}$$

with $\Theta(t) = \rho^{\frac{\varepsilon}{1-\varepsilon}}A(t)^{\frac{1}{1-\varepsilon}}$.

Proof. The proof is similar to that of Theorem 1 and therefore is omitted. ■

Theorem 4 *Let*

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

assume that there exist strictly positive solutions A, B to the ordinary differential equations (67), (68) such that (69), (70) and

$$\lim_{t\to\infty}(F(t,1-\xi_s)-\min(r,\nu_{0m}(t)))<0,$$
$$\lim_{t\to\infty}(F(t,-\xi_m)-\max(0,r-\nu_0(t)+\theta^2/(2\gamma))-A(t))<0,$$

hold true and define

$$\mathcal{L}_m(t,H)=\int_t^\infty e^{-\int_t^\tau (A(s)-F(s,-\xi_m))ds}\lambda_{m1}(\tau)H^{-\xi_m}/(\gamma_m-1)d\tau,$$
$$\mathcal{L}_s(t,H)=\int_t^\infty e^{-\int_t^\tau (r-F(s,1-\xi_s))ds}\lambda_{s1}(\tau)\phi(\tau)B(\tau)(\lambda_{s0}(\tau)-\eta(\tau))H^{1-\xi_s}.$$

Up to a first order approximation the indirect utility of an alive agent is

$$V_1(t,W,H)=V_0(t,W,H)+\Theta(t)\mathcal{L}_m(t,H)N_0(t,W,H)+\Theta(t)\mathcal{L}_s(t,H)$$

and generates the approximate optimal consumption, portfolio, health insurance and health investment strategy given by

$$\begin{aligned}
c_{1t}&=c_{0t}+A(t)(1-\varepsilon)\mathcal{L}_m(t,H_{t-})N_0(t,W_{t-},H_{t-})+A(t)\mathcal{L}_s(t,H_{t-}),\\
\pi_{1t}&=\pi_{0t}+(\theta/(\gamma\sigma_S))\mathcal{L}_s(t,H_{t-}),\\
x_{1t}&=x_{0t}+\chi(t,\xi_m)(1-1/\gamma_s)\mathcal{L}_m(t,H_{t-})N_0(t,W_{t-},H_{t-})\\
&\qquad+\chi(t,\xi_s-1)\mathcal{L}_s(t,H_{t-})\\
I_{1t}&=I_{0t}-(\xi_m K(t)/(1-\alpha))\mathcal{L}_m(t,H_{t-})N_0(t,W_{t-},H_{t-})\\
&\qquad-((\xi_s-1)K(t)/(1-\alpha))\mathcal{L}_s(t,H_{t-})
\end{aligned}$$

with $K(t)=\alpha^{\frac{1}{1-\alpha}}B(t)^{\frac{\alpha}{1-\alpha}}$.

Proof. The proof is similar to that of Theorem 2 and therefore is omitted. ■

C Data

We rely on a sample of 30'961 individuals obtained by pooling the 1999, 2001, 2003, 2005, and 2007 waves of the Panel Study of Income Dynamics (PSID, http://psidonline.isr.umich.edu/).

All nominal variables correspond to per-capita values (i.e., household values divided by household size) scaled by 10^{-4}. The explanatory variables used in the estimation of the model are the agents' wealth and health which are constructed from the PSID data as follows:

Health We associate values of 1.5 (poor health), 2.0 (fair), 2.5 (good), 3.0 (very good) and 3.5 (excellent health) to the self-reported health variable corresponding to the household head.

Wealth We use financial wealth defined as risky plus riskless assets. Risky assets are stocks in publicly held corporations, mutual funds, investment trusts, private annuities, IRA's or pension plans. Riskless assets are checking accounts plus bonds plus remaining IRA's and pension assets.

The observed portfolios, consumption, health expenditure and health insurance used in the estimation are constructed from the PSID data as follows:

Portfolio Value of financial wealth held in risky assets.

Consumption The consumption measure that we rely on is inferred from the food, utility and transportation expenditures available in PSID, using the Skinner (1987) method with the updated shares of Guo (2010).

Health expenditures Total out-of-pocket expenditures paid by household on hospital, nursing home, doctor, outpatient surgery, dental expenditures, prescriptions in-home medical care.

Health insurance Total amount paid for health insurance premium.